SIAM J. FINANCIAL MATH.
Vol. 14, No. 4, pp. 959–1003

Cubature Method for Stochastic Volterra Integral Equations*

Qi Feng† and Jianfeng Zhang‡

Abstract. In this paper, we introduce the cubature formula for stochastic Volterra integral equations. We first derive the stochastic Taylor expansion in this setting, by utilizing a functional Itô formula, and provide its tail estimates. We then introduce the cubature measure for such equations, and construct it explicitly in some special cases, including a long memory stochastic volatility model. We shall provide the error estimate rigorously. Our numerical examples show that the cubature method is much more efficient than the Euler scheme, provided certain conditions are satisfied.

Key words. stochastic Volterra integral equations, cubature formula, stochastic Taylor expansion, fractional stochastic volatility model, rough volatility model

MSC codes. 60H20, 65C30, 91G60

DOI. 10.1137/22M146889X

1. Introduction.

Consider a stock price in a Brownian setting under risk neutral measure $\mathbb{P}$:

$$dS_t = S_t \sigma_t dB_t. \tag{1.1}$$

In the Black–Scholes model, the volatility process $\sigma_t \equiv \sigma_0$ is a constant. There is a large literature on stochastic volatility models where σ is also a diffusion process; see, e.g., Fouque et al. [19]. Strongly supported by empirical studies, the fractional stochastic volatility models and rough volatility models have received very strong attention in recent years, where σ satisfies the following stochastic Volterra integral equation (SVIE):

$$\sigma_t = \sigma_0 + \int_0^t K(t,r)V_0(\sigma_r)dr + \int_0^t K(t,r)V_1(\sigma_r)d\tilde{B}_r. \tag{1.2}$$

Here $\tilde{B}$ is another Brownian motion possibly correlated with B, V_i's are appropriate deterministic functions, and the deterministic two time variable function K has a Hurst parameter $H>0$, in the sense that $K(t,r) \sim (t-r)^{H-\frac{1}{2}}$ and $\partial_t K(t,r) \sim (t-r)^{H-\frac{3}{2}}$ when $t-r>0$ is small. Such a model was first proposed by Comte and Renault [10] for $H>\frac{1}{2}$ to model the long memory property of the volatility process. Another notable work is Gatheral, Jaisson, and Rosenbaum [23], which finds market evidence that volatility's high-frequency behavior

*Received by the editors January 4, 2022; accepted for publication (in revised form) July 5, 2023; published electronically October 10, 2023.
https://doi.org/10.1137/22M146889X
Funding: The first author's research was supported in part by NSF grant DMS-2306769. The second author's research was supported in part by NSF grants DMS-1908665 and DMS-2205972.
†Department of Mathematics, Florida State University, Tallahassee, FL 32306 USA (qfeng2@fsu.edu).
‡Department of Mathematics, University of Southern California, Los Angeles, CA 90089 USA (jianfenz@usc.edu).

could be modeled as a rough path with $H < \frac{1}{2}$. We remark that one special case of (1.2) is the fractional Brownian motion, where $V_0 \equiv 0$ and $V_1 \equiv 1$; see, e.g. Nualart [44].

Our goal in this paper is to understand and more importantly to numerically compute the option price in this market: assuming zero interest rate for simplicity,

$$\mathbb{E}[G(S_T)]. \tag{1.3}$$

Note that the volatility process σ in (1.2) is in general neither a Markov process nor a semimartingale.[1] Consequently, S_t is highly non-Markovian, in the sense that one cannot Markovianize it by adding finitely many extra states, and correspondingly the option price is characterized as a path dependent PDE (PPDE, for short); see Viens and Zhang [51]. This imposes significant challenges, both theoretically and numerically. Indeed, compared to the huge literature on numerical methods for PDEs, there are very few works on efficient numerical methods for such PPDEs. Besides the standard Euler scheme (see Zhang [55]), we refer the reader to Wen and Zhang [53] for an improved rectangular method; Jacquier and Oumgari [31] and Ruan and Zhang [50] on numerical methods for high dimensional (nonlinear) PPDEs driven by SVIEs; Richard, Tan, and Yang [48, 49] on discrete-time simulation schemes, including the Euler and Milstein schemes, and the corresponding multilevel Monte Carlo method; Ma, Yang, and Cui [41] by using Markov chain approximation; and Alfonsi and Kebaier [1], Bayer and Breneis [3], and Harms [29] by using the Laplace transform for singular kernel functions. In recent years, there has also been a growing interest on the convergence analysis and error estimates for SVIEs; see, e.g., Bayer, Fukasawa, and Nakahara [4], Bayer, Hall, and Tempone [5], Bonesini, Jacquier, and Pannier [8], Friz, Salkeld, and Wagenhofer [20], Fukasawa and Ugai [21], Gassiat [22], Li, Huang, and Hu [35], and Nualart and Saikia [45]. In this paper, we propose the cubature method for the above option price (1.3). This is a deterministic method, and our numerical examples show that, under certain conditions, it is much more efficient than the simulation methods such as the Euler scheme.

The cubature method was first introduced by the seminal works Lyons and Victoir [40] and Litterer and Lyons [37] for diffusion processes; see also Gyurkó and Lyons [26], Litterer and Lyons [38], Ninomiya and Shinozaki [42], and Ninomiya and Victoir [43] for its extensive numerical implementations. The method builds upon the stochastic Taylor expansion for smooth G:

$$G(S_T) = I_N + R_N; \tag{1.4}$$

see, e.g., Kloeden and Platen [34], where I_N is a linear combination of multiple integrals against the Brownian motion B (typically in Stratonovich form), called the signatures of B, and R_N is the remainder term. The main idea is to introduce a discrete measure Q to match the expectations of the signatures: recalling that $\mathbb{E} = \mathbb{E}^{\mathbb{P}}$ is the expectation under $\mathbb{P}$,

$$\mathbb{E}[I_N] = \mathbb{E}^Q[I_N]. \tag{1.5}$$

Then we will have an approximation $\mathbb{E}[G(S_T)] \approx \mathbb{E}^Q[G(S_T)]$. Since Q is discrete and it is easy to compute the exact value of $\mathbb{E}^Q[G(S_T)]$ (without involving simulations), the algorithm is

[1]When $H > 1$, X is actually a semimartingale; see, e.g., [51]. However, it is still highly non-Markovian, so the numerical challenge remains in this case.

very efficient, provided sufficient technical conditions to make the approximation error small enough.

In this paper, we shall consider the general SVIE (see (2.1) below), and our goal is to approximate $\mathbb{E}[G(X_T)]$. To introduce the cubature method for X, our first step is to derive the stochastic Taylor expansion in this setting. Note that (1.4) relies heavily on the Itô formula, but the solution X to the SVIE is not a semimartinagle, which prohibits us from applying the Itô formula directly. To overcome this difficulty, we utilize an auxiliary two time variable process Θ_t^s introduced by Wang [52] and Viens and Zhang [51]; see (2.6) below. This process satisfies $\Theta_t^t = X_t$ and enjoys the desired semimartingale property: for fixed s, the process $t \in [0,s] \mapsto \Theta_t^s$ is a semimartingale. In particular, [51] established a functional Itô formula, which enables us to derive the desired stochastic Taylor expansion with more involved signatures for the SVIEs than the diffusion case. We then introduce a discrete cubature measure Q for X, in the spirit of (1.4) and (1.5), and prove the following error estimate: for some constant C_N which depends on the regularity of the coefficients,

$$\left|\mathbb{E}[G(X_T)] - \mathbb{E}^Q[G(X_T)]\right| \le C_N T^{\frac{N+1}{2}}. \tag{1.6}$$

The above result is desirable only when T is small. For general T, we follow the idea of [40, Theorem 3.3] and utilize the flow property of the path dependent value function established in [51]. To be precise, we consider a uniform partition of $[0,T]$: $0 = T_0 < \cdots < T_M = T$ and construct a cubature measure Q_m on each subinterval $[T_m, T_{m+1}]$. Let Q be the independent composition of $\{Q_m\}_{0\le m<M}$; we then have the following estimate:

$$\left|\mathbb{E}[G(X_T)] - \mathbb{E}^Q[G(X_T)]\right| \le C_N \frac{T^{\frac{N+1}{2}}}{M^{\frac{N-1}{2}}}, \tag{1.7}$$

where C_N is independent of M. The above estimate clearly converges to 0 as $M \to \infty$.

We remark that, while our stochastic Taylor expansion can be developed for any kernel K with Hurst parameter $H > 0$, the cubature formula becomes much more subtle when $H < \frac{1}{2}$. In this paper, we restrict ourselves to the case $H \ge \frac{1}{2}$ and leave the case $H < \frac{1}{2}$ to future study. For applications, we refer the reader to Comte and Renault [10] for the long memory model with $\frac{1}{2} < H < 1$, Gulisashvili, Viens, and Zhang [24] for the integrated variance model with $1 < H < 2$, and El Omari [15] for the mixed fractional Brownian motion model with more general H. We also refer the reader to Beran [6, section 4.2] for applications in hydrology, Loussot et al. [39] for applications in image processing, Gupta, Singh, and Karlekar [25] for applications in signal classification, Blu and Unser [7] for fractional spline estimators, and Perrin et al. [47] for the theory of higher order fractional Brownian motions with general $H > 1$. However, we should point out that our result does not cover the rough volatility models in Gatheral, Jaisson, and Rosenbaum [23] with $0 < H < \frac{1}{2}$. Moreover, the estimates (1.6) and (1.7) require the coefficients to be sufficiently smooth, as we will specify in the paper. In particular, the constant C_N will depend on such regularity.

The efficiency of our cubature method comes down to the construction of the cubature measure Q in (1.7), which will involve $(2W)^M$ deterministic paths for some constant W. When the dimension of X is large, or when N is large, the W will be large; and when T is large, in light of (1.7) we will require M to be large. We refer the reader to section 7.5 for more precise comments on the efficiency issue. When all the conditions are satisfied so that $(2W)^M$ is at a reasonable level, our numerical examples show that the cubature method is much more efficient than the Euler scheme.

We should remark that the above efficiency issue was already present for the cubature method in the standard Brownian setting. There have been great efforts in the literature to overcome this difficulty and to apply the idea of the cubature method to more general models; see, e.g., Crisan and Manolarakis [11, 12], Crisan and McMurray [13], de Raynal and Trillos [14], Filipović, Larsson, and Pulido [16], Foster, Lyons, and Oberhauser [17], Foster, dos Reis, and Strange [18], and Hayakawa, Oberhauser, and Lyons [30]. It will be very interesting to explore whether these ideas can help to improve the cubature method in the Volterra framework. We would also like to mention the very interesting connection between the signature, the kernel method, and machine learning; see Chevyrev and Oberhauser [9], Kidger et al. [32], Király and Oberhauser [33], Liao et al. [36], and the references therein.

Finally, we note that, while sharing many properties, the SVIE (1.2) is different from the following SDE driven by a fractional Brownian motion $B^H_t := \int_0^t K(t,r)dr + \int_0^t K(t,r)d\tilde{B}_r$:

$$\sigma'_t = \sigma_0 + \int_0^t V_0(\sigma'_r)dr + \int_0^t V_1(\sigma'_r)dB^H_r. \tag{1.8}$$

We refer the reader to Baudoin and Coutin [2] and Passeggeri [46] for some works on signatures for fractional Brownian motions and Harang and Tindel (see [27, 28]) on signatures defined for " Volterra path." We shall remark that, unlike our signature, which is directly for the solution σ_t to the SVIE (1.2) (instead of for the driving Brownian motion $\tilde{B}$), these signatures are for the driving fractional Brownian motion B^H or "Volterra path," which has much simpler structure. In particular, their signatures do not lead to the desired stochastic Taylor expansion which is crucial for the cubature method.

The rest of the paper is organized as follows. In section 2, we derive the stochastic Taylor expansions for the general SVIEs and prove the tail estimate. In section 3, we introduce the cubature formula when T is small, and in section 4 we modify the cubature formula when $[0,T]$ is decomposed into M parts. We construct the cubature measure Q explicitly for a one dimensional SVIE in section 5 and for the two dimensional fractional stochastic volatility model in section 6. In section 7, we present various numerical examples and compare their efficiency with the Euler scheme. Finally, we present some technical proofs in the appendix.

2. Stochastic Taylor expansions.

Throughout this paper, let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a filtered probability space, let $B^0_t := t$, and let$B = (B^1, \dots, B^d)$ be a d-dimensional Brownian motion. Let $T > 0$ be a fixed terminal time. We consider d_1-dimensional state process $X = (X^1, \dots, X^{d_1})$ solving the following SVIE under Stratonovich integration $\circ$: given $x = (x_1, \dots, x_{d_1}) \in \mathbb{R}^{d_1}$,

$$X^i_t = x_i + \sum_{j=0}^{d} \int_0^t K_i(t,r) V^i_j(X_r) \circ dB^j_r, \quad i = 1, \dots, d_1, \tag{2.1}$$

and we are interested in the efficient numerical computation of

$$Y_0 := \mathbb{E}[G(X_T)]. \tag{2.2}$$

Throughout the paper, the following hypotheses will always be enforced: for some $N \geq 1$ which will be specified in the context, the following hold:

(H0) Each $K_i:\{(t,r):0\le r\le t\le T\}\to[0,\infty)$ is infinitely smooth on $\{r<t\}$, and either $K_i\equiv 1$ or K_i has Hurst parameter $H_i>\frac{1}{2}$, that is, $K_i(t,r)\sim(t-r)^{H_i-\frac{1}{2}}$ and $\partial_t K_i(t,r)\sim(t-r)^{H_i-\frac{3}{2}}$ when $t-r>0$ is small.

(HN) The functions $V^i_j, G\in C^N(\mathbb{R}^{d_1};\mathbb{R})$ with all the derivatives up to the order N bounded.

For later purposes, we will also need the following stronger version of **(H0)**:

(H0-N) Each $K_i:\{(t,r):0\le r\le t\le T\}\to[0,\infty)$ is infinitely smooth on $\{r<t\}$, and either $K_i\equiv 1$ or K_i has Hurst parameter $H_i>N+\frac{1}{2}$, in the sense that $\frac{\partial^{\alpha+\beta}}{\partial t^\alpha\partial r^\beta}K_i(t,r)\sim(t-r)^{H_i-\alpha-\beta-\frac{1}{2}}$ when $t-r>0$ is small for all integers $\alpha,\beta\ge 0$ such that $\alpha+\beta\le N$.

Remark 2.1. (i) When K_i has Hurst parameter $H_i>\frac{1}{2}$, X^i is Hölder-$(H_i\wedge 1-\varepsilon)$ continuous in t for any small $\varepsilon>0$. This implies that $V(X^i_t)\circ dB^j_t=V(X^i_t)dB^j_t$ for any smooth function V, where $V(X^i_t)dB^j_t$ denotes the Itô integral, and they coincide with the Young's pathwise integral. On the other hand, when $K_i\equiv 1$, X^i is clearly a semimartingale. So, letting I_0 denote the set of $i>0$ such that $K_i\equiv 1$, we may rewrite (2.1) in Itô's form, and in particular they are well-posed under **(H0)** and **(H2)**:

$$X^i_t=x_i+\sum_{j=0}^d\int_0^t K_i(t,r)V^i_j(X_r)dB^j_r+\frac{1}{2}\sum_{k\in I_0}\int_0^t K_i(t,r)\partial_k V^i_j(X_r)V^k_j(X_r)dr. \tag{2.3}$$

(ii) Consider a special case: $x_1=0$, $K_1\equiv 1$, $V^1_0\equiv 1$, $V^1_j\equiv 0$ for $j\ge 2$. Then we can easily see that $X^1_t=t$. So the system (2.1) actually covers the case that the coefficients V^i_j depend on the time variable t.

Remark 2.2. (i) In fractional stochastic volatility models, where X^i is interpreted as the volatility (or variance) process of a certain underlying asset price, the assumption $\frac{1}{2}<H_i<1$ implies that the volatility has "long memory"; see Comte and Renault [10]. We also refer the reader to [6, 7, 15, 24, 25, 39, 47] for applications and theory when $H_i>1$.

(ii) The case $H_i<\frac{1}{2}$, supported by the empirical studies in Gatheral, Jaisson, and Rosenbaum [23], has received very strong attention in the mathematical finance literature in recent years. The singularity of K_i in this case will make the theory much more involved; for example, one may need to consider the weak solution to (2.1), and consequently the numerical algorithms will be less efficient. We shall leave this important and challenging case to future study.

2.1. The functional Itô formula. Note that X^i is not a semimartingale when $K_i\ne 1$, which prohibits us from applying many stochastic analysis tools such as the Itô formula directly. To get around this difficulty, in this subsection we introduce a functional Itô formula, which is established in Viens and Zhang [51] but tailored for the purpose of this paper.

Denote $\mathbb{X}_t:=C^0([t,T];\mathbb{R}^{d_1})\cap C^1((t,T];\mathbb{R}^{d_1})$ for each $t\in[0,T]$, equipped with the uniform norm. For each $t\in[0,T]$ and $\phi:\mathbb{X}_t\to\mathbb{R}$, let $\partial_{\mathbf{x}}\phi$ denote the Fréchet derivative of u. That is, $\partial_{\mathbf{x}}\phi(\mathbf{x}):\mathbb{X}_t\to\mathbb{R}$ is a linear mapping satisfying

$$\phi(\mathbf{x}+\eta)-\phi(\mathbf{x})=\langle\partial_{\mathbf{x}}\phi(\mathbf{x}),\eta\rangle+o(\|\eta\|)\quad\forall\mathbf{x},\eta\in\mathbb{X}_t. \tag{2.4}$$

Similarly, we may define the second order derivate $\partial_{\mathbf{xx}}\phi(\mathbf{x})$ as a bilinear mapping on $\mathbb{X}_t\times\mathbb{X}_t$:

$$\langle\partial_{\mathbf{x}}\phi(\mathbf{x}+\eta_2),\eta_1\rangle-\langle\partial_{\mathbf{x}}\phi(\mathbf{x}),\eta_1\rangle=\langle\partial_{\mathbf{xx}}\phi(\mathbf{x}),(\eta_1,\eta_2)\rangle+o(\|\eta_2\|)\quad\forall\eta_1,\eta_2\in\mathbb{X}_t. \tag{2.5}$$

We may continue to define higher order derivatives $\partial_{\mathbf{x}}^{(n)}\phi(\mathbf{x})$ in an obvious manner and let $C^n(\mathbb{X}_t)$ denote the set of continuous functions $\phi:\mathbb{X}_t\to\mathbb{R}$ which has uniformly continuous derivatives up to order n. Moreover, as in [51] (see also an earlier work [52]), we introduce a two time variable process $\Theta_t^s=(\Theta_t^{1,s},\dots,\Theta_t^{d_1,s})$ for $0\le t\le s\le T$:

$$\Theta_t^{i,s}=x_i+\sum_{j=0}^{d}\int_0^t K_i(s,r)V_j^i(X_r)\circ dB_r^j. \tag{2.6}$$

This process enjoys the following nice properties:

- For fixed s, the process $t\in[0,s]\to\Theta_t^s$ is an $\mathbb{F}$-progressively measurable semimartingale.
- For fixed t, the process $s\in[t,T]\to\Theta_t^s$ is $\mathcal{F}_t$-measurable, continuous on $[t,T]$, infinitely smooth on $(t,T]$, and with "initial" condition $\Theta_t^t=X_t$. In particular, $\Theta_t\in\mathbb{X}_t$ a.s.

Then we have the following functional Itô formula, which is essentially the same as [51, Theorem 3.10] but in Stratonovich form instead of Itô form.

Proposition 2.3. *Let* **(H0)** *and* **(H2)** *hold, and let $\phi\in C^2(\mathbb{X}_{T'})$ for some $0<T'<T$. Then*

$$d\phi(\Theta_t^{[T',T]})=\sum_{i=1}^{d_1}\sum_{j=0}^{d}\langle\partial_{\mathbf{x}_i}\phi(\Theta_t^{[T',T]}),K_{i,t}^{[T',T]})\rangle V_j^i(X_t)\circ dB_t^j,\quad 0\le t\le T'. \tag{2.7}$$

Here $\Theta_t^{[T',T]}$ and $K_{i,t}^{[T',T]}$ denote the paths $\Theta_t^s, K_i(s,t)$, $s\in[T',T]$, respectively.

We now turn to the problem (2.2). For any $t\in[0,T]$ and $\theta\in\mathbb{X}_t$, introduce

$$u(t,\theta)=\mathbb{E}[G(X_T^{t,\theta})],\quad X_s^{t,\theta,i}=\theta_s^i+\sum_{j=0}^{d}\int_t^s K_i(s,r)V_j^i(X_r^{t,\theta})\circ dB_r^j,\quad i=1,\dots,d_1. \tag{2.8}$$

Since θ is differentiable, by Remark 2.1(i) it is clear that the above Volterra SDE is well-posed. Moreover,

$$u(T,\tilde{x})=G(\tilde{x})\ \forall\tilde{x}\in\mathbb{X}_T=\mathbb{R}^{d_1};\quad Y_0=u(0,x),\text{ where } x\in\mathbb{X}_0 \text{ is a constant path,}$$

and we have the following simple result, whose proof is postponed to the appendix.

Proposition 2.4. *Under* **(H0)** *and* **(HN)***, we have $u(t,\cdot)\in C^{N-1}(\mathbb{X}_t)$ for any $t\in[0,T]$. Moreover, all the involved derivatives are bounded by $C_Ne^{C_NT}$, where C_N depends only on the parameters in* **(H0)** *and* **(HN)***.*

2.2. The stochastic Taylor expansion. Fix $M\ge 1$, and set $T_m:=m\delta$, $m=0,\dots,M$, where $\delta:=\delta_M:=\frac{T}{M}$. In this subsection, we fix m and consider the stochastic Taylor expansion of $u(T_{m+1},\cdot)$ at T_m. We first introduce some notation: for any $n\ge 1$ and $s\in[T_m,T_{m+1}]$,

$$\mathbb{T}_n^m(s):=\{\vec{t}=(t_1,\dots,t_n):T_m\le t_n\le\cdots\le t_1\le s\},\quad t_0^m:=T_{m+1},\ \mathbb{T}_n^m:=\mathbb{T}_n(t_0^m). \tag{2.9}$$

Assume $u(T_{m+1},\cdot)$ is sufficiently smooth; by (2.7), we have

$$\begin{aligned} u(T_{m+1},\Theta_{T_{m+1}}^{[T_{m+1},T]})&=u(t_0^m,\Theta_{t_0^m}^{[t_0^m,T]})\\ &=u(t_0^m,\Theta_{T_m}^{[t_0^m,T]})+\sum_{i_1=1}^{d_1}\sum_{j_1=0}^{d}\int_{T_m}^{t_0^m}\langle\partial_{\mathbf{x}_{i_1}}u(t_0^m,\Theta_{t_1}^{[t_0^m,T]}),K_{i_1,t_1}^{[t_0^m,T]}\rangle V_{j_1}^{i_1}(\Theta_{t_1}^{t_1})\circ dB_{t_1}^{j_1}, \end{aligned} \tag{2.10}$$

where we used the fact that $X_{t_1}=\Theta^{t_1}_{t_1}$. Now fix $t_1\in[T_m,t_0^m]$, note that $t_2\in[T_m,t_1]\mapsto\Theta^{t_1}_{t_2}$ is a semimartingale, and note that $\theta\in\mathbb{X}_{t_0^m}\mapsto\langle\partial_{\mathbf{x}_{i_1}}u(t_0^m,\theta),K^{[t_0^m,T]}_{i_1,t_1}\rangle$ is in $C^2(\mathbb{X}_{t_0^m})$. Then

$$dV^{i_1}_{j_1}(\Theta^{t_1}_{t_2})=\sum_{i_2=1}^{d_1}\sum_{j_2=0}^{d}\partial_{x_{i_2}}V^{i_1}_{j_1}(\Theta^{t_1}_{t_2})K_{i_2}(t_1,t_2)V^{i_2}_{j_2}(X_{t_2})\circ dB^{j_2}_{t_2},$$

$$d\langle\partial_{\mathbf{x}_{i_1}}u(t_0^m,\Theta^{[t_0^m,T]}_{t_2}),K^{[t_0^m,T]}_{i_1,t_1}\rangle$$
$$=\sum_{i_2=1}^{d_1}\sum_{j_2=0}^{d}\langle\partial_{\mathbf{x}_{i_1}\mathbf{x}_{i_2}}u(t_0^m,\Theta^{[t_0^m,T]}_{t_2}),\,(K^{[t_0^m,T]}_{i_1,t_1},K^{[t_0^m,T]}_{i_2,t_2})\rangle V^{i_2}_{j_2}(\Theta^{t_2}_{t_2})\circ dB^{j_2}_{t_2}.$$

Applying Itô's formula and plugging these into (2.10), we obtain

(2.11)

$$\begin{aligned}u(t_0^m,\Theta^{[t_0^m,T]}_{t_0^m})&=u(t_0^m,\Theta^{[t_0^m,T]}_{T_m})+\sum_{i_1=1}^{d_1}\sum_{j_1=0}^{d}\int_{T_m}^{t_0^m}\langle\partial_{\mathbf{x}_{i_1}}u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),K^{[t_0^m,T]}_{i_1,t_1}\rangle V^{i_1}_{j_1}(\Theta^{t_1}_{T_m})\circ dB^{j_1}_{t_1}\\&+\sum_{i_1,i_2=1}^{d_1}\sum_{j_1,j_2=0}^{d}\int_{\mathbb{T}_2^m}\Big[\langle\partial_{\mathbf{x}_{i_1}\mathbf{x}_{i_2}}u(t_0^m,\Theta^{[t_0^m,T]}_{t_2}),(K^{[t_0^m,T]}_{i_1,t_1},K^{[t_0^m,T]}_{i_2,t_2})\rangle V^{i_1}_{j_1}(\Theta^{t_1}_{t_2})V^{i_2}_{j_2}(\Theta^{t_2}_{t_2})\\&+\langle\partial_{\mathbf{x}_{i_1}}u(t_0^m,\Theta^{[t_0^m,T]}_{t_2}),K^{[t_0^m,T]}_{i_1,t_1}\rangle K_{i_2}(t_1,t_2)\partial_{x_{i_2}}V^{i_1}_{j_1}(\Theta^{t_1}_{t_2})V^{i_2}_{j_2}(\Theta^{t_2}_{t_2})\Big]\circ dB^{j_2}_{t_2}\circ dB^{j_1}_{t_1}.\end{aligned}$$

The formulae (2.10) and (2.11) are the first order and second order expansions of $u(t_0^m,\Theta^{[t_0^m,T]}_{t_0^m})$. For higher order expansions, we introduce the following notation. For any $n\ge 1$, denote $\mathcal{I}_n:=\{1,\dots,d_1\}^n$ with elements $\vec{i}=(i_1,\dots,i_n)$ and $\mathcal{J}_n:=\{0,\dots,d\}^n$ with elements $\vec{j}=(j_1,\dots,j_n)$, and introduce a set of mappings for the indices:

$$\mathcal{S}_n:=\Big\{\vec{\kappa}=(\kappa_1,\dots,\kappa_n):\kappa_l\in\{0,1,\dots,l-1\},l=1,\dots,n\Big\}.$$

Given $\vec{i}\in\mathcal{I}_n$, $\vec{j}\in\mathcal{J}_n$, $\vec{t}\in\mathbb{T}_n$, $\vec{\kappa}\in\mathcal{S}_n$, $\vec{x}=(x_1,\dots,x_n)\in(\mathbb{R}^{d_1})^n$, $\theta\in\mathbb{X}_{t_0^m}$, $\varphi:\mathbb{T}_n^m\to\mathbb{R}$, and $\psi:\mathbb{R}^{d_1}\to\mathbb{R}$, denote

$$\begin{aligned}&\mathcal{N}_\alpha(\vec{\kappa}):=\{l\in\{1,\dots,n\}:\kappa_l=\alpha\},\quad\alpha=0,\dots,n,\\&\mathcal{K}(\vec{i},\vec{\kappa};\vec{t}):=\prod_{l=1}^{n}K_{i_l}(t_{\kappa_l},t_l),\quad\mathcal{K}_+(\vec{i},\vec{\kappa};\vec{t}):=\prod_{\alpha=1}^{n}\prod_{l\in\mathcal{N}_\alpha(\vec{\kappa})}K_{i_l}(t_{\kappa_l},t_l),\\&\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t}):=(K^{[t_0^m,T]}_{i_l,t_l})_{l\in\mathcal{N}_0(\vec{\kappa})},\\&\partial^{\vec{\kappa}}_{\vec{i}}u(t_0^m,\theta):=\partial_{\mathbf{x}_{i_{l_1}}}\cdots\partial_{\mathbf{x}_{i_{l_k}}}u(t_0^m,\theta),\text{ where }\{l_1,\dots,l_k\}=\mathcal{N}_0(\vec{\kappa}),\\&\partial^{\vec{\kappa},\alpha}_{\vec{i}}\psi(x):=\partial_{x_{i_{l_1}}}\dots\partial_{x_{i_{l_k}}}\psi(x),\text{ where }\{l_1,\dots,l_k\}=\mathcal{N}_\alpha(\vec{\kappa}),\\&\mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\vec{x},\theta):=\prod_{\alpha=1}^{n}\partial^{\vec{\kappa},\alpha}_{\vec{i}}V^{i_\alpha}_{j_\alpha}(x_\alpha)\mathcal{K}_+(\vec{i},\vec{\kappa};\vec{t})\,\langle\partial^{\vec{\kappa}}_{\vec{i}}u(t_0^m,\theta),\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t})\rangle,\\&\Theta^{\vec{t}}_s:=(\Theta^{t_1}_s,\dots,\Theta^{t_n}_s),\ s\le t_n,\quad\varphi(\vec{t})\circ dB^{\vec{j}}_{\vec{t}}:=\varphi(\vec{t})\circ dB^{j_n}_{t_n}\circ\cdots\circ dB^{j_1}_{t_1}.\end{aligned}\tag{2.12}$$

Note that $\vec{\mathcal{K}}_0$ and $\mathcal{V}$ here actually depend on m, but we omit this dependence for notational simplicity. We then have the following expansion, whose proof is postponed to the appendix.

Proposition 2.5. *For any $N \geq 1$, under* **(H0)** *and* $\mathbf{(H(N+3))}$, *we have*

$$u(t_0^m, \Theta_{t_0^m}^{[t_0^m,T]}) = u(t_0^m, \Theta_{T_m}^{[t_0^m,T]}) + \sum_{n=1}^{N} \sum_{\vec{\imath}\in\mathcal{I}_n, \vec{\jmath}\in\mathcal{J}_n, \vec{\kappa}\in\mathcal{S}_n} \int_{\mathbb{T}_n^m} \mathcal{V}(\vec{\imath},\vec{\jmath},\vec{\kappa};\vec{t},\Theta_{T_m}^{\vec{t}},\Theta_{T_m}^{[t_0^m,T]}) \circ dB_{\vec{t}}^{\vec{\jmath}}$$
$$(2.13) \qquad + \sum_{\vec{\imath}\in\mathcal{I}_{N+1}, \vec{\jmath}\in\mathcal{J}_{N+1}, \vec{\kappa}\in\mathcal{S}_{N+1}} \int_{\mathbb{T}_{N+1}^m} \mathcal{V}(\vec{\imath},\vec{\jmath},\vec{\kappa};\vec{t},\Theta_{t_{N+1}}^{\vec{t}},\Theta_{t_{N+1}}^{[t_0^m,T]}) \circ dB_{\vec{t}}^{\vec{\jmath}}.$$

2.3. The remainder estimate. In this subsection, we estimate the remainder term in Taylor expansion, which will provide a guideline for our numerical algorithm later. For an appropriate function $\varphi : \mathbb{T}_n^m \times \Omega \to \mathbb{R}$ and $T_m \leq s \leq t_0^m$, denote

$$(2.14) \qquad \|\varphi(\cdot)\|_{s,\vec{\jmath}}^2 := \|\varphi(\cdot)\|_{[T_m,s],\vec{\jmath}}^2 := \mathbb{E}_m\left[\left|\int_{\mathbb{T}_n^m(s)} \varphi(\vec{t}) \circ dB_{\vec{t}}^{\vec{\jmath}}\right|^2\right], \quad \text{where} \quad \mathbb{E}_m := \mathbb{E}_{\mathcal{F}_{T_m}}.$$

Moreover, for $\vec{\jmath} \in \mathcal{J}_n$, $\vec{t} \in \mathbb{T}_n$, and $1 \leq l \leq n$, denote

$$(2.15) \qquad \vec{\jmath}_l := (j_1, \ldots, j_l), \quad \vec{\jmath}_{-l} := (j_{l+1}, \ldots, j_n), \quad \vec{t}_{-l} := (t_{l+1}, \ldots, t_n).$$

We first have the following simple but crucial lemma, whose proof is postponed to the appendix.

Lemma 2.6. *Fix $n \geq 2$, $\vec{\jmath} \in \mathcal{J}_n$, and let $\varphi : \mathbb{T}_n^m \times \Omega \to \mathbb{R}$ be bounded, jointly measurable in all variables, and, for each $\vec{t} \in \mathbb{T}_n^m$, $\varphi(\vec{t})$ is $\mathcal{F}_{t_n}$-measurable in ω. There exists a universal constant $C > 0$ such that, for any $T_m \leq s \leq t_0^m$,*

$$(2.16) \qquad \mathbb{E}_m\left[\int_{\mathbb{T}_n^m(s)} \varphi(\vec{t}) \circ dB_{\vec{t}}^{\vec{\jmath}}\right] = \begin{cases} \int_{T_m}^{s} \mathbb{E}_m\left[\int_{\mathbb{T}_{n-1}^m(t_1)} \varphi(t_1,\vec{t}_{-1}) \circ dB_{\vec{t}_{-1}}^{\vec{\jmath}_{-1}}\right] dt_1, & j_1 = 0, \\ 0, & j_1 \neq 0, j_2, \\ \frac{1}{2}\int_{T_m}^{s} \mathbb{E}_m\left[\int_{\mathbb{T}_{n-2}^m(t_1)} \varphi(t_1,t_1,\vec{t}_{-2}) \circ dB_{\vec{t}_{-2}}^{\vec{\jmath}_{-2}}\right] dt_1, & j_1 = j_2 > 0, \end{cases}$$

$$(2.17) \qquad \|\varphi(\cdot)\|_{s,\vec{\jmath}}^2 \leq \begin{cases} C\delta^2 \operatorname*{ess\,sup}_{T_m \leq s' \leq s} \|\varphi(s',\cdot)\|_{s',\vec{\jmath}_{-1}}^2, & j_1 = 0, \\ C\delta \operatorname*{ess\,sup}_{T_m \leq s' \leq s} \|\varphi(s',\cdot)\|_{s',\vec{\jmath}_{-1}}^2, & j_1 \neq 0, j_2, \\ C\delta \operatorname*{ess\,sup}_{T_m \leq s' \leq s} \|\varphi(s',\cdot)\|_{s',\vec{\jmath}_{-1}}^2 + C\delta^2 \operatorname*{ess\,sup}_{T_m \leq s' \leq s} \|\varphi(s',s',\cdot)\|_{s',\vec{\jmath}_{-2}}^2, & j_1 = j_2 > 0. \end{cases}$$

Note that B^0 and $\{B^j\}_{j\geq 1}$ contribute differently in (2.16) and (2.17). Alternatively, we note that $B_t^0 = t$ is Lipschitz continuous, but B_t^j is Hölder-$(\frac{1}{2} - \varepsilon)$ continuous for $j \geq 1$. To provide a more coherent error estimate, we shall modify (2.13) slightly. For any $1 \leq n \leq N$ and $p \geq 1$, denote

$$\mathcal{J}_{n,N} := \Big\{\vec{j} \in \mathcal{J}_n : \|\vec{j}\| \le N\Big\}, \quad \text{where} \quad \|\vec{j}\| := n + \sum_{l=1}^{n} \mathbf{1}_{\{j_l=0\}},$$
$$A_N^m := \sup_{n\le N} \sup_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_n,\|\vec{j}\|=N,\vec{t}\in\mathbb{T}_n^m} \sup_{\vec{x}\in(\mathbb{R}^{d_1})^n,\theta\in\mathbb{X}_{t_0^m}} \Big|\sum_{\vec{\kappa}\in\mathcal{S}_N} \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\vec{x},\theta)\Big|. \tag{2.18}$$

We then have the following tail estimate, whose proof is postponed to the appendix.

Theorem 2.7. *Let* **(H0)** *and* **(H(N + 3))** *hold, and let* R_N^m *be determined by*

$$u(t_0^m, \Theta_{t_0^m}^{[t_0^m,T]}) = I_N^m + R_N^m, \quad \textit{where}$$
$$I_N^m := u(t_0^m, \Theta_{T_m}^{[t_0^m,T]}) + \sum_{n=1}^{N} \sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_{n,N},\vec{\kappa}\in\mathcal{S}_n} \int_{\mathbb{T}_n^m} \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta_{T_m}^{\vec{t}},\Theta_{T_m}^{[t_0^m,T]}) \circ dB_{\vec{t}}^{\vec{j}}. \tag{2.19}$$

Then there exists a constant $C_N>0$*, which depends only on* N *and* d, d_1*, such that*

$$|\mathbb{E}_m[R_N^m]| \le \Big(\mathbb{E}_m[|R_N^m|^2]\Big)^{\frac{1}{2}} \le C_N\Big[A_{N+1}^m\delta^{\frac{N+1}{2}} + A_{N+2}^m\delta^{\frac{N+2}{2}} + A_{N+3}^m\delta^{\frac{N+3}{2}}\Big]. \tag{2.20}$$

Remark 2.8. Clearly, for fixed N, the error in (2.20) will be smaller when δ is smaller, when G and V_j^i are smoother (so that u is smoother), and when the dimensions d and d_1 are smaller (so that C_N is smaller). This is consistent with our numerical results later. ∎

3. The cubature formula: The one period case.

Note that (2.20) is effective when δ is small. In this section, we consider the case that T is small. Then we may simply set $M = 1$, and thus $\delta = T$. We shall apply the results in section 2 with $m = 0$. In particular, in this case, $\mathbb{E}_0 = \mathbb{E}$. For notational simplicity, in this section we shall omit the superscript ^{0}, e.g., $\mathbb{T}_n = \mathbb{T}_n^0$, $I_N = I_N^0$, and $R_N = R_N^0$.

3.1. Simplification of the stochastic Taylor expansion.

In this case, we have the following: denoting $t_0 := T$,

$$\begin{gathered}\Theta_{T_m}^{t} = \Theta_0^t = x, \quad \mathbb{X}_{t_0^m} = \mathbb{X}_T = \mathbb{R}^{d_1}, \quad u(T,x) = G(x),\, x\in\mathbb{R}^{d_1},\\ \langle\partial_{\vec{i}}^{\vec{\kappa}} u(t_0^m,x), \vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t})\rangle = \partial_{\vec{i}}^{\vec{\kappa},0} G(x) \prod_{l\in\mathcal{N}_0(\vec{\kappa})} K_{i_l}(T,t_l),\\ \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},(x,\dots,x),x) = \mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};x)\mathcal{K}(\vec{i},\vec{\kappa};\vec{t}),\\ \text{where} \quad \mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};x) := \prod_{\alpha=1}^{n} \partial_{\vec{i}}^{\vec{\kappa},\alpha} V_{j_\alpha}^{i_\alpha}(x)\partial_{\vec{i}}^{\vec{\kappa},0}G(x).\end{gathered} \tag{3.1}$$

Thus, (2.19) becomes

$$G(X_T) = I_N + R_N = G(x) + \sum_{n=1}^{N} \sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_{n,N},\vec{\kappa}\in\mathcal{S}_n} \mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};x) \int_{\mathbb{T}_n} \mathcal{K}(\vec{i},\vec{\kappa};\vec{t}) \circ dB_{\vec{t}}^{\vec{j}} + R_N. \tag{3.2}$$

Moreover, by abusing notation we may modify A_N^0 and define A_N as follows:

$$A_N := \sup_{n\le N} \sup_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_n,\|\vec{j}\|=N,\vec{t}\in\mathbb{T}_n^m} \sup_{x\in\mathbb{R}^{d_1}} \Big|\sum_{\vec{\kappa}\in\mathcal{S}_N} \mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};x)\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\Big|. \tag{3.3}$$

Remark 3.1. Motivated from the Taylor expansion (3.2), the step-N Volterra signature should have the following form in the space $\bigoplus_{n=0}^{N}(\mathbb{R}^{d_1+1})^{\otimes n}$:

$$\sum_{n=0}^{N}\sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_n,\vec{\kappa}\in\mathcal{S}_n}\Big(\int_{\mathbb{T}_n}\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\circ dB_{\vec{t}}^{\vec{j}}\Big)(e_{j_1}\otimes\cdots\otimes e_{j_n}), \tag{3.4}$$

where $\{e_j\}_{j=0,1,\dots,d_1}$ denotes the canonical basis of $\mathbb{R}^{d_1+1}$. Below, we shall focus on the expectation of the Volterra signature at any step.

To facilitate the cubature method in the next subsection, we shall rewrite (3.2) slightly further. Note that, for fixed N, the mapping $(\vec{i},\vec{j},\vec{\kappa})\in\bigcup_{n\le N}\mathcal{I}_n\times\mathcal{J}_{n,N}\times\mathcal{S}_n\to\mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};\cdot)$ (as a function of x) is not one to one, so we may combine the terms with the same $\mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};\cdot)$. That is, we may rewrite (3.2) as

$$\begin{aligned} &G(X_T)=G(x)+\sum_{\phi\in\mathbb{V}_N}\phi(x)\Gamma_N^{\phi}+R_N, \quad \text{where}\\ &\mathbb{V}_N:=\bigcup_{n\le N}\Big\{\mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};\cdot):(\vec{i},\vec{j},\vec{\kappa})\in\mathcal{I}_n\times\mathcal{J}_{n,N}\times\mathcal{S}_n\Big\}\subset C(\mathbb{R}^{d_1};\mathbb{R}),\\ &\Gamma_N^{\phi}:=\sum_{n=1}^{N}\sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_{n,N},\vec{\kappa}\in\mathcal{S}_n}\mathbf{1}_{\{\mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};\cdot)=\phi\}}\int_{\mathbb{T}_n}\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\circ dB_{\vec{t}}^{\vec{j}}. \end{aligned} \tag{3.5}$$

Since it requires rather complicated notations to characterize $\phi\in\mathbb{V}_N$ precisely in the general case, we leave it to the special cases that we will actually compute numerically.

3.2. The cubature formula. We now extend the cubature formula for Brownian motion in [40] to the Volterra setting, especially for the Taylor expansion (3.5). From now on, we set $\Omega:=C([0,T];\mathbb{R}^d)$ as the canonical space and B as the canonical process, and thus $\mathbb{P}$ is the Wiener measure so that B is a $\mathbb{P}$-Brownian motion. For some $W\ge 1$, $L\ge 1$, we introduce a discrete probability measure Q on Ω: for some constants $a_{k,l}=(a_{k,l}^1,\dots,a_{k,l}^d)\in\mathbb{R}^d$, $k=1,\dots,W$, $l=1,\dots,L$,

$$\begin{aligned} &Q:=\sum_{k=1}^{2W}\lambda_k\delta_{\omega_k}, \quad \text{where } \delta_\cdot \text{ denotes the Dirac measure}, \quad \lambda_k>0, \quad \sum_{k=1}^{2W}\lambda_k=1,\\ &\lambda_{W+k}=\lambda_k, \quad \omega_{W+k}=-\omega_k, \quad k=1,\dots,W,\\ &\omega_{k,0}=0, \quad \omega_{k,t}=\omega_{k,s_{l-1}}+\frac{a_{k,l}}{\sqrt{T}}[t-s_{l-1}],\ t\in(s_{l-1},s_l],\ s_l:=\frac{l}{L}T,\ 0=1,\dots,L. \end{aligned} \tag{3.6}$$

Here the second line implies that Q is symmetric, since Brownian motion is symmetric. Also, it is OK to consider nonuniform partition $0=s_0<\cdots<s_L=T$. Recall (2.12), for each piecewise linear $\omega=(\omega^1,\dots,\omega^{d_1})$ as in (3.6), $\vec{j}\in\mathcal{J}_n$, and $\varphi:\mathbb{T}_n\to\mathbb{R}$, and denote

$$\int_{\mathbb{T}_n}\varphi(\vec{t})\,d\omega_{\vec{t}}^{\vec{j}}:=\int_{\mathbb{T}_n}\varphi(\vec{t})\,d\omega_{t_n}^{j_n}\cdots d\omega_{t_1}^{j_1}, \quad \text{where} \quad \omega_t^0:=t. \tag{3.7}$$

Then we have

$$\mathbb{E}^Q\Big[\int_{\mathbb{T}_n}\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\circ dB_{\vec{t}}^{\vec{j}}\Big]=\sum_{k=1}^{2W}\lambda_k\int_{\mathbb{T}_n}\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\,d(\omega_k)_{\vec{t}}^{\vec{j}}. \tag{3.8}$$

Definition 3.2. *Let $N \ge 1$, $W \ge 1$. We say that Q defined in (3.6) is an N-Volterra cubature formula on $[0,T]$ if, recalling $\mathbb{E} = \mathbb{E}^{\mathbb{P}}$,*

$$\mathbb{E}^Q[\Gamma_N^\phi] = \mathbb{E}[\Gamma_N^\phi] \quad \text{for all} \quad \phi \in \mathbb{V}_N, \qquad \text{and hence} \quad \mathbb{E}^Q[I_N] = \mathbb{E}[I_N]. \tag{3.9}$$

Recall our goal (2.1)–(2.2). Our main idea is the following approximation:

$$Y_0 := \mathbb{E}[G(X_T)] \approx Y_0^Q := \sum_{k=1}^{2W} \lambda_k G(X_T(\omega_k)), \quad \text{where} \tag{3.10}$$

$$X_t^i(\omega) = x_i + \sum_{j=0}^{d} \int_0^t K_i(t,r) V_j^i(X_r(\omega)) d\omega_r^j, \quad i = 1, \dots, d_1. \tag{3.11}$$

We now have the main result of this section, whose proof is postponed to the appendix.

Theorem 3.3. *Under* **(H0)** *and* **(H(N + 3))***, we have the following: recalling (3.3) and (3.1),*

$$\begin{aligned} |Y_0 - Y_0^Q| \le C_N \Big[A_{N+1}(1 + C_Q^{N-1}) T^{\frac{N+1}{2}} + A_{N+2}(1 + C_Q^{N-2}) T^{\frac{N+2}{2}} + A_{N+3} T^{\frac{N+3}{2}} \Big], \\ \text{where} \quad C_Q := \max_{1 \le k \le W, 1 \le j \le d, 1 \le l \le L} |a_{k,l}^j|. \end{aligned} \tag{3.12}$$

In particular, if each K_i is rescalable, in the sense that there exists an $\alpha_i \in [0,\infty)$ (not necessarily the same as $H_i - \frac{1}{2}$) such that

$$K_i(ct, cr) = c^{\alpha_i} K(t,r) \quad \text{for all } 0 \le r < t, \tag{3.13}$$

then all the $a_{k,l}$ and hence C_Q are independent of T.

3.3. A simplification of the cubature formula. Due to the symmetric properties of Brownian motion and Q, we may simplify the requirement (3.9). Recalling (2.12) and abusing notation, for $\vec{j} \in \mathcal{J}_n$ we denote

$$\mathcal{N}_\alpha(\vec{j}) := \{l \in \{1, \dots, n\} : j_l = \alpha\}, \quad \alpha = 0, \dots, d. \tag{3.14}$$

Lemma 3.4. *Let* **(H0)** *hold, and let $\vec{j} \in \mathcal{J}_n$ be such that $|\mathcal{N}_\alpha(\vec{j})|$ is odd for some $\alpha = 1, \dots, d$; in particular, if $\|\vec{j}\|$ is odd, then*

$$\mathbb{E}\left[\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}} \right] = 0 = \mathbb{E}^Q\left[\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}} \right]. \tag{3.15}$$

Proof. One may easily derive the first equality from (2.16) by induction on n. The second equality follows directly from the symmetric properties of Q. ■

Note further that, when $\vec{j} = (0, \dots, 0) \in \mathcal{J}_n$, we have $dB_{\vec{t}}^{\vec{j}} = d\omega_{\vec{t}}^{\vec{j}} = dt_n \cdots dt_1$. This, together with Lemma 3.4, implies the following result immediately.

Theorem 3.5. *Let* **(H0)** *and* **(H(N+3))** *hold, and denote*

$$\begin{aligned}
&\bar{\mathcal{J}}_{n,N} := \{\vec{j} \in \mathcal{J}_{n,N} \backslash \{(0,\dots,0)\} : |\mathcal{N}_\alpha(\vec{j})| \text{ is even for all } \alpha = 1,\dots,n\},\\
&\bar{\mathbb{V}}_N := \bigcup_{n\le N} \left\{\mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};\cdot) : (\vec{i},\vec{j},\vec{\kappa}) \in \mathcal{I}_n \times \bar{\mathcal{J}}_{n,N} \times \mathcal{S}_n\right\} \subset \mathbb{V}_N,\\
&\bar{\Gamma}^\phi_N := \sum_{n=1}^{N} \sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\bar{\mathcal{J}}_{n,N},\vec{\kappa}\in\mathcal{S}_n} \mathbf{1}_{\{\mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};\cdot)=\phi\}} \int_{\mathbb{T}_n} \mathcal{K}(\vec{i},\vec{\kappa};\vec{t}) \circ dB^{\vec{j}}_{\vec{t}}.
\end{aligned} \tag{3.16}$$

Then Q satisfies (3.9) *if and only if*

$$\mathbb{E}^Q[\bar{\Gamma}^\phi_N] = \mathbb{E}[\bar{\Gamma}^\phi_N] \quad \textit{for all}\, \phi \in \bar{\mathbb{V}}_N. \tag{3.17}$$

When N is odd, note that $\bar{\mathcal{J}}_{N,N} = \emptyset$, so we will get the cubature formula for free at the Nth order. Therefore, we shall always consider odd N.

Example 3.6. (i) In the case $N=3$, obviously we have

$$\bar{\mathcal{J}}_{1,3} = \bar{\mathcal{J}}_{3,3} = \emptyset, \quad \bar{\mathcal{J}}_{2,3} = \{(j,j) : 1 \le j \le d\}. \tag{3.18}$$

(ii) In the case $N=5$, we have

$$\begin{aligned}
&\bar{\mathcal{J}}_{1,5} = \bar{\mathcal{J}}_{5,5} = \emptyset, \quad \bar{\mathcal{J}}_{2,5} = \{(j,j) : 1 \le j \le d\}, \quad \bar{\mathcal{J}}_{3,5} = \{(j,j,0),(j,0,j),(0,j,j)\},\\
&\bar{\mathcal{J}}_{4,5} = \{(j,j,j,j),(j,j,\tilde{j},\tilde{j}),(j,\tilde{j},j,\tilde{j}),(j,\tilde{j},\tilde{j},j) : 1 \le j \ne \tilde{j} \le d\}.
\end{aligned} \tag{3.19}$$

4. The cubature formula: The multiple period case. In this case, we consider general T, and we use the setting in section 2, in particular $\delta := \frac{T}{M}$.

4.1. The cubature formula on each subinterval $[T_m, T_{m+1}]$. Recall (2.19). Note that in (3.2), $\mathcal{V}_0(\vec{i},\vec{j},\vec{\kappa};x)$ and $\int_{\mathbb{T}_n} \mathcal{K}(\vec{i},\vec{\kappa};\vec{t}) \circ dB^{\vec{j}}_{\vec{t}}$ are separated and the cubature measure Q is determined only by $\int_{\mathbb{T}_n} \mathcal{K}(\vec{i},\vec{\kappa};\vec{t}) \circ dB^{\vec{j}}_{\vec{t}}$. In (2.19), however,

$$\mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta^{\vec{t}}_{T_m},\Theta^{[t^m_0,T]}_{T_m}) = \prod_{\alpha=1}^{n} \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{t_\alpha}_{T_m}) \mathcal{K}_+(\vec{i},\vec{\kappa};\vec{t}) \,\langle \partial^{\vec{\kappa}}_{\vec{i}} u(t^m_0,\Theta^{[t^m_0,T]}_{T_m}), \vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t})\rangle$$

and we are not able to move the term $\prod_{\alpha=1}^{n} \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{t_\alpha}_{T_m})$ outside of the stochastic integral, which prohibits us from constructing a desirable Q_m to match the conditional expectations of I^m_N: $\mathbb{E}^{Q_m}_m[I^m_N] = \mathbb{E}_m[I^m_N]$. In light of (2.20), we shall instead content ourselves with

$$\left|\mathbb{E}^{Q_m}_m[I^m_N] - \mathbb{E}_m[I^m_N]\right| \le C\delta^{\frac{N+1}{2}}. \tag{4.1}$$

We shall remark, though, that in general, conditional expectations are only defined in the a.s. sense, which requires specifying the probability on $\mathcal{F}_{T_m}$. However, here we will construct Q_m only on the paths on $[T_m, T_{m+1}]$. For this purpose, we interpret the conditional expectations in a pathwise sense, as we explain in the remark below, so that (4.1) could make sense.

Remark 4.1. Under our conditions, one can easily see that $\mathbb{E}_m[I^m_N] = v_m(\Theta^{[T_m,T]}_{T_m})$ for a deterministic function $v_m \in C(\mathbb{X}_{T_m})$. Similarly, for the Q_m we are going to construct, we will interpret it as a regular conditional probability distribution, and thus we also have the structure $\mathbb{E}_m^{Q_m}[I_N^m] = \tilde{v}_m(\Theta_{T_m}^{[T_m,T]})$ for a deterministic function $\tilde{v}_m \in C(\mathbb{X}_{T_m})$. Then by (4.1) we actually mean a stronger result:

$$|\tilde{v}_m(\theta) - v_m(\theta)| \le C\delta^{\frac{N+1}{2}} \quad \text{for all} \quad \theta \in \mathbb{X}_{T_m}. \tag{4.2}$$

We refer the reader to [54, Chapter 9] for more details of the pathwise stochastic analysis. In this paper, since our main focus is the approximation, to avoid introducing further complicated notations, we abuse notation slightly and write them as conditional expectations.

From now on, we shall assume K_i is sufficiently smooth in (t,r). Then, recalling (2.6), the mapping $s \in [T_m, T] \to \Theta_{T_m}^s$ is smooth: for any $\alpha \ge 0$,

$$\frac{\partial^\alpha}{\partial s^\alpha}\Theta_{T_m}^{i,s} = \sum_{j=0}^d \int_0^{T_m} \frac{\partial^\alpha}{\partial s^\alpha} K_i(s,r) V_j^i(X_r) \circ dB_r^j. \tag{4.3}$$

Then the mapping $\vec{t} \in \mathbb{T}_n^m \to \check{\mathcal{V}}(\vec{t}) := \mathcal{V}(\vec{i}, \vec{j}, \vec{\kappa}; \vec{t}, \Theta_{T_m}^{\vec{t}}, \Theta_{T_m}^{[t_0^m, T]})$ is also smooth. Our idea is to introduce further the Taylor expansion of $\check{\mathcal{V}}(\vec{t})$ at $\vec{T}_m := (T_m, \ldots, T_m)$. For any $\vec{\alpha} = (\alpha_1, \ldots, \alpha_n)$ with $\alpha_l = 0, 1, \ldots,$ denote $\|\vec{\alpha}\| := \sum_{l=1}^n \alpha_l$ and $\vec{\alpha}! := \prod_{l=1}^n (\alpha_l!)$. Then by the standard Taylor expansion formula we have, for any $k \ge 0$,

$$\begin{aligned} &\check{\mathcal{V}}(\vec{t}) = \sum_{\|\vec{\alpha}\| \le k} \frac{1}{\vec{\alpha}!} \frac{\partial^{\|\vec{\alpha}\|}}{\partial \vec{t}^{\vec{\alpha}}} \check{\mathcal{V}}(\vec{T}_m) \prod_{l=1}^n (t_l - T_m)^{\alpha_l} + \check{R}_k(\vec{t}), \text{ where } \frac{\partial^{\|\vec{\alpha}\|}}{\partial \vec{t}^{\vec{\alpha}}} := \frac{\partial^{\|\vec{\alpha}\|}}{\partial t_1^{\alpha_1} \cdots \partial t_n^{\alpha_n}} \\ &\quad \text{and} \quad |\check{R}_k(\vec{t})| \le C_{n,k} \sup_{\|\vec{\alpha}\| = k+1} \sup_{s_l \in [T_m, t_l], l=1,\ldots,n} \left| \frac{\partial^{k+1}}{\partial \vec{t}^{\vec{\alpha}}} \check{\mathcal{V}}(s_1, \ldots, s_n) \right| \delta^{k+1}. \end{aligned} \tag{4.4}$$

We now extend Theorem 2.7. Recall Lemma 3.4 and Theorem 3.5.

Theorem 4.2. *Let N be odd, and let* $\mathbf{(H0\text{-}\frac{N-1}{2})}$, $\mathbf{(H(N+3))}$ *hold. Let $\check{R}_N^m$ be determined by*

$$I_N^m = \check{I}_N^m + \check{R}_N^m, \quad \textit{where} \tag{4.5}$$

$$\begin{aligned} \check{I}_N^m := u(t_0^m, \Theta_{T_m}^{[t_0^m,T]}) + \sum_{n=1}^N \sum_{\vec{i} \in \mathcal{I}_n, \vec{\kappa} \in \mathcal{S}_n} \Bigg[&\sum_{\vec{j} \in \mathcal{J}_{n,N} \setminus \bar{\mathcal{J}}_{n,N}} \int_{\mathbb{T}_n^m} \mathcal{V}(\vec{i}, \vec{j}, \vec{\kappa}; \vec{t}, \Theta_{T_m}^{\vec{t}}, \Theta_{T_m}^{[t_0^m,T]}) \circ dB_{\vec{t}}^{\vec{j}} \\ &+ \sum_{\vec{j} \in \bar{\mathcal{J}}_{n,N}} \sum_{\vec{\alpha}: \|\vec{\alpha}\| \le \frac{N - \|\vec{j}\|}{2}} \frac{1}{\vec{\alpha}!} \frac{\partial^{\|\vec{\alpha}\|}}{\partial \vec{t}^{\vec{\alpha}}} \mathcal{V}(\vec{i}, \vec{j}, \vec{\kappa}; \vec{T}_m, \Theta_{T_m}^{\vec{T}_m}, \Theta_{T_m}^{[t_0^m,T]}) \int_{\mathbb{T}_n^m} \prod_{l=1}^n (t_l - T_m)^{\alpha_l} \circ dB_{\vec{t}}^{\vec{j}} \Bigg]. \end{aligned}$$

Then there exists a constant C_N^m, which depends on N, H_i, and the upper bounds of V_j^i and their derivatives up to the order $N+2$, such that

$$\mathbb{E}_m[|\check{R}_N^m|^2] \le C_N^m e^{C_N^m T} \delta^{N+1}. \tag{4.6}$$

Proof. For each $\vec{j}\in\bar{\mathcal{J}}_{n,N}$, noting that $\|\vec{j}\|$ is even and N is odd, set $k:=\frac{N-\|\vec{j}\|-1}{2}$. Using the notations in (4.4), one can see that $\frac{\partial^{k+1}}{\partial \vec{t}^{\vec{\alpha}}}\check{\mathcal{V}}$ involves the derivatives of V^i_j and $u(t^m_0,\cdot)$ up to the order $n+k+1$ and the derivatives of K_i up to the order $k+1$. Note that $2\le\|\vec{j}\|\le N-1$, and then

$$n+k+1\le\|\vec{j}\|+\frac{N-\|\vec{j}\|-1}{2}+1=\frac{N+\|\vec{j}\|+1}{2}+1\le N+1,$$
$$k+1=\frac{N-\|\vec{j}\|-1}{2}+1\le\frac{N-3}{2}+1\le\frac{N-1}{2}.$$

Recall Proposition 2.4 for the bounds of the derivatives of u. Now, following the arguments in Theorem 2.7, one can easily see that, for some appropriate constant C depending on the parameters specified in this theorem,

$$\mathbb{E}\left[\left|\int_{\mathbb{T}^m_n}\check{R}_k(\vec{t})\circ dB^{\vec{j}}_{\vec{t}}\right|^2\right]\le Ce^{C^m_N T}\delta^{2(k+1)+\|\vec{j}\|}=Ce^{C^m_N T}\delta^{N+1}.$$

This implies (4.6) immediately. ∎

We next introduce Q_m as in (3.6) but on paths on $[T_m,T_{m+1}]$:

$$\begin{gathered}Q_m:=\sum_{k=1}^{2W}\lambda_k\delta_{\omega_k},\quad \lambda_k>0,\quad \sum_{k=1}^{2W}\lambda_k=1,\\ \lambda_{W+k}=\lambda_k,\quad \omega_{W+k}=-\omega_k,\quad 1\le k\le W,\\ \omega_{k,0}=T_m,\ \omega_{k,t}=\omega_{k,s_{l-1}}+\frac{a_{k,l}}{\sqrt{\delta}}[t-s_{l-1}],\ t\in(s_{l-1},s_l],\ s_l:=T_m+\frac{l}{L}\delta,\ 0\le l\le L.\end{gathered}\tag{4.7}$$

Recall Remark 4.1.

Definition 4.3. *Let $N\ge 1$ be odd, and fix m. We say Q_m defined in (4.7) is a modified N-Volterra cubature formula on $[T_m,T_{m+1}]$ if, for all $n\le N$, $\vec{j}\in\bar{\mathcal{J}}_{n,N}$, and $\|\vec{\alpha}\|\le\frac{N-\|\vec{j}\|-1}{2}$,*

$$\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_n}\prod_{l=1}^n(t_l-T_m)^{\alpha_l}\circ dB^{\vec{j}}_{\vec{t}}\right]=\mathbb{E}_m\left[\int_{\mathbb{T}^m_n}\prod_{l=1}^n(t_l-T_m)^{\alpha_l}\circ dB^{\vec{j}}_{\vec{t}}\right].\tag{4.8}$$

Remark 4.4. (i) The equations in (4.8) corresponding to $\|\vec{\alpha}\|=0$ exactly characterize the cubature measures for Brownian motions. That is, our modified N-Volterra cubature formula is a cubature formula for the standard one but not vice versa in general. In the case $N=3$, however, as we will see in Example 4.6(i) below, the two are equivalent.

(ii) The kernel $\prod_{l=1}^n(t_l-T_m)^{\alpha_l}$ in (4.8) is rescalable in the sense of (3.13). Then, by the same arguments as in Theorem 3.3, $C_{Q_m}:=\max_{1\le k\le W,1\le j\le d,1\le l\le L}|a^j_{k,l}|$ is independent of δ (or M). Indeed, as in (8.9) below, Q_m is a modified N-Volterra cubature formula on $[T_m,T_{m+1}]$ if and only if the following Q^*_N is a modified N-Volterra cubature formula on $[0,1]$:

$$\begin{gathered}Q^*_N:=\sum_{k=1}^{2W}\lambda_k\delta_{\omega_k},\quad \lambda_k>0,\quad \sum_{k=1}^{2W}\lambda_k=1,\\ \lambda_{W+k}=\lambda_k,\quad \omega_{W+k}=-\omega_k,\quad 1\le k\le W,\\ \omega_{k,0}=0,\ \omega_{k,t}=\omega_{k,s_{l-1}}+a_{k,l}[t-s^*_{l-1}],\ t\in(s^*_{l-1},s^*_l],\ s^*_l:=\frac{l}{L},\ 0\le l\le L.\end{gathered}\tag{4.9}$$

We emphasize that Q^*_N is universal, in the sense that it depends only on N, the dimensions, and our construction of the cubature measure, but it does not depend on T, M, or even K. In particular, $C_{Q^*_N} := \max_{1\le k\le W, 1\le j\le d, 1\le l\le L} |a^j_{k,l}|$ is independent of T, M, or $\delta$and $C_{Q_m} = C_{Q^*_N}$.

Theorem 4.5. *Let N be odd, let* $(\mathbf{H0}\text{-}\frac{N-1}{2})$, $(\mathbf{H(N+3)})$ *hold, and let Q_m be as in Definition 4.3. Then, for the C^m_N as in Theorem 4.2 and $C_{Q^*_N}$ in Remark 4.4, we have*

$$\begin{aligned}\Big|\mathbb{E}^{Q_m}_m[u(T_{m+1}, \Theta^{[T_{m+1},T]}_{T_{m+1}})] - \mathbb{E}_m[u(T_{m+1}, \Theta^{[T_{m+1},T]}_{T_{m+1}})]\Big| &\le C^m_N(1+C^{N-1}_{Q^*_N})e^{C^m_N T}\delta^{\frac{N+1}{2}} \\ &+ C_N\Big[A^m_{N+1}(1+C^{N-1}_{Q^*_N})\delta^{\frac{N+1}{2}} + A^m_{N+2}(1+C^{N-2}_{Q^*_N})\delta^{\frac{N+2}{2}} + A^m_{N+3}\delta^{\frac{N+3}{2}}\Big].\end{aligned} \tag{4.10}$$

This proof is also postponed to the appendix.

Example 4.6. (i) When $N=3$, recall (3.18) and note that $\frac{N-\|\vec{j}\|-1}{2} = 0$ for $\vec{j} = (j,j) \in \bar{\mathcal{J}}_{2,3}$, and we see that (4.8) is equivalent to the cubature formula for standard Brownian motions:

$$\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_2} \circ dB^{(j,j)}_{\vec{t}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}^m_2} \circ dB^{(j,j)}_{\vec{t}}\right] = \frac{\delta}{2}. \tag{4.11}$$

(ii) In the case $N=5$, recall (3.19) and note that $\frac{N-\|\vec{j}\|-1}{2} = 1$ for $\vec{j} = (j,j) \in \bar{\mathcal{J}}_{2,5}$ and $\frac{N-\|\vec{j}\|-1}{2} = 0$ for $\vec{j} \in \bar{\mathcal{J}}_{3,5} \cup \bar{\mathcal{J}}_{4,5}$; then (4.8) is equivalent to the folllowing: for $1 \le j \ne \tilde{j} \le d$, $l = 1, 2$,

$$\begin{aligned}
&\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_2} \circ dB^{(j,j)}_{\vec{t}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}^m_2} \circ dB^{(j,j)}_{\vec{t}}\right] = \frac{\delta}{2},\\
&\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_2} (t_l - T_m)\circ dB^{(j,j)}_{\vec{t}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}^m_2} (t_l - T_m)\circ dB^{(j,j)}_{\vec{t}}\right] = \frac{\delta^2}{4},\\
&\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_3} \circ dB^{\vec{j}}_{\vec{t}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}^m_3} \circ dB^{\vec{j}}_{\vec{t}}\right] = \frac{\delta^2}{4}, \quad \vec{j} = (j,j,0), (0,j,j),\\
&\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_3} \circ dB^{(j,0,j)}_{\vec{t}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}^m_3} \circ dB^{(j,0,j)}_{\vec{t}}\right] = 0,\\
&\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_4} \circ dB^{\vec{j}}_{\vec{t}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}^m_4} \circ dB^{\vec{j}}_{\vec{t}}\right] = \frac{\delta^2}{8}, \quad \vec{j} = (j,j,j,j), (j,j,\tilde{j},\tilde{j}),\\
&\mathbb{E}^{Q_m}_m\left[\int_{\mathbb{T}^m_4} \circ dB^{\vec{j}}_{\vec{t}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}^m_4} \circ dB^{\vec{j}}_{\vec{t}}\right] = 0, \quad \vec{j} = (j,\tilde{j},j,\tilde{j}), (j,\tilde{j},\tilde{j},j).
\end{aligned} \tag{4.12}$$

4.2. The cubature formula on the whole interval $[0, T]$.

Recall that Q_m is defined on $C([T_m, T_{m+1}]; \mathbb{R}^{d_1})$. We shall now compose all the Q_m:

$$Q := Q_0 \otimes \cdots \otimes Q_{M-1}. \tag{4.13}$$

Here $\otimes$ refers to independent composition. Then Q is a probability measure on $\Omega = C([0,T]; \mathbb{R}^{d_1})$. Similarly, let $\mathbb{P}_m$ denote the Wiener measure on $C([T_m, T_{m+1}]; \mathbb{R}^{d_1})$; then $\mathbb{P} = \mathbb{P}_0 \otimes \cdots \otimes \mathbb{P}_{M-1}$. The following result extends [40, Theorem 3.3] to our setting.

Theorem 4.7. *Let N be odd, let* $\mathbf{(H0\text{-}\frac{N-1}{2})}$, $\mathbf{(H(N+4))}$ *hold, and let Q be defined by* (4.13) *with each Q_m as in Definition* 4.3. *Then, for the C_N^m in Theorem* 4.5 *and $C_{Q_N^*}$ in Remark* 4.4, *we have*

$$\begin{aligned}\Big|\mathbb{E}^Q[G(X_T)]-\mathbb{E}[G(X_T)]\Big| &\le \sum_{m=0}^{M-1}\Big[C_N^m(1+C_{Q_N^*}^{N-1})e^{C_N^m T}\delta^{\frac{N+1}{2}}\\ &+C_N[A_{N+1}^m(1+C_{Q_N^*}^{N-1})\delta^{\frac{N+1}{2}}+A_{N+2}^m\delta^{\frac{N+2}{2}}(1+[C_{Q_N^*}\sqrt{\delta}]^{N-2})+A_{N+3}^m\delta^{\frac{N+3}{2}}]\Big].\end{aligned} \tag{4.14}$$

Moreover, for a possibly larger C_N which may depend on the bounds of the derivatives of V_j^i up to the order $N+4$ and the $C_{Q_N^*}$, but not on M, we have[2]

$$\Big|\mathbb{E}^Q[G(X_T)]-\mathbb{E}[G(X_T)]\Big|\le C_N M e^{C_N T}\delta^{\frac{N+1}{2}} = C_N e^{C_N T}\frac{T^{\frac{N+1}{2}}}{M^{\frac{N-1}{2}}}, \tag{4.15}$$

which converges to 0 as $M\to\infty$.

Proof. Note that, recalling (2.8),

$$\begin{aligned}&\Big|\mathbb{E}^Q[G(X_T)]-\mathbb{E}[G(X_T)]\Big| = \Big|\mathbb{E}^{Q_0\otimes\cdots\otimes Q_{M-1}}[G(X_T)]-\mathbb{E}^{\mathbb{P}_0\otimes\cdots\otimes\mathbb{P}_{M-1}}[G(X_T)]\Big|\\ &\le\sum_{m=0}^{M-1}\Big|\mathbb{E}^{Q_0\otimes\cdots\otimes Q_m\otimes\mathbb{P}_{m+1}\otimes\cdots\otimes\mathbb{P}_{M-1}}[G(X_T)]-\mathbb{E}^{Q_0\otimes\cdots Q_{m-1}\otimes\mathbb{P}_m\otimes\cdots\otimes\mathbb{P}_{M-1}}[G(X_T)]\Big|\\ &\le\sum_{m=0}^{M-1}\Big|\mathbb{E}^{Q_0\otimes\cdots\otimes Q_m}[u(T_{m+1},\Theta_{T_{m+1}}^{[T_{m+1},T]})]-\mathbb{E}^{Q_0\otimes\cdots Q_{m-1}\otimes\mathbb{P}_m}[u(T_{m+1},\Theta_{T_{m+1}}^{[T_{m+1},T]})]\Big|\\ &\le\sum_{m=0}^{M-1}\mathbb{E}^{Q_0\otimes\cdots\otimes Q_{m-1}}\Big[\Big|\mathbb{E}_m^{Q_m}[u(T_{m+1},\Theta_{T_{m+1}}^{[T_{m+1},T]})]-\mathbb{E}_m^{\mathbb{P}_m}[u(T_{m+1},\Theta_{T_{m+1}}^{[T_{m+1},T]})]\Big|\Big].\end{aligned}$$

Recall Remark 4.1, and note that $\mathbb{E}_m^{\mathbb{P}_m}=\mathbb{E}_m$; then by (4.2) we see that (4.14) follows directly from (4.10). Finally, by (2.12), (2.18), and Proposition 2.4 we obtain (4.15). ■

Remark 4.8. (i) Compared to Theorem 3.3, the above Theorem 4.7 allows us to deal with large T. Moreover, compared with the Q in (3.17), it is easier to construct the cubature measure Q_m in (4.8) and the Q in (4.13). The price to pay, however, is that (4.8) requires higher regularity of K_i in order to have the desired convergence rate.

(ii) Provided sufficient regularity $\mathbf{(H0\text{-}\frac{N-1}{2})}$ on K_i (and $\mathbf{(H(N+4))}$ on V_j^i and G), we have the convergence and its rate in (4.15) as $M\to\infty$, which is very desirable in theory. However, by (4.7) and (4.13), we see that each Q_m will involve $2W$ paths, and thus the independent composition Q will involve $(2W)^M$ paths. Therefore, practically we still don't want to make M too large, which in turn means that T cannot be too large. We note that the same difficulty arises in the Brownian setting, and there have been various ideas on improving

[2]The constant $e^{C_N T}$ below is due to the estimate for the derivatives of u in Proposition 2.4. If one can improve this estimate, under certain technical conditions, then one can replace $e^{C_N T}$ with the new bound for the derivatives of u up to the order $N+3$. This comment is valid for the estimates in (4.6), (4.10), (4.14) as well.

the efficiency, for example the recombination schemes in [30, 38]. It will be very interesting to explore these ideas and see if they can be extended to the Volterra setting.

(iii) Note that the choice of Q_m is not unique. In particular, (4.8) involves a certain number of equations. To make it solvable, we need to allow for a sufficient number of parameters λ_k, $a_{k,l}$, $1 \le k \le W$, $0 \le l \le L-1$. As mentioned in (ii), the complexity of our cubature algorithm increases dramatically for large W but is much less sensitive to the value of L. So, whenever possible, we would prefer a small W while allowing for a reasonably large L. We shall remark that, when the dimension d_1 is large, typically we need a large W. This consideration is not serious for the one period case, which, however, requires T to be small.

(iv) Clearly, we have a better rate for a larger N (again, provided sufficient regularity). However, a larger N implies more equations in (4.8), which in turn requires larger values of W and/or L. In the meantime, a larger N implies a larger C_N in (4.15). So the algorithm may not be always more efficient for a larger N.

5. A one dimensional model.

In this section, we focus on the following model with $d = d_1 = 1$:

$$X_t = x_0 + \int_0^t K(t,r)V(X_r) \circ dB_r, \quad K(t,r) = (t-r)^{H-\frac{1}{2}}, \tag{5.1}$$

where the Brownian motion B is one dimensional and the Hurst parameter $H > \frac{1}{2}$. We investigate a few cases in detail and compute the desired Q. We shall illustrate the efficiency of our algorithm in these cases by several numerical examples in section 7 below.

Note that in this case $V_0^1 \equiv 0$; then there is no need to consider $j = 0$. So, for $n \le N$, by abusing notation we may view $\mathcal{I}_n = \{(1,\dots,1)\}$, $\mathcal{J}_n = \mathcal{J}_{n,N} = \{(1,\dots,1)\}$. We may omit $\vec{i} = (1,\dots,1)$, $\vec{j} = (1,\dots,1)$ inside $\mathcal{K}$ and $\mathcal{V}$ in (2.12).

5.1. The multiple period case with order $N = 3$.

We shall construct Q_m as in (4.7) for a fixed m. In the numerical examples in section 7, we may simply compose these Q_m independently as in (4.13).

In this case, (4.11) consists of only one equation:

$$\mathbb{E}^{Q_m}\left[\int_{T_m}^{T_{m+1}} \int_{T_m}^{t_1} dB_{t_2} \circ dB_{t_1}\right] = \frac{\delta}{2}. \tag{5.2}$$

To construct Q_m, we set $W = 1$ and $L = 1$ in (4.7):

$$\lambda_1 = \frac{1}{2}, \quad d\omega_{1,t} = \frac{a}{\sqrt{\delta}}dt, \, t \in [T_m, T_{m+1}]. \tag{5.3}$$

Then (5.2) becomes

$$\frac{\delta}{2} = \int_{T_m}^{T_{m+1}} \int_{T_m}^{t_1} d\omega_{t_2} d\omega_{t_1} = \frac{a^2}{\delta} \int_{T_m}^{T_{m+1}} \int_{T_m}^{t_1} dt_2 dt_1 = \frac{a^2\delta}{2}.$$

Thus,

$$a = 1, \quad \text{and hence} \quad \omega_{1,t} = \frac{t - T_m}{\sqrt{\delta}}, \, t \in [T_m, T_{m+1}]. \tag{5.4}$$

We remark that the above computation does not involve H, in fact, as we saw in Example 4.6(i), the cubature measure in this case coincides with that of standard Brownian motion. However, in order to have the desired error estimate, by Theorem 4.5 we need $H>\frac{3}{2}$.

5.2. The multiple period case with order $N=5$. While we may apply (4.12) directly, in this one dimensional case actually we may simplify the problem further. Note that the corresponding term which requires the further expansion (4.4) or (4.5) is the following: recalling (2.11) and abusing the notation $\check{\mathcal{V}}(\vec{t})$,

$$\sum_{\vec{\kappa}\in\mathcal{S}_2}\int_{\mathbb{T}_2^m}\mathcal{V}(\vec{\kappa};\vec{t},\Theta^{\vec{t}}_{T_m},\Theta^{[t_0^m,T]}_{T_m})\circ dB^{(1,1)}_{\vec{t}}=\int_{\mathbb{T}_2^m}\check{\mathcal{V}}(\vec{t})\circ dB^{(1,1)}_{\vec{t}},$$
$$\begin{aligned}\text{where}\quad \check{\mathcal{V}}(\vec{t}):=&\langle\partial_{\mathbf{xx}}u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),(K^{[t_0^m,T]}_{t_1},K^{[t_0^m,T]}_{t_2})\rangle\, V(\Theta^{t_1}_{T_m})V(\Theta^{t_2}_{T_m})\\ &+\langle\partial_{\mathbf{x}}u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),K^{[t_0^m,T]}_{t_1}\rangle\, K(t_1,t_2)\partial_xV(\Theta^{t_1}_{T_m})V(\Theta^{t_2}_{T_m}).\end{aligned}$$

For $N=5$, we shall assume **(H0-2)**, namely $H>\frac{5}{2}$; then $K(T_m,T_m)=\partial_tK(T_m,T_m)=\partial_rK(T_m,T_m)=0$. Thus,

$$\begin{aligned}\partial_{t_1}\check{\mathcal{V}}(\vec{t})|_{\vec{t}=(T_m,T_m)}=&\langle\partial_{\mathbf{xx}}u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),(\partial_rK^{[t_0^m,T]}_r|_{r=T_m},K^{[t_0^m,T]}_{T_m})\rangle\, V(X_{T_m})V(X_{T_m})\\ &+\langle\partial_{\mathbf{xx}}u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),(K^{[t_0^m,T]}_{T_m},K^{[t_0^m,T]}_{T_m})\rangle\, V'(X_{T_m})V(X_{T_m})\partial_s\Theta^s_{T_m}|_{s=T_m},\\ \partial_{t_2}\check{\mathcal{V}}(\vec{t})|_{\vec{t}=(T_m,T_m)}=&\langle\partial_{\mathbf{xx}}u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),(K^{[t_0^m,T]}_{T_m},\partial_rK^{[t_0^m,T]}_r|_{r=T_m})\rangle\, V(X_{T_m})V(X_{T_m})\\ &+\langle\partial_{\mathbf{xx}}u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),(K^{[t_0^m,T]}_{T_m},K^{[t_0^m,T]}_{T_m})\rangle\, V'(X_{T_m})V(X_{T_m})\partial_s\Theta^s_{T_m}|_{s=T_m}.\end{aligned}$$

Note that $\langle\partial_{xx}u,(\eta_1,\eta_2)\rangle=\langle\partial_{xx}u,(\eta_2,\eta_1)\rangle$; then

$$\partial_{t_1}\check{\mathcal{V}}(\vec{t})|_{\vec{t}=(T_m,T_m)}=\partial_{t_2}\check{\mathcal{V}}(\vec{t})|_{\vec{t}=(T_m,T_m)}.$$

This leads to the following expansion:

$$\begin{aligned}\int_{\mathbb{T}_2^m}\check{\mathcal{V}}(\vec{t})\circ dB^{(1,1)}_{\vec{t}}=&\check{\mathcal{V}}(T_m,T_m)\int_{\mathbb{T}_2^m}dB^{(1,1)}_{\vec{t}}\\ &+\partial_{t_1}\check{\mathcal{V}}(\vec{t})|_{\vec{t}=(T_m,T_m)}\int_{\mathbb{T}_2^m}[(t_1-T_m)+(t_2-T_m)]\circ dB^{(1,1)}_{\vec{t}}+\check{R}(\vec{t}),\end{aligned}$$

where $\check{R}(\vec{t})$ satisfies the desired estimate. Consequently, we may merge the two equations in the second line of (4.12) into one equation, in the same spirit of (3.5), by considering only their sum. Therefore, in this case, (4.12) reduces to three equations:

$$\begin{aligned}&\mathbb{E}^{Q_m}\left[\int_{T_m}^{T_{m+1}}\int_{T_m}^{t_1}dB_{t_2}\circ dB_{t_1}\right]=\frac{\delta}{2},\\ &\mathbb{E}^{Q_m}\left[\int_{T_m}^{T_{m+1}}\int_{T_m}^{t_1}[(t_1-T_m)+(t_2-T_m)]dB_{t_2}\circ dB_{t_1}\right]=\frac{\delta^2}{2},\\ &\mathbb{E}^{Q_m}\left[\int_{T_m}^{T_{m+1}}\int_{T_m}^{t_1}\int_{T_m}^{t_2}\int_{T_m}^{t_3}dB_{t_4}\circ dB_{t_3}\circ dB_{t_2}\circ dB_{t_1}\right]=\frac{\delta^2}{8}.\end{aligned}\tag{5.5}$$

To construct Q_m, we set $W = 2$ and $L = 1$ in (4.7):

$$\lambda_1 + \lambda_2 = \frac{1}{2}, \qquad d\omega_{k,t} = \frac{a_k}{\sqrt{\delta}} dt,\ k = 1, 2. \tag{5.6}$$

Note that, in light of Remark 4.8, we would prefer a small W. However, if we set $W = 1$ here, the cubature measure Q_m does not exist for any value L. By straightforward calculation, we see that (5.5) becomes

$$\sum_{k=1}^{2} \lambda_k a_k^2 \delta = \frac{\delta}{2}; \qquad \sum_{k=1}^{2} \lambda_k a_k^2 \delta^2 = \frac{\delta^2}{2}; \qquad \sum_{k=1}^{2} \lambda_k \frac{a_k^4}{12} \delta^2 = \frac{\delta^2}{8}. \tag{5.7}$$

In particular, the first two equations coincide, and we obtain

$$a_1^2 = 4\left[1 + \sqrt{\frac{2\lambda_2}{\lambda_1}}\right], \qquad a_2^2 = 4\left[1 - \sqrt{\frac{2\lambda_1}{\lambda_2}}\right]. \tag{5.8}$$

This requires $\lambda_1 \le \frac{1}{6}$ so that $\sqrt{\frac{2\lambda_1}{\lambda_2}} \le 1$. Then, for any $0 \le \lambda_1 \le \frac{1}{6}$, we would obtain a solution by (5.8). One particular solution is

$$\lambda_1 := \frac{1}{6}, \quad \lambda_2 := \frac{1}{3}, \quad a_1 = \sqrt{3}, \quad a_2 := 0. \tag{5.9}$$

We remark that, in this case, $-\omega_2 = \omega_2 = 0$, so we actually have a total of three paths, instead of four paths: by abusing the notation λ_2,

$$\lambda_1 = \frac{1}{6}, \quad d\omega_{1,t} = \sqrt{\frac{3}{\delta}} dt, \quad \lambda_2 = \frac{2}{3}, \quad \frac{d}{dt}\omega_{2,t} = 0, \quad \lambda_3 = \frac{1}{6}, \quad d\omega_{3,t} = -\sqrt{\frac{3}{\delta}} dt. \tag{5.10}$$

5.3. The one period case with order $N = 3$. Recall (3.16) and (3.18), in particular we shall only consider $n = 2$. Then one can verify straightforwardly that

$$\begin{aligned}
&\mathcal{S}_2 = \{(0,0); (0,1)\}, \quad \bar{\mathbb{V}}_3 = \left\{\mathcal{V}_0(\vec{\kappa};\cdot) : \vec{\kappa} \in \mathcal{S}_2\right\} = \{G''V^2, G'V'V\},\\
&\bar{\Gamma}_3^{G''V^2} = \int_{\mathbb{T}_2} \mathcal{K}((0,0);\vec{t}) \circ dB_{\vec{t}}^{(1,1)} = \int_0^T \int_0^{t_1} [(T-t_1)(T-t_2)]^{H-\frac{1}{2}} dB_{t_2} \circ dB_{t_1},\\
&\bar{\Gamma}_3^{G'V'V} = \int_{\mathbb{T}_2} \mathcal{K}((0,1);\vec{t}) \circ dB_{\vec{t}}^{(1,1)} = \int_0^T \int_0^{t_1} [(T-t_1)(t_1-t_2)]^{H-\frac{1}{2}} dB_{t_2} \circ dB_{t_1}.
\end{aligned} \tag{5.11}$$

By (2.16), one can easily compute that

$$\mathbb{E}[\bar{\Gamma}_3^{G''V^2}] = \frac{T^{2H}}{4H}, \quad \mathbb{E}\left[\bar{\Gamma}_3^{G'V'V}\right] = 0. \tag{5.12}$$

To construct Q, we set $W = 1$ and $L = 2$ in (3.6) :

$$\lambda_1 = \frac{1}{2}, \quad d\omega_{1,t} = \left[\frac{a_1}{\sqrt{T}} \mathbf{1}_{[0,\frac{T}{2}]}(t) + \frac{a_2}{\sqrt{T}} \mathbf{1}_{(\frac{T}{2},T]}(t)\right] dt. \tag{5.13}$$

For notational simplicity, we introduce

$$H_- := H - \frac{1}{2} > 0, \quad H_+ := H + \frac{1}{2} > 1. \tag{5.14}$$

Then one may compute

$$\begin{aligned}\mathbb{E}^Q[\bar{\Gamma}_3^{G''V^2}] &= \int_0^T \int_0^{t_1} [(T-t_1)(T-t_2)]^{H_-} d\omega_{1,t_2} d\omega_{1,t_1} = \frac{T^{2H}}{2H_+^2}\left[\left(1 - \frac{1}{2^{H_+}}\right)a_1 + \frac{a_2}{2^{H_+}}\right]^2,\\ \mathbb{E}^Q[\bar{\Gamma}_3^{G'V'V}] &= \int_0^T \int_0^{t_1} [(T-t_1)(t_1-t_2)]^{H_-} d\omega_{1,t_2} d\omega_{1,t_1} = T^{2H}[c_1 a_1^2 + c_2 a_1 a_2 + c_3 a_2^2],\end{aligned} \tag{5.15}$$

where

$$\begin{aligned} c_1 &:= \int_0^{\frac{1}{2}} \int_0^{t_1} [(1-t_1)(t_1-t_2)]^{H_-} dt_2 dt_1 = \frac{1}{H_+}\int_0^{\frac{1}{2}} (1-t)^{H_-} t^{H_+} dt,\\ c_2 &:= \int_{\frac{1}{2}}^1 \int_0^{\frac{1}{2}} [(1-t_1)(t_1-t_2)]^{H_-} dt_2 dt_1 = \frac{1}{H_+}\int_{\frac{1}{2}}^1 (1-t)^{H_-}\left[t^{H_+} - \left(t - \frac{1}{2}\right)^{H_+}\right] dt,\\ c_3 &:= \int_{\frac{1}{2}}^1 \int_{\frac{1}{2}}^{t_1} [(1-t_1)(t_1-t_2)]^{H_-} dt_2 dt_1 = \frac{1}{H_+}\int_{\frac{1}{2}}^1 (1-t)^{H_-}\left(t-\frac{1}{2}\right)^{H_+} dt.\end{aligned} \tag{5.16}$$

Combining (5.12) and (5.15), we obtain from (3.9) that

$$\left[(2^{H_+} - 1)a_1 + a_2\right]^2 = \frac{H_+^2 2^{2H_+}}{2H}, \quad c_1 a_1^2 + c_2 a_1 a_2 + c_3 a_2^2 = 0. \tag{5.17}$$

First, by the second equation we obtain (we may use the other one as well)

$$\frac{a_2}{a_1} = c_4 := \frac{-c_2 + \sqrt{c_2^2 - 4c_1c_3}}{2c_3}. \tag{5.18}$$

Plugging this into the first equation in (5.17), we obtain one solution:

$$a_1 = \frac{H_+ 2^{H_+}}{\sqrt{2H}[2^{H_+} + c_4 - 1]}, \quad a_2 = \frac{H_+ 2^{H_+} c_4}{\sqrt{2H}[2^{H_+} + c_4 - 1]}. \tag{5.19}$$

Example 5.1. Setting $H = \frac{3}{2}$ as above, then one may compute straightforwardly that

$$\begin{gathered} c_1 = \frac{5}{384}, \quad c_2 = \frac{10}{384}, \quad c_3 = \frac{1}{384}, \quad c_4 = -5 + 2\sqrt{5},\\ a_1 = \frac{\sqrt{5}+1}{\sqrt{3}}, \quad a_2 = \frac{5 - 3\sqrt{5}}{\sqrt{3}},\end{gathered}$$

and we obtain Q through (5.13).

5.4. The one period case with order $N = 5$. Recall (3.16) and (3.19); in particular, we shall only consider $n = 2$ and $n = 4$. Clearly, $\bar{\mathbb{V}}_3 \subset \bar{\mathbb{V}}_5$ for the $\bar{\mathbb{V}}_3$ in (5.11). Moreover, again omitting $\vec{i} = (1, \ldots, 1)$, $\vec{j} = (1, \ldots, 1)$,

$$\bar{\mathbb{V}}_5 \backslash \bar{\mathbb{V}}_3 = \left\{\mathcal{V}_0(\vec{\kappa};\cdot) : \vec{\kappa} \in \mathcal{S}_4\right\}.$$

Recall (5.14) and the Gamma function $\Gamma(\alpha,\beta) := \int_0^1 (1-t)^{\alpha-1} t^{\beta-1} dt$, and denote by $\varphi^{(k)}$ the kth derivative of φ. We then have the following result.

Lemma 5.2. *For the above model, we have* $\bar{\mathbb{V}}_5\backslash\bar{\mathbb{V}}_3=\{\phi_\alpha\}_{1\le\alpha\le7}$, *where*

$$\begin{gathered}\phi_1=G^{(4)}V^4,\quad \phi_2=G^{(3)}V'V^3,\quad \phi_3=G''V''V^3,\quad \phi_4=G''(V')^2V^2,\\ \phi_5=G'V^{(3)}V^3,\quad \phi_6=G'V''V'V^2,\quad \phi_7=G'(V')^3V.\end{gathered}\tag{5.20}$$

Moreover, denoting $\gamma_4^{\phi_\alpha}:=\mathbb{E}[\bar{\Gamma}_4^{\phi_\alpha}]$, *we have*

$$\begin{gathered}\gamma^{\phi_1}=\frac{T^{4H}}{32H^2},\quad \gamma_4^{\phi_2}=\frac{T^{4H}}{4H}\Gamma(2H,H_+),\\ \gamma_4^{\phi_3}=\gamma_4^{\phi_4}=\frac{T^{4H}}{4H}\Gamma(2H,2H_+),\quad \gamma_4^{\phi_5}=\gamma_4^{\phi_6}=\gamma_4^{\phi_7}=0.\end{gathered}\tag{5.21}$$

Proof. Note that $\mathcal{S}_4=\bigcup_{\alpha=1}^7\mathcal{S}_{4,\alpha}$:

(5.22)
$$\begin{aligned}&\mathcal{S}_{4,1}:=\{(0,0,0,0)\},\quad \mathcal{S}_{4,5}:=\{(0,1,1,1)\},\quad \mathcal{S}_{4,7}:=\{(0,1,2,3)\},\\ &\mathcal{S}_{4,2}:=\{(0,0,0,1),(0,0,1,0),(0,1,0,0),(0,0,0,2),(0,0,2,0),(0,0,0,3)\},\\ &\mathcal{S}_{4,3}:=\{(0,0,1,1),(0,1,0,1),(0,1,1,0),(0,0,2,2)\},\\ &\mathcal{S}_{4,4}:=\{(0,0,1,2),(0,0,2,1),(0,0,1,3),(0,0,2,3),(0,1,0,2),(0,1,2,0),(0,1,0,3)\},\\ &\mathcal{S}_{4,6}:=\{(0,1,1,2),(0,1,2,1),(0,1,2,2),(0,1,1,3)\}.\end{aligned}$$

By (3.1), one can check that $\mathcal{V}_0(\vec{\kappa};\cdot)=\phi_\alpha$ for all $\vec{\kappa}\in\mathcal{S}_{4,\alpha}$, $\alpha=1,\dots,7$. Then, by (2.16),

$$\begin{aligned}&\mathbb{E}\left[\int_{\mathbb{T}_4}\mathcal{K}(\vec{\kappa};\vec{t})\circ dB_{\vec{t}}^{(1,1,1,1)}\right]=\frac{1}{4}\int_0^T\int_0^{t_1}\mathcal{K}(\vec{\kappa};(t_1,t_1,t_3,t_3))dt_3dt_1\\ &=\frac{T^{4H}}{4}\int_0^1\int_0^{t_1}[(1-t_1)(t_{\kappa_2}-t_1)(t_{\kappa_3}-t_3)(t_{\kappa_4}-t_3)]^{H_-}dt_3dt_1.\end{aligned}$$

Now the expectations in (5.21) follow from straightforward computation. ∎

We next construct a desired Q. Note that (3.17) consists of two equations for $n=2$ and seven equations for $n=4$. To allow for sufficient flexibility, we set $W=2$ and $L=4$ in (3.6):

$$\lambda_1+\lambda_2=\frac{1}{2},\quad d\omega_{k,t}=\sum_{l=1}^4\frac{a_{k,l}}{\sqrt{T}}\mathbf{1}_{[s_{l-1},s_l)}(t)dt,\quad k=1,2.\tag{5.23}$$

Similar to (5.15), the following result is obvious.

Lemma 5.3. *For the above model* (5.23), *we have, for* $k=1,2$,

$$\begin{aligned}&\int_{\mathbb{T}_2}\mathcal{K}((0,0);(t_1,t_2))d(\omega_k)_{\vec{t}}^{(1,1)}=\frac{1}{2H_+^2T}\left[\sum_{l=1}^4[s_l^{H_+}-s_{l-1}^{H_+}]a_{k,l}\right]^2,\\ &\int_{\mathbb{T}_2}\mathcal{K}((0,1);(t_1,t_2))d(\omega_k)_{\vec{t}}^{(1,1)}=\frac{1}{T}\sum_{1\le l_2\le l_1\le4}c(l_1,l_2)a_{k,l_1}a_{k,l_2},\\ &\int_{\mathbb{T}_4}\mathcal{K}(\vec{\kappa};\vec{t})d(\omega_k)_{\vec{t}}^{(1,1,1,1)}=\frac{1}{T^2}\sum_{1\le l_4\le l_3\le l_2\le l_1\le4}c(\vec{\kappa},\vec{l})a_{k,l_1}a_{k,l_2}a_{k,l_3}a_{k,l_4},\end{aligned}\tag{5.24}$$

where

$$
(5.25)\quad
\begin{aligned}
c(l,l) &:= \int_{s_{l-1}}^{s_l}\int_{s_{l-1}}^{t_1}[(T-t_1)(t_1-t_2)]^{H_-}dt_2dt_1 = \frac{1}{H_+}\int_{s_{l-1}}^{s_l}(T-t)^{H_-}(t-s_{l-1})^{H_+}dt,\\
c(l_1,l_2) &:= \int_{s_{l_1-1}}^{s_{l_1}}\int_{s_{l_2-1}}^{s_{l_2}}[(T-t_1)(t_1-t_2)]^{H_-}dt_2dt_1\\
&= \frac{1}{H_+}\int_{s_{l_1-1}}^{s_{l_1}}(T-t)^{H_-}[(t-s_{l_2-1})^{H_+}-(t-s_{l_2})^{H_+}]dt, \quad l_1>l_2,\\
c(\vec{\kappa},\vec{l}) &:= \int_{s_{l_1-1}}^{s_{l_1}}\int_{s_{l_2-1}}^{s_{l_2}\wedge t_1}\int_{s_{l_3-1}}^{s_{l_3}\wedge t_2}\int_{s_{l_4-1}}^{s_{l_4}\wedge t_3}\mathcal{K}(\vec{\kappa};\vec{t})dt_4dt_3dt_2dt_1.
\end{aligned}
$$

Combine (5.12), (5.21), and (5.24), we have the following result.

Theorem 5.4. Equation (3.9) *is equivalent to the following equations:*

$$
(5.26)\quad
\begin{gathered}
\lambda_1+\lambda_2=\frac{1}{2},\\
\frac{1}{H_+^2T}\sum_{k=1}^{2}\lambda_k\left[\sum_{l=1}^{4}[s_l^{H_+}-s_{l-1}^{H_+}]a_{k,l}\right]^2=\frac{T^{2H}}{4H},\\
\sum_{k=1}^{2}\lambda_k\sum_{1\le l_2\le l_1\le 4}c(l_1,l_2)a_{k,l_1}a_{k,l_2}=0,\\
\frac{2}{4!T^2}\sum_{\vec{l}\in\{1,\dots,4\}^4}\sum_{\vec{\kappa}\in\mathcal{S}_{4,\alpha}}\sum_{k=1}^{2}\lambda_k c(\vec{\kappa},\vec{l})a_{k,l_1}a_{k,l_2}a_{k,l_3}a_{k,l_4}=\gamma_4^{\phi_\alpha}, \quad \alpha=1,\dots,7.
\end{gathered}
$$

We remark that (5.26) consists of 10 equations with 10 unknowns: $\lambda_k, a_{k,l}$, $l=1,2,3,4$, $k=1,2$. Since these equations are nonlinear, in particular they involve fourth order polynomials of $a_{k,l}$, in general we are not able to derive explicit solutions as in (5.19). Indeed, even the existence of solutions is not automatically guaranteed, and in that case we can actually increase W and/or L in (5.23) to allow for more unknowns. Nevertheless, we can solve (5.26) numerically, and our numerical examples in the next section show that the numerical solutions of (5.26) serve for our purpose well.

6. A fractional stochastic volatility model.

Consider a financial market where S_t denotes the underlying asset price and U_t is the volatility process:

$$
(6.1)\quad
\begin{aligned}
S_t &= S_0+\int_0^t b_1(r,S_r,U_r)dr+\int_0^t\sigma_1(r,S_r,U_r)\circ dB_r^1,\\
U_t &= U_0+\int_0^t K(t,r)b_2(r,U_r)ds+\int_0^t K(t,r)\sigma_2(r,U_r)\circ dB_r^2.
\end{aligned}
$$

Here B^1,B^2 are correlated Brownian motions with constant correlation $\rho\in[-1,1]$, $K(t,r)=(t-r)^{H-\frac{1}{2}}$ with Hurst parameter $H>\frac{1}{2}$. Assume for simplicity the interest rate is 0, and our goal is to compute the option price $\mathbb{E}[G(S_T)]$.

Note that (6.1) involves the time variable t, so we are in the situation with $d=2$ and $d_1=3$ in (2.1). Indeed, denoting $X_t=(X_t^0,X_t^1,X_t^2):=(t,S_t,U_t)$, then we have

$$
(6.2)\quad
\begin{gathered}
x=(0,S_0,U_0), \quad K_0=K_1=1,\ K_2=K,\\
V^0=(1,0,0), \quad V^1=(b_1,\sigma_1,0), \quad V^2=(b_2,0,\sigma_2).
\end{gathered}
$$

Here, for notational simplicity, we use indices $(0,1,2)$ instead of $(1,2,3)$ for X. We shall emphasize that, although B^1, B^2 are correlated here, the Taylor expansions (2.13) and (2.19) will remain the same, but the expectations in Lemma 2.6 need to be modified in an obvious way. In particular, (3.15) will hold true only when $\|\vec{j}\|$ is odd. Therefore, in this section we modify the $\bar{\mathcal{J}}_{n,N}$ in (3.16), still denoted as $\bar{\mathcal{J}}_{n,N}$, by abusing notation:

$$\bar{\mathcal{J}}_{n,N} := \{\vec{j} \in \mathcal{J}_{n,N} \backslash \{(0,\dots,0)\} : \|\vec{j}\| \text{ is even}\}. \tag{6.3}$$

Then all the results in the previous sections remain true. Alternatively, we may express B^1, B^2 as linear combinations of independent Brownian motions. However, this will make V^1, V^2 more complicated and does not really simplify the analysis below.

6.1. The multiple period case with $N=3$. By (6.3), we see that

$$\bar{\mathcal{J}}_{1,3} = \bar{\mathcal{J}}_{3,3} = \emptyset, \quad \bar{\mathcal{J}}_{2,3} = \{(j_1, j_2) : j_1, j_2 > 0\}.$$

Thus, the cubature measure Q_m should satisfy the following: for $\vec{j} \in \bar{\mathcal{J}}_{2,3}$,

$$\mathbb{E}_m^{Q_m}\left[\int_{\mathbb{T}_n^m} \circ dB_{\vec{t}}^{\vec{j}}\right] = \mathbb{E}_m\left[\int_{\mathbb{T}_n^m} \circ dB_{\vec{t}}^{\vec{j}}\right]. \tag{6.4}$$

This is the same as the Brownian motion case. More precisely,

$$\begin{aligned} \mathbb{E}_m^{Q_m}\left[\int_{\mathbb{T}_2^m} \circ dB_{\vec{t}}^{(1,1)}\right] &= \mathbb{E}_m^{Q_m}\left[\int_{\mathbb{T}_2^m} \circ dB_{\vec{t}}^{(2,2)}\right] = \frac{\delta}{2}, \\ \mathbb{E}_m^{Q_m}\left[\int_{\mathbb{T}_2^m} \circ dB_{\vec{t}}^{(1,2)}\right] &= \mathbb{E}_m^{Q_m}\left[\int_{\mathbb{T}_2^m} \circ dB_{\vec{t}}^{(2,1)}\right] = \frac{\rho\delta}{2}. \end{aligned} \tag{6.5}$$

To construct Q_m, we set $W=2$ and $L=1$ in (4.7): noting that ω is two dimensional,

$$\lambda_1 = \lambda_2 = \frac{1}{4}, \quad d\omega_{k,t} = d(\omega_{k,t}^1, \omega_{k,t}^2) = \left(\frac{a_k^1}{\sqrt{\delta}}, \frac{a_k^2}{\sqrt{\delta}}\right) dt,\ k = 1,2. \tag{6.6}$$

Plugging these into (6.5), we have

$$\frac{\delta}{4}[|a_1^1|^2 + |a_2^1|^2] = \frac{\delta}{4}[|a_1^2|^2 + |a_2^2|^2] = \frac{\delta}{2}, \quad \frac{\delta}{4}[a_1^1 a_1^2 + a_2^1 a_2^2] = \frac{\rho\delta}{2}. \tag{6.7}$$

One can easily solve the above equations:

$$\begin{gathered} a_1^1 = \sqrt{2}\sin(\theta_1), \quad a_2^1 = \sqrt{2}\cos(\theta_1), \quad a_1^2 = \sqrt{2}\sin(\theta_2), \quad a_2^2 = \sqrt{2}\cos(\theta_2) \\ \text{for any } \theta_1, \theta_2 \text{ satisfying } \cos(\theta_1 - \theta_2) = \rho. \end{gathered} \tag{6.8}$$

6.2. The one period case with $N=3$. We first note that, due to the multiple dimensionality here, the system corresponding to (3.17) will be pretty large, especially when $N=5$ in the next subsection. However, since $(X_t^0, X_t^1) = (t, S_t)$ are not of Volterra type, the system can be simplified significantly. Furthermore, we shall modify the cubature method slightly as follows.

Remark 6.1. Recall that in (3.5) and (3.16) we group the terms with the same $\mathcal{V}_0(\vec{i}, \vec{j}, \vec{\kappa}; \cdot)$. Note that the mapping $(\vec{i}, \vec{j}, \vec{\kappa}) \to \int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}}$ is also not one to one, and since $(X_t^0, X_t^1) =$

(t, S_t), many terms do not appear in the Taylor expansion (or, say, the corresponding $\mathcal{V}_0(\vec{i}, \vec{j}, \vec{\kappa}; \cdot) = 0$). It turns out that it will be more convenient to group the terms based on $\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}}$ for this model, as we will do in this and the next subsections. To be precise, let $\tilde{\mathbb{V}}_N^0$ denote the terms $\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}}$ with $(\vec{i}, \vec{j}, \vec{\kappa}) \in \mathcal{I}_n \times (\mathcal{J}_{n,N} \setminus \{(0, \dots, 0)\}) \times \mathcal{S}_n$ appearing in the Taylor expansion, and let$\Delta \tilde{\mathbb{V}}_N^0 := \tilde{\mathbb{V}}_N^0 \setminus \tilde{\mathbb{V}}_{N-1}^0$. We emphasize that, unlike in (3.5), the elements here are not $\mathcal{V}_0$. We then modify Definition 3.2 by replacing (3.9) with

$$\mathbb{E}^Q\left[\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}}\right] = \mathbb{E}\left[\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}}\right] \quad \text{for all the terms in } \tilde{\mathbb{V}}_N^0. \tag{6.9}$$

We remark that if $\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}', \vec{\kappa}'; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}'} = \int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t}) \circ dB_{\vec{t}}^{\vec{j}}$ (as random variables), then automatically we have $\int_{\mathbb{T}_n} \mathcal{K}(\vec{i}', \vec{\kappa}'; \vec{t})\, d\omega_{\vec{t}}^{\vec{j}'} = \int_{\mathbb{T}_n} \mathcal{K}(\vec{i}, \vec{\kappa}; \vec{t})\, d\omega_{\vec{t}}^{\vec{j}}$ for all ω in (3.6).

Recalling from (6.3) that Lemma 3.4 remains true when $\|\vec{j}\|$ is odd, we shall only find $\Delta \tilde{\mathbb{V}}_2^0$. Instead of applying the Taylor expansion (3.2) directly on (6.1), we first expand $\mathbb{E}[G(S_T)]$ as in (3.2). Indeed, note that $\Theta_t^s = (t, S_t, \Theta_t^{s,2})$, where

$$\Theta_t^{s,2} = U_0 + \int_0^t K(s,r) b_2(r, U_r) ds + \int_0^t K(s,r) \sigma_2(r, U_r) \circ dB_r^2. \tag{6.10}$$

Then, applying the chain rule, we have the following: denoting $\tilde{b}_1 := G' b_1$ and $\tilde{\sigma}_1 := G' \sigma_1$,

$$\begin{aligned} G(S_T) &= G(S_0) + \int_0^T \left[\tilde{b}_1(\Theta_{t_1}^{t_1}) dt_1 + \tilde{\sigma}_1(\Theta_{t_1}^{t_1}) \circ dB_{t_1}^1\right] \\ &= R_{\neq 2} + \int_{\mathbb{T}_1} \tilde{b}_1(\Theta_{t_1}^{t_1}) dt_1 + \int_{\mathbb{T}_2} \sum_{i=1}^2 \partial_i \tilde{\sigma}_1(\Theta_{t_2}^{t_1}) \sigma_i(\Theta_{t_2}^{t_2}) K_i(t_1, t_2) \circ dB_{t_2}^i \circ dB_{t_1}^1. \end{aligned} \tag{6.11}$$

Here $R_{\neq n}$ is a generic term whose order is not equal to n. Then, recalling Remark 6.1, by (6.2) we can easily obtain

$$\begin{aligned} &\Delta \tilde{\mathbb{V}}_2^0 := \{\tilde{\Gamma}_2^{(1,1)}, \tilde{\Gamma}_2^{(1,2)}\}, \quad \text{where} \quad \tilde{\Gamma}_2^{(1,i)} := \int_{\mathbb{T}_2} K_i(t_1, t_2) dB_{t_2}^i \circ dB_{t_1}^1, \\ &\mathbb{E}[\tilde{\Gamma}_2^{(1,i)}] = \frac{1}{2} \int_{\mathbb{T}_1} K_i(t_1, t_1) d\langle B^i, B^1 \rangle_{t_1} = \begin{cases} \frac{T}{2}, & i = 1, \\ 0, & i = 2. \end{cases} \end{aligned} \tag{6.12}$$

To construct Q, we set $W = 1$ and $L = 1$ in (3.6) : $d\omega_{1,t} = (\frac{a_1}{\sqrt{T}}, \frac{a_2}{\sqrt{T}}) dt$. Then

$$\mathbb{E}^Q[\tilde{\Gamma}_2^{(1,1)}] = \frac{|a_1|^2}{T} \int_{\mathbb{T}_2} d\vec{t} = \frac{T}{2} |a_1|^2, \quad \mathbb{E}^Q[\tilde{\Gamma}_2^{(1,2)}] = \frac{a_1 a_2}{T} \int_{\mathbb{T}_2} K_2(t_1, t_2) d\vec{t} = \frac{T^{H_+}}{H_+(H_+ + 1)} a_1 a_2.$$

By (6.12), we have

$$\frac{T}{2} |a_1|^2 = \frac{T}{2}, \qquad \frac{T^{H_+}}{H_+(H_+ + 1)} a_1 a_2 = 0, \quad \text{implying} \quad a_1 = 1, \quad a_2 = 0.$$

We remark that this solution is independent of ρ. In fact, numerical results (which are not reported in the paper) show that this does not provide a good approximation, even when T is small. So we shall move to the order $N = 5$ in the next subsection, although it becomes much more involved to find the cubature measure.

6.3. The one period case with $N=5$. In this case, besides the $\Delta\tilde{\mathbb{V}}_2^0$ in (6.12), we also need $\Delta\tilde{\mathbb{V}}_4^0$. For this purpose, we need the second order expansion of $\tilde{b}_1(\Theta_{t_1}^{t_1})$ and the third order expansion of $\tilde{\sigma}_1(\Theta_{t_1}^{t_1})$. These derivations are straightforward but rather tedious. We thus postpone them to the appendix and turn to numerical examples first.

7. Numerical examples.

7.1. The algorithm. Our numerical algorithm consists of the following five steps. We are illustrating only the algorithm in section 5.4. The algorithms in the other subsections, especially those in section 6, need to be modified slightly in the obvious manner.

Step 1. Compute $\mathbb{E}\Big[\int_{\mathbb{T}_n}\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\circ dB_{\vec{t}}^{\vec{j}}\Big]$ for each n and $(\vec{i},\vec{j},\vec{\kappa})$ by using (2.16), and then compute $\mathbb{E}[\bar{\Gamma}_N^{\phi}]$ in (3.17) for each $\phi\in\bar{\mathbb{V}}_N$.

Step 2. Compute $\mathbb{E}^Q\Big[\int_{\mathbb{T}_n}\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\circ dB_{\vec{t}}^{\vec{j}}\Big]$ by (3.8), and then compute $\mathbb{E}^Q[\bar{\Gamma}_N^{\phi}]$ in (3.17) for each $\phi\in\bar{\mathbb{V}}_N$.

Step 3. Establish the equations (3.17) with unknowns $\lambda_k, a_{k,l}$, $k=1,\dots,W$, $l=1,\dots,L$, from (3.6), and then solve these equations to obtain a desired Q. We may in general use numerical methods to solve these (polynomial) equations when explicit solutions are not available, see Remark 7.1 below.

Step 4. For each ω_k obtained in Step 3, solve the (deterministic) ODE (3.11) by discretizing $[0,T]$ equally into D pieces. That is, denoting $h:=\frac{T}{D}$ and $\omega_{-h}:=0$,

$$X_{lh}^{D,i}(\omega)=x_i+\sum_{j=0}^{d}\sum_{\alpha=0}^{l-1}K_i(lh,\alpha h)V_j^i(X_{\alpha h}^D(\omega))\frac{\omega_{(\alpha+1)h}^j-\omega_{(\alpha-1)h}^j}{2},\quad l=0,\dots,D. \tag{7.1}$$

For convenience, we typically set D as a multiple of ML for the M,L in (3.9).

Step 5. We obtain the approximation by (3.10): $Y_0^Q\approx Y_0^{Q,D}:=\sum_{k=1}^{2W}\lambda_k G(X_T^D(\omega_k))$.

We note that, provided the conditions in Remark 2.8, our algorithm is deterministic and is much more efficient than the probabilistic methods, e.g., the Euler scheme in Zhang [55].

Remark 7.1. (i) In subsections 5.1, 5.2, 5.3, 6.1, and 6.2, we have obtained the cubature paths; then we can move to Step 4 directly.

(ii) We remark that Steps 1–3 depend only on the model, more specifically only on K_i, but not on the specific forms of V_j^i or G. So, given the model, we may compute the desired Q offline, and then for each V and G we only need to complete Steps 4 and 5.

(iii) To illustrate the idea for Step 3, we consider the equations in (5.26). We shall use the steepest decent method to minimize the following weighted sum:

$$\begin{aligned}\inf_{\lambda_k\ge 0,a_{k,l}\in\mathbb{R},k=1,2,l=1,2,3,4}\Bigg[&\beta_1\Big|\lambda_1+\lambda_2-\frac{1}{2}\Big|^2+\beta_2\Big|\sum_{k=1}^{2}\lambda_k\sum_{1\le l_2\le l_1\le 4}c(l_1,l_2)a_{k,l_1}a_{k,l_2}\Big|^2\\&+\beta_3\Big|\frac{1}{H_+^2}\sum_{k=1}^{2}\lambda_k\Big[\sum_{l=1}^{4}[s_l^{H_+}-s_{l-1}^{H_+}]a_{k,l}\Big]^2-\frac{T^{2H}}{4H}\Big|^2\\&+\sum_{\alpha=1}^{7}\bar{\beta}_\alpha\Big|\frac{2}{4!}\sum_{\vec{l}\in\{1,\dots,4\}^4}\sum_{\vec{\kappa}\in\mathbb{S}_{4,\alpha}}\sum_{k=1}^{2}\lambda_k c(\vec{\kappa},\vec{l})a_{k,l_1}a_{k,l_2}a_{k,l_3}a_{k,l_4}-\Gamma_4^{\phi_\alpha}\Big|^2\Bigg],\end{aligned}$$

where $\beta_i, \bar{\beta}_\alpha > 0$, $i = 1, 2, 3$, and $\alpha = 1, \dots, 7$, are some appropriate weights.

(iv) For (iii) above, we may replace it with any efficient solver for equations (5.26).

(v) While our algorithm is more sensitive to W than to L, as pointed out in Remark 4.8(iii), a large L will increase the difficulty to solve equations like (5.26). However, since this can be done offline, the impact of L is less serious.

Remark 7.2. In this paper, we focus on the impacts of M and N, but do not analyze rigorously the impact of D (or h) in Step 4, which can be chosen much larger than ML. We shall only comment on it heuristically in this remark.

(i) First, by (7.1), and in particular due to the path dependence, it is clear that the running time of the algorithm grows quadratically (rather than linearly) in D.

(ii) By standard arguments, under mild regularity conditions one can easily show that

$$|\mathbb{E}^Q[G(X_T)] - Y_0^{Q,D}| \le C \sup_{k,l} \frac{|a_{k,l}|}{\sqrt{\delta}} h = C C_{Q_N^*} \sqrt{\frac{M}{T}} \frac{T}{D} = C C_{Q_N^*} \frac{\sqrt{MT}}{D}.$$

So there is a balance between the quadratic running cost and this error estimate. Theoretically, given an error level ε, we shall choose the parameters M, D, and Q which satisfy $C C_{Q_N^*} \frac{\sqrt{MT}}{D} \le \varepsilon$ and minimize the computational cost. Since the algorithm is much more sensitive to M, Q than to D, we content ourselves in this paper to choose a reasonably large D and we identify $Y_0^{Q,D}$ with Y_0^Q to emphasize the dependence on Q. Indeed, our numerical results show that the total error is not sensitive to D; see Example 7.7 below.

In the rest of this section, all the numerics are based on the use of Python 3.7.6 under Quad-Core Intel Core i5 CPU (3.4 GHz). For the running time, we use s and ms to denote second and millisecond, respectively.

7.2. An illustrative one dimensional linear model. In this subsection, we present a one dimensional numerical example:

$$X_t = x_0 + \int_0^t (t-r)^{H-\frac{1}{2}} dB_r, \quad Y_0 = \mathbb{E}[G(X_T)]. \tag{7.2}$$

In this case, $X_T \sim \text{Normal}(x_0, \frac{T^{2H}}{2H})$, so essentially we can compute the true value of Y_0:

$$Y_0^{true} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} G\left(x_0 + \sqrt{\frac{T^{2H}}{2H}} x\right) e^{-\frac{x^2}{2}} dx. \tag{7.3}$$

We use Y_0^{cub}, $Y_0^{mul,M}$, and Y_0^{Euler} to denote the values computed by using the one period cubature formula, the multiple period cubature formula with M periods, and the Euler scheme, respectively. We shall compare our numerical results with this true value. In particular, since the cubature method is deterministic, while Y_0^{Euler} is random, we shall explain how we compare the numerical results of the cubature methods with Y_0^{Euler}.

Our first example shows that, when T is small, the one period method in (3.17) is more efficient than the multiple period method in (4.8) with $M = 1$, especially when H is small. We remark that, although $M = 1$, (3.17) and (4.8) have different kernels and thus have different cubature paths.

Example 7.3. Consider (7.2) with $G(x)=x^2$ and $x_0=0$. Then, with order $N=3$,

$$|Y_0^{cub}-Y_0^{true}|=0<|Y_0^{mul,1}-Y_0^{true}|=\frac{H_-^2}{2HH_+^2}T^{2H}. \tag{7.4}$$

We see that the last error is small when T is small or H is large. However, it is still larger than the error of Y_0^{cub}, which is 0 in this case.

Proof. First, by (7.3) it is clear that $Y_0^{true}=\frac{T^{2H}}{2H}$.

For the one period cubature method with $N=3$, by (3.11) and (5.13) we have

$$X_T(\omega_1)=\int_0^{\frac{T}{2}}(T-t)^{H_-}\frac{a_1}{\sqrt{T}}dt+\int_{\frac{T}{2}}^{T}(T-t)^{H_-}\frac{a_2}{\sqrt{T}}dt$$
$$=\frac{a_1}{\sqrt{T}}\frac{1}{H_+}\left[T^{H_+}-\left(\frac{T}{2}\right)^{H_+}\right]+\frac{a_2}{\sqrt{T}}\frac{1}{H_+}\left(\frac{T}{2}\right)^{H_+}=\left[a_1(2^{H_+}-1)+a_2\right]\frac{T^H}{H_+2^{H_+}},$$
$$X_T(\omega_2)=-X_T(\omega_1).$$

Then, by (3.10) and (5.17),

$$Y_0^{cub}=\frac{1}{2}\sum_{k=1}^{2}|X_T(\omega_k)|^2=[a_1(2^{H_+}-1)+a_2]^2\frac{T^{2H}}{H_+^2 2^{2H_+}}=\frac{H_+^2 2^{2H_+}}{2H}\frac{T^{2H}}{H_+^2 2^{2H_+}}=\frac{T^{2H}}{2H}=Y_0^{true}.$$

However, for the multiple period cubature method with $M=1$ and $N=3$, by (5.3) and (5.4) we have the following: by abusing the notations ω_k,

$$X_T(\omega_1)=\int_0^T(T-t)^{H_-}\frac{1}{\sqrt{T}}dt=\frac{T^H}{H_+},\quad X_T(\omega_2)=-X_T(\omega_1),$$
$$Y_0^{mul,1}=\frac{1}{2}\left[|X_T(\omega_1)|^2+|X_T(\omega_2)|^2\right]=\frac{T^{2H}}{H_+^2},$$
$$|Y_0^{mul,1}-Y_0^{true}|=\left|\frac{1}{H_+^2}-\frac{1}{2H}\right|T^{2H}=\frac{H_-^2}{2HH_+^2}T^{2H}. \qquad \blacksquare$$

Our next example shows that, when T is large, the one period algorithm fails, but the multiple period algorithm does converge when M becomes large, as we expect.

Example 7.4. Consider (7.2) with $T=3$, $H=5/2$, $G(x)=(x-1/2)^+$, $x_0=0.56$. We compute the value Y_0^{cub} from the one period model with $N=3$ and $Y_0^{mul,M}$ from the multiple period model with $N=3$ and M from 1 to 5. The cubature paths are constructed as in Example 7.3, with $D=300$ in Step 4. Numerical results are reported in Table 1.

We remark that here the function G is not as smooth as required in Theorems 3.3 and 4.7. Nevertheless, we see from the numerical results that the cubature method still works well.

Table 1

The numerical results for Example 7.4.

Y_0^{true}	Y_0^{cub}	$Y_0^{mul,1}$	$Y_0^{mul,2}$	$Y_0^{mul,3}$	$Y_0^{mul,4}$	$Y_0^{mul,5}$
2.8112	3.5157	2.6281	3.2450	3.1967	3.0340	2.8883

We now present an example to compare the accuracy between $N=3$ and $N=5$.

Example 7.5. Consider (7.2) with $T = 0.3$, $H = 3/2$, and try three different G with corresponding initial value x_0. We choose $D=30$ in Step 4 and compute $Y_0^{mul,2}$ with $N=3$ and $N=5$, respectively. The numerical results are reported in Table 2.

We note that, although we have a better rate of convergence in Theorem 4.7 when $N=5$, the numerical results in this example do not show such improvement. One explanation is that the constant C_N in (4.15) becomes larger when we increase N; see Remark 4.8(iv). The numerical results for the one period cubature method do not show significant improvement either when we increase the order from $N=3$ to $N=5$. However, for the fractional stochastic volatility model (6.1), as we saw in section 6.2, the one period cubature method with $N=3$ does not depend on ρ at all, and thus it is clearly not as good as the one period cubature method with $N=5$.

Our next example compares the one period cubature method with the Euler scheme. More examples concerning this comparison will be presented in the next two subsections.

Example 7.6. Consider (7.2) with $H=\frac{3}{2}$, $T=0.2$, $G(x)=\cos(x)$, and $x_0=1$. We choose $D = 12$ in Step 4, and Y_0^{cub} is computed with $N = 5$. The numerical results are shown in Table 3.

We now explain the numerical results in Table 3. First, in this case, $Y_0^{true} = 0.53959$. Next, by solving (7.2) numerically, we report the approximate solution of (5.26) in Table 4;

Table 2
The numerical results for Example 7.5.

$G(x)\,/\,x_0$	$\cos(x)\,/\,1$	$x^2\,/\,1$	$(x-1/2)^+\,/\,0.56$
Y_0^{true}	0.5378641	1.0090376	0.0751964
$Y_0^{mul,2}(N=3)$	0.5380251	1.0084375	0.0740474
$Y_0^{mul,2}(N=5)$	0.5380277	1.0084375	0.0751558

Table 3
The numerical results for Example 7.6.

$\widetilde{M}_{Euler}$	100	500	1000
e^{cub} (cubature time)	0.00046 (1.86ms)	0.00046 (1.86ms)	0.00046 (1.86ms)
e^{Euler}_{mean} (Euler time)	0.00338 (14.2ms)	0.00156 (70ms)	0.00114 (141ms)
SD^{Euler} (percentile)	0.0026 (19.4%)	0.00119 (24%)	0.00081 (27.5%)

Table 4
The approximate cubature paths in Example 7.6.

k	1	2
λ_k	0.15332891	0.34667109
$a_{k,1}$	−3.04533315	1.57981296
$a_{k,2}$	0.71729258	−2.08974376
$a_{k,3}$	−0.60085202	2.33258457
$a_{k,4}$	0.12029985	−4.5060389

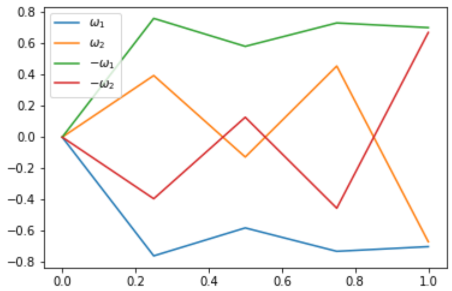

Figure 1. *The approximate cubature paths in Example 7.6 (rescaling to $T=1$).*

see also Figure 1 for the plot of the four approximate cubature paths (after rescaling to $T=1$). We then obtain $Y_0^{cub}=0.54005$. In Table 3, the cubature error $e^{cub}:=|Y_0^{cub}-Y_0^{true}|=0.00046$. The reported running time is for Steps 4 and 5 only, since Steps 1–3 can be completed offline once for all.

For the Euler scheme, we also set time discretization step $D=12$. Let $\widetilde{M}_{Euler}$ denote the sample size in the Euler scheme, namely the number of simulated paths of the Brownian motion. Clearly, both the approximate value Y_0^{Euler} and the running time depend on $\widetilde{M}_{Euler}$; in particular, the latter is proportional to $\widetilde{M}_{Euler}$. Note that Y_0^{Euler} is random. We repeat the Euler scheme 1000 times, each with sample size $\widetilde{M}_{Euler}$, and obtain $Y_0^{Euler,i}$ and the corresponding Euler scheme errors $e_i^{Euler}:=|Y_0^{Euler,i}-Y_0^{true}|$, $1\le i\le 1000$. We shall use the sample median e_{median}^{Euler} of $\{e_i^{Euler}\}_{1\le i\le 1000}$ to measure the accuracy of the Euler scheme. We can then compare e^{cub} and e_{median}^{Euler}, and to have a more precise comparison, we will actually compute the percentile of the cubature error e^{cub} among the Euler scheme errors $\{e_i^{Euler}\}_{1\le i\le 1000}$: α-percentile means that about $1000\times\alpha\%=10\alpha$ of $\{e_i^{Euler}\}_{1\le i\le 1000}$ are smaller than e^{cub}. So 50% roughly means $e^{cub}=e_{median}^{Euler}$ and the two methods have the same accuracy, while $\alpha\%<50\%$ means that $e^{cub}<e_{median}^{Euler}$ and the cubature method has better accuracy: the smaller $\alpha\%$ is, the better the cubature method outperforms. Moreover, since $\{e_i^{Euler}\}_{1\le i\le 1000}$ are i.i.d., we may use the normal approximation to compute the percentile. We will report their mean $e_{mean}^{Euler}:=\frac{1}{1000}\sum_{i=1}^{1000}e_i^{Euler}\approx e_{median}^{Euler}$ and standard deviation SD^{Euler}, and then the percentile $\alpha\%\approx\Phi(\frac{e^{cub}-e_{mean}^{Euler}}{SD^{Euler}})$, where Φ is the cdf of the standard normal.

For the above example, we test three cases, $\widetilde{M}_{Euler}=100,500,1000$, and the numerical results are reported in Table 3. As we see, when $\widetilde{M}_{Euler}=500$, the Euler scheme takes 70 milliseconds (for each run, not for 1000 runs), which is about 38 times slower than the 1.86 milliseconds used for the cubature method, and the percentile of the cubature method is 24%. So the cubature method outperforms the Euler scheme both in running time and in accuracy. When we increase the sample size $\widetilde{M}_{Euler}$ to 1000, the percentile increases to 27%, so the cubature method still outperforms in accuracy. In this case, the running time of the Euler scheme increases to 141 milliseconds, which is about 76 times slower than the cubature method. On the other hand, if we decrease the sample size $\widetilde{M}_{Euler}$ to 100, the running time of the Euler scheme drops to 14.2 milliseconds, which is still 7.6 times slower than the cubature method, but the accuracy deteriorates further with a percentile 19.4%. So, in all three cases,

Table 5
The numerical results for Example 7.7.

D	12	60	120
e^{cub} (cubature time)	0.00046 (1.86ms)	0.00046 (15.3ms)	0.00046 (54.7ms)
e^{Euler}_{mean} (time, $\tilde{M}=100$)	0.00338 (14.2ms)	0.0034 (314ms)	0.00336 (1.24s)
e^{Euler}_{mean} (time, $\tilde{M}=500$)	0.00156 (70ms)	0.00161 (1.54s)	0.00108 (5.9s)
e^{Euler}_{mean} (time, $\tilde{M}=1000$)	0.00114 (141ms)	0.00113 (3.11s)	0.00107 (11.5s)

the cubature method outperforms the Euler scheme significantly both in running time and in accuracy.

We conclude this subsection with an example concerning the impact of D.

Example 7.7. Consider the same setting as in Example 7.6, but try three different D's. The numerical results are shown in Table 5.[3]

As we can see, the cubature method is not sensitive to D. The Euler scheme does rely on our choices of D and $\tilde{M}_{Euler}$. However, in all the above choices, the cubature method outperforms both in running time and in accuracy.

7.3. A one dimensional nonlinear model. We now consider the following nonlinear model, but still in the one dimensional setting:

$$X_t = x_0 + \int_0^t (t-r)^{H-\frac{1}{2}} \cos(X_s) dB_r, \quad Y_0 = \mathbb{E}[G(X_T)]. \tag{7.5}$$

We shall use the one period cubature method with $N=5$ when T is small, and we shall use the multiple period cubature method with $N=3$ and appropriate M when T is large. Our main purpose is to compare the efficiency of the cubature method with that of the Euler scheme.

We first note that, by Remark 7.1(ii), the cubature paths for (7.5) are the same as those for (7.2). In particular, for the one period method with $N=5$, we may continue to use the paths in Table 4. For comparison purposes, we will use the same D for the cubature method and the Euler scheme. For the $\widetilde{M}_{Euler}$ in the Euler scheme, there is an obvious tradeoff between the running time and the accuracy. While one may try to find an "optimal" $\tilde{M}_{Euler}$ for a given D, such an analysis relies on a precise idea on the constants involved in the error estimates, as in Remark 7.2(ii). Since our main focus is the cubature method, and since our examples show that the cubature method outperforms significantly (under our strong conditions), we do not go through that analysis. Instead, unless stated otherwise, for simplicity in the rest of this section we shall always set

$$\widetilde{M}_{Euler} = 500, \quad \text{and we repeat the Euler scheme 1000 times,}$$

each time with $\widetilde{M}_{Euler}$ simulation paths. We use $\{Y_0^{Euler,i}\}_{1\le i\le 1000}$ and Y_0^{true} to compute e^{Euler}_{mean} and SD^{Euler}. However, in this case we are not able to compute the exact value of Y_0^{true}

[3] The running time for the Euler scheme grows quadratically in D, as expected. However, the running time for the cubature method grows slower than quadratically, especially when D is not that large. This is possibly because in our code, for the sake of readability, there is a relatively time consuming step whose cost grows linearly in D. The efficiency of our cubature method could be improved slightly further, when D is small, if we write the code in a more straightforward way.

Table 6
The numerical results for Example 7.8.

$G(x)$ / x_0	$\cos(x)$ / 1	x^2 / 1	$(x-0.5)^+$ / 0.56
Y_0^{true}	0.5401	1.00073	0.0617
Y_0^{cub}	0.5402	1.00041	0.0601
e^{cub} (time)	0.0001(1.75ms)	0.00032(1.66ms)	0.0016 (1.63ms)
e^{Euler}_{mean} (time)	0.0008(80.4ms)	0.00200(82ms)	0.00147(75ms)
SD^{Euler}(percentile)	0.00064(18%)	0.00149(20.1%)	0.0011(54.3%)

Table 7
The numerical results for Example 7.9.

T	0.2	0.5	0.8
Y_0^{true}	1.00073	1.0109	1.045
Y_0^{cub}	1.00041	1.0056	1.022
e^{cub} (time)	0.00032(1.66ms)	0.0054(5.25ms)	0.0234(11ms)
e^{Euler}_{mean} (time)	0.00200 (82ms)	0.0332(445ms)	0.0152(1.16s)
SD^{Euler}(percentile)	0.00149(20.1%)	0.0098(57.4%)	0.0115(69.4%)

as in (7.3). Since the convergence of the Euler scheme approximations is well understood, for comparison purposes we shall set the true value as the sample mean of the Euler scheme approximations:

$$Y_0^{true} = \frac{1}{1000}\sum_{i=1}^{1000} Y_0^{Euler,i}. \tag{7.6}$$

In the first example, we show the impact of the regularity of G.

Example 7.8. Consider (7.5) with $H = \frac{3}{2}$, $T = 0.2$, and consider three different G's with corresponding x_0. We choose $D = 12$ and compare the one period cubature method with $N = 5$ with the Euler scheme. The numerical results are reported in Table 6.

As we can see, the cubature method outperforms the Euler scheme in all three cases. However, when G becomes less smooth, the advantage of the cubature method fades away, which is consistent with the theoretical observation in Remark 2.8.

The next example illustrates the impact of T.

Example 7.9. Consider (7.5) with $H = \frac{3}{2}$, $G(x) = x^2$, $x_0 = 1$ and three different values of T: 0.2, 0.5, 0.8. We choose $D = 12$, 30, 48, respectively, and compare the one period cubature method with $N = 5$ with the Euler scheme. The numerical results are reported in Table 7.

Again consistent with our theoretical result, the performance of the cubature method decays when T gets large. In the above example, the cubature method obviously outperforms the Euler scheme when $T = 0.2$, and still works better when $T = 0.5$, but in the case $T = 0.8$, there is a tradeoff between the speed and the accuracy and it is hard to claim the cubature method is more efficient. In the last case, we shall use the multiple period cubature method, as we do in the next example, and we can easily see that the cubature method outperforms the Euler scheme again.

Table 8
The numerical results for Example 7.10.

$G(x)$ / x_0	$\cos(x)$ / 1	x^2 / 1	$(x-0.5)^+$ / 0.56
Y_0^{true}	0.5136	1.098	0.2275
$Y_0^{mul,5}$	0.5186	1.084	0.2297
$e^{mul,5}$ (time)	$0.005(300ms)$	$0.014(302ms)$	$0.0022(302ms)$
e^{Euler}_{mean} (time)	0.0074(4.2s)	0.023 (4.36s)	0.011 (4.41s)
SD^{Euler}(percentile)	0.0056 (37.37%)	0.017 (35.1%)	0.0089 (21%)

Table 9
The numerical results for Example 7.11.

$G(x)$ / S_0	$\cos(x)$ / 1	x^2 / 1	$(x-0.5)^+$ / 0.56
Y_0^{true}	0.4270	1.2947	0.1320
Y_0^{cub}	0.4257	1.2967	0.1283
e^{cub} (time)	$0.0013(8.53ms)$	$0.0020(8.9ms)$	$0.0037(8.68ms)$
e^{Euler}_{mean} (time)	0.0063 (273ms)	0.0157(283ms)	0.0037(272ms)
SD^{Euler}(percentile)	0.0047 (21.3%)	0.0119 (19.2%)	0.0028 (50%)

Example 7.10. Consider (7.5) with $H=\frac{3}{2}$, $T=1$, and three different G's with corresponding x_0. We choose $D=100$, and compare the multiple period cubature method with $M=5$ and $N=3$ with the Euler scheme. The numerical results are reported in Table 8.

7.4. A fractional stochastic volatility model. In this section, we consider the following special case of (6.1) with $H=\frac{3}{2}$ and $\rho=\frac{1}{2}$:

$$\begin{aligned} dS_t &= S_t b_1(U_t)dt + S_t\sigma_1(U_t)\circ dB^1_t,\\ U_t &= 1+\int_0^t [t-s]\left[\frac{1}{2}-\frac{1}{3}U_s\right]ds + \int_0^t [t-s]\sigma_2(U_s)\circ dB^2_t. \end{aligned} \tag{7.7}$$

Again we will use one period cubature method with $N=5$ when T is small, and the multiple period cubature method with $N=3$ and appropriate M when T is large, and we shall compare the efficiency between the cubature method and the Euler scheme.

Example 7.11. Consider (7.7) with $T=0.1$, $b_1(U)=U$, $\sigma_1(U)=\sigma_2(U)=\cos(U)$, and consider three different G's with corresponding S_0. We choose $D=12$ and compare the one period cubature method with $N=5$ with the Euler scheme. The numerical results are reported in Table 9.

For the cubature method, we first compute the cubature paths following the same idea as in Remark 7.1. By section B.1 below, we choose $W=5$ and $L=4$. Then we obtain

$$\lambda_1=0.0247245002,\quad \lambda_2=0.0561159547,\quad \lambda_3=0.417734596e,$$
$$\lambda_4=0.00142494883,\quad \lambda_5=4.44061201e-17$$

and the 10 paths are plotted in Figure 2 (after rescaling to $T=1$).

Example 7.12. Consider the same setting as in Example (7.7), except that $b_1(U)=\sigma_1(U)=\sigma_2(U)=\sqrt{U}$. The numerical results are reported in Table 10.

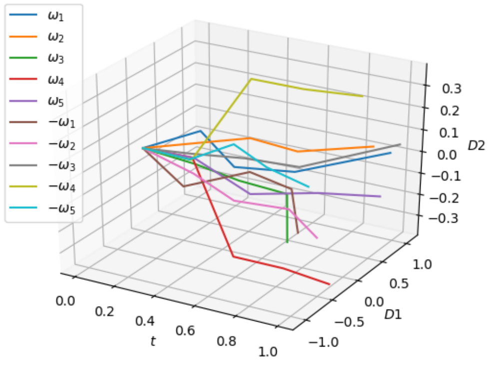

Figure 2. *The cubature paths for the model (7.7) (rescaling to $T = 1$).*

Table 10
The numerical results for Example 7.12.

$G(x)$ / S_0	$\cos(x)$ / 1	x^2 / 1	$(x-0.5)^+$ / 0.56
Y_0^{true}	0.37897	1.4932	0.17098
Y_0^{cub}	0.3698	1.4887	0.17797
e^{cub} (time)	0.00917(8.06ms)	0.0045(8.39ms)	0.00699(8.07ms)
e_{mean}^{Euler} (time)	0.0119(306ms)	0.0361(311ms)	0.007(304ms)
SD^{Euler}(percentile)	0.0092(40.8%)	0.0270(19.1%)	0.0052(50%)

Table 11
The numerical results for Example 7.13.

H	1	3/2	5/2
Y_0^{true}	1.36	1.33	1.2957
$Y_0^{mul,3}$	1.299	1.302	1.286
$e^{mul,3}$(time)	0.061(2.23s)	0.028(2.26s)	0.0097(2.42s)
e_{mean}^{Euler}(time)	0.033 (15.8s)	0.036(15.8s)	0.037(15.5s)
SD^{Euler}(percentile)	0.024 (81.1%)	0.028 (41.7%)	0.0260 (21%)

We remark that Example 7.12 uses the same cubature paths as in Example 7.11.

Example 7.13. Consider (7.7) with $T = 1$, $b_1(U) = U$, $\sigma_1(U) = \sigma_2(U) = \cos(U)$, $G(x) = (x - 1/2)^+$, $S_0 = 0.56$, and consider three different H. We shall compare the efficiency of the multiple period cubature method with $M = 3$, $N = 3$ and the Euler scheme. We set $D = 100$. Recalling $\rho = \frac{1}{2}$, for the cubature method we use $\theta_1 = \frac{\pi}{6}$ and $\theta_2 = -\frac{\pi}{6}$ in (6.8). We repeat the Euler scheme 100 times (instead of 1000 times). The numerical results are reported in Table 11.

We see that the cubature method outperforms the Euler scheme when $H = \frac{3}{2}$ and $H = \frac{5}{2}$, especially in the latter case, but it does not seem to work well when $H = 1$. This is consistent with our theoretical result.

7.5. Some concluding remarks. We first note that our theoretical convergence analysis for the cubature method, namely Theorems 3.3 and 4.7, is complete, provided sufficient regularities on K and (V, G) (corresponding to N). In particular, it holds true for arbitrary dimensions and arbitrarily large T (for Theorem 4.7).

For the numerical efficiency, in its realm, the cubature method has clear advantages over the Euler scheme. In light of Remark 4.8, the cubature method requires the following three conditions, though: (i) sufficient regularity, so as to obtain the desired error estimate; (ii) low dimension (and relatively small N), so that W can be relatively small; and (iii) not too large T, so that M can be relatively small. We remark that the standard cubature method in [40] for diffusions also requires these conditions. However, the constraints are more severe in the Volterra setting here, for example, the constant C_N in (4.15) is larger here, and for given dimensions and N, there are more equations required in (4.8), and hence we may need a larger W. When the number of cubature paths $(2W)^M$ is large (recalling again Remark 4.8(ii)), it will be interesting to explore whether the approach in [11, 12] could help reduce the complexity, which we shall leave for future research. The less smooth case, especially when $H < \frac{1}{2}$, requires a novel idea to extend our approach.

We shall also remark that the parameters M, N in the cubature method cannot be too big. For the Euler scheme, by increasing the sample size $\widetilde{M}_{Euler}$ gradually one may improve the accuracy "continuously" at the price of sacrificing the speed. For the cubature method, we have only limited choices on M, N and thus lose the flexibility of improving its accuracy "continuously." Consequently, the cubature method is more appropriate in situations where one has a strong requirement on the speed but is less stringent on the accuracy.

8. Appendix. *Proof of Proposition* 2.4. Recall (2.8). Following rather standard arguments, we see that $\partial_{\mathbf{x}} u(t, \cdot)$ exists and has the following representation:

$$\begin{aligned} &\langle \partial_{\mathbf{x}} u(t,\theta), \eta\rangle = \mathbb{E}[\partial_x G(X_T^{t,\theta}) \cdot \nabla_\eta X_T^{t,\theta}], \quad \theta, \eta \in \mathbb{X}_t, \\ &\text{where} \quad \nabla_\eta X_s^{t,\theta,i} = \eta_s^i + \sum_{j=0}^{d} \int_t^s K_i(s,r) \partial_x V_j^i(X_r^{t,\theta}) \nabla_\eta X_r^{t,\theta,j} \circ dB_r^j, \quad i = 1, \ldots, d_1. \end{aligned} \tag{8.1}$$

We note that above we need for the second derivative of V_j^i to exist so that the Stratonovich integration $\circ$ makes sense. Then we see immediately that $\|\partial_{\mathbf{x}} u(t,\theta)\| \leq C e^{C_N^m (T-t)}$. By similar arguments, we can prove the results for higher order derivatives. In particular, we note that the $(N-1)$th derivative of u would involve the Nth derivative of V_j^i. ∎

Proof of Proposition 2.5. We first note that, by Proposition 2.4, $u(t_0^m, \cdot) \in C^{N+2}(\mathbb{X}_{t_0^m})$, so the right-hand side of (2.13) makes sense.

When $N = 1, 2$, one may verify easily that (2.13) reduces to (2.10), (2.11), respectively. Assume (2.13) holds for $N-1$. For $\vec{i} \in \mathcal{I}_N$, $\vec{j} \in \mathcal{J}_N$, $\vec{\kappa} \in \mathcal{S}_N$, $\vec{t} \in \mathbb{T}_N$, we have

$$
\begin{aligned}
&\Delta_{N+1}(\vec{i},\vec{j},\vec{\kappa};\vec{t}) := \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta^{\vec{t}}_{t_N},\Theta^{[t_0^m,T]}_{t_N}) - \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta^{\vec{t}}_{T_m},\Theta^{[t_0^m,T]}_{T_m})\\
&= \prod_{\alpha=1}^{N} \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{t_\alpha}_{t_N})\mathcal{K}_+(\vec{i},\vec{\kappa};\vec{t})\,\langle \partial^{\vec{\kappa}}_{\vec{i}} u(t_0^m,\Theta^{[t_0^m,T]}_{t_N}),\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t})\rangle\\
&\quad - \prod_{\alpha=1}^{N} \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{t_\alpha}_{T_m})\mathcal{K}_+(\vec{i},\vec{\kappa};\vec{t})\,\langle \partial^{\vec{\kappa}}_{\vec{i}} u(t_0^m,\Theta^{[t_0^m,T]}_{T_m}),\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t})\rangle.
\end{aligned}
$$

Note that, for $s\in[T_m,t_N]$, by Itô's formula and Proposition 2.3 we have

$$
\begin{aligned}
&d\partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{t_\alpha}_s) = \sum_{\tilde{i}=1}^{d_1}\sum_{\tilde{j}=0}^{d} \partial_{x_{\tilde{i}}}[\partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}](\Theta^{t_\alpha}_s)K_{\tilde{i}}(t_\alpha,\tilde{t})V^{\tilde{i}}_{\tilde{j}}(X_s)\circ dB^{\tilde{j}}_s,\\
&d\langle \partial^{\vec{\kappa}}_{\vec{i}} u(t_0^m,\Theta^{[t_0^m,T]}_s),\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t})\rangle\\
&= \sum_{\tilde{i}=1}^{d_1}\sum_{\tilde{j}=0}^{d} \langle \partial_{\mathbf{x}_{\tilde{i}}}\partial^{\vec{\kappa}}_{\vec{i}} u(t_0^m,\Theta^{[t_0^m,T]}_s),(\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t}),K^{[t_0^m,T]}_{\tilde{i},s})\rangle V^{\tilde{i}}_{\tilde{j}}(X_s)\circ dB^{\tilde{j}}_s.
\end{aligned}
\tag{8.2}
$$

Then, by Itô's formula, we have, for given $\vec{i}\in\mathcal{I}_N$, $\vec{j}\in\mathcal{J}_N$, $\vec{t}\in\mathbb{T}_N$,

$$
\begin{aligned}
&\Delta_{N+1}(\vec{i},\vec{j},\vec{\kappa};\vec{t}) = \sum_{\tilde{i}=1}^{d_1}\sum_{\tilde{j}=0}^{d}\int_{T_m}^{t_N} \Xi_{N+1}((\vec{i},\tilde{i}),(\vec{j},\tilde{j}),\vec{\kappa};(\vec{t},s))\circ dB^{\tilde{j}}_s, \quad \text{where}\\
&\Xi_{N+1}((\vec{i},\tilde{i}),(\vec{j},\tilde{j}),\vec{\kappa};(\vec{t},s)) := \sum_{\tilde{\alpha}=1}^{N}\prod_{\alpha\in\{1,\dots,N\}\setminus\{\tilde{\alpha}\}} \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{t_\alpha}_s)\\
&\times \partial_{x_{\tilde{i}}}[\partial^{\vec{\kappa},\tilde{\alpha}}_{\vec{i}} V^{i_{\tilde{\alpha}}}_{j_{\tilde{\alpha}}}](\Theta^{t_{\tilde{\alpha}}}_s)V^{\tilde{i}}_{\tilde{j}}(\Theta^s_s)\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})K_{\tilde{i}}(t_{\tilde{\alpha}},s)\,\langle \partial^{\vec{\kappa}}_{\vec{i}} u(t_0^m,\Theta^{[t_0^m,T]}_s),\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t})\rangle\\
&+ \prod_{\alpha=1}^{N} \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{t_\alpha}_s)V^{\tilde{i}}_{\tilde{j}}(\Theta^s_s)\mathcal{K}(\vec{i},\vec{\kappa};\vec{t})\,\langle \partial_{\mathbf{x}_{\tilde{i}}}\partial^{\vec{\kappa}}_{\vec{i}} u(t_0^m,\Theta^{[t_0^m,T]}_s),(\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};\vec{t}),K^{[t_0^m,T]}_{\tilde{i},s})\rangle.
\end{aligned}
\tag{8.3}
$$

Thus, since (2.13) holds for $N-1$ by induction assumption, we have

$$
\begin{aligned}
&u(t_0^m,\Theta^{[t_0^m,T]}_{t_0^m}) - u(t_0^m,\Theta^{[t_0^m,T]}_{T_m})\\
&= \sum_{n=1}^{N-1}\sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_n,\vec{\kappa}\in\mathcal{S}_n}\int_{\mathbb{T}^m_n} \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta^{\vec{t}}_{T_m},\Theta^{[t_0^m,T]}_{T_m})\circ dB^{\vec{j}}_{\vec{t}}\\
&\quad + \sum_{\vec{i}\in\mathcal{I}_N,\vec{j}\in\mathcal{J}_N,\vec{\kappa}\in\mathcal{S}_N}\int_{\mathbb{T}^m_N} \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta^{\vec{t}}_{t_N},\Theta^{[t_0^m,T]}_{t_N})\circ dB^{\vec{j}}_{\vec{t}}\\
&= \sum_{n=1}^{N}\sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_n,\vec{\kappa}\in\mathcal{S}_n}\int_{\mathbb{T}^m_n} \mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta^{\vec{t}}_{T_m},\Theta^{[t_0^m,T]}_{T_m})\circ dB^{\vec{j}}_{\vec{t}}\\
&\quad + \sum_{(\vec{i},\tilde{i})\in\mathcal{I}_{N+1},(\vec{j},\tilde{j})\in\mathcal{J}_{N+1},\vec{\kappa}\in\mathcal{S}_N}\int_{\mathbb{T}^m_{N+1}} \Xi_{N+1}((\vec{i},\tilde{i}),(\vec{j},\tilde{j}),\vec{\kappa};(\vec{t},s))\circ dB^{(\vec{j},\tilde{j})}_{(\vec{t},s)}.
\end{aligned}
$$

Then it suffices to show that

$$
\Xi_{N+1}((\vec{i},\tilde{i}),(\vec{j},\tilde{j}),\vec{\kappa};(\vec{t},s)) = \sum_{\tilde{\alpha}=0}^{N}\mathcal{V}((\vec{i},\tilde{i}),(\vec{j},\tilde{j}),(\vec{\kappa},\tilde{\alpha});(\vec{t},s),\Theta^{(\vec{t},s)}_s,\Theta^{[t_0^m,T]}_s).
\tag{8.4}
$$

Indeed, recalling (2.12), one can verify that

$$
\begin{aligned}
&\Xi_{N+1}((\vec{i},\tilde{i}),(\vec{j},\tilde{j}),\vec{\kappa};(\vec{t},s))\\
&=\sum_{\tilde{\alpha}=1}^{N}\prod_{\alpha=1}^{N}\partial_{(\vec{i},\tilde{i})}^{(\vec{\kappa},\tilde{\alpha}),\alpha}V_{j_\alpha}^{i_\alpha}(\Theta_s^{t_\alpha})V_{\tilde{j}}^{\tilde{i}}(\Theta_s^{s})\mathcal{K}((\vec{i},\tilde{i}),(\vec{\kappa},\tilde{\alpha});(\vec{t},s))\\
&\qquad\qquad\times\langle\partial_{(\vec{i},\tilde{i})}^{(\vec{\kappa},\tilde{\alpha})}u(t_0^m,\Theta_s^{[t_0^m,T]}),\vec{\mathcal{K}}_0((\vec{i},\tilde{i}),(\vec{\kappa},\tilde{\alpha});(\vec{t},s))\rangle\\
&\quad+\prod_{\alpha=1}^{N}\partial_{(\vec{i},\tilde{i})}^{(\vec{\kappa},0),\alpha}V_{j_\alpha}^{i_\alpha}(\Theta_s^{t_\alpha})V_{\tilde{j}}^{\tilde{i}}(\Theta_s^{s})\mathcal{K}((\vec{i},\tilde{i}),(\vec{\kappa},0);(\vec{t},s))\\
&\qquad\qquad\times\langle\partial_{(\vec{i},\tilde{i})}^{(\vec{\kappa},0)}u(t_0^m,\Theta_s^{[t_0^m,T]}),(\vec{\mathcal{K}}_0((\vec{i},\tilde{i}),(\vec{\kappa},0);(\vec{t},s))\rangle\\
&=\sum_{\tilde{\alpha}=0}^{N}\prod_{\alpha=1}^{N+1}\partial_{(\vec{i},\tilde{i})}^{(\vec{\kappa},\tilde{\alpha}),\alpha}V_{j_\alpha}^{i_\alpha}(\Theta_s^{t_\alpha})\mathcal{K}((\vec{i},\tilde{i}),(\vec{\kappa},\tilde{\alpha});(\vec{t},s))\\
&\qquad\qquad\times\langle\partial_{(\vec{i},\tilde{i})}^{(\vec{\kappa},\tilde{\alpha})}u(t_0^m,\Theta_s^{[t_0^m,T]}),\vec{\mathcal{K}}_0((\vec{i},\tilde{i}),(\vec{\kappa},\tilde{\alpha});(\vec{t},s))\rangle\\
&=\sum_{\tilde{\alpha}=0}^{N}\mathcal{V}((\vec{i},\tilde{i}),(\vec{j},\tilde{j}),(\vec{\kappa},\tilde{\alpha});(\vec{t},s),\Theta_s^{(\vec{t},s)},\Theta_s^{[t_0^m,T]}).
\end{aligned}
$$

This proves (8.4), and hence (2.13) for N. ■

Proof of Lemma 2.6. Recall that $B_t^0=t$; the case $j_1=0$ is obvious. We now assume $j_1>0$. Fix $T_m\le s\le t_0^m$, and denote

$$
\psi(t_1,t_2):=\int_{\mathbb{T}_{n-2}^m(t_2)}\varphi(t_1,t_2,\vec{t}_{-2})\circ dB_{\vec{t}_{-2}}^{\vec{j}_{-2}}.
$$

Then, when $j_2\neq j_1$, we have

$$
\int_{\mathbb{T}_n^m(s)}\varphi(\vec{t})\circ dB_{\vec{t}}^{\vec{j}}=\int_{T_m}^{s}\left[\int_{T_m}^{t_1}\psi(t_1,t_2)\circ dB_{t_2}^{j_2}\right]\circ dB_{t_1}^{j_1}=\int_{T_m}^{s}\left[\int_{T_m}^{t_1}\psi(t_1,t_2)\circ dB_{t_2}^{j_2}\right]dB_{t_1}^{j_1},
$$

coinciding with the Itô integral, and when $j_1=j_2$,

$$
\int_{\mathbb{T}_n^m(s)}\varphi(\vec{t})\circ dB_{\vec{t}}^{\vec{j}}=\int_{T_m}^{s}\left[\int_{T_m}^{t_1}\psi(t_1,t_2)\circ dB_{t_2}^{j_2}\right]dB_{t_1}^{j_1}+\frac{1}{2}\int_{T_m}^{s}\psi(t_1,t_1)dt_1.
$$

Then one may verify (2.16) and (2.17) straightforwardly. ■

Proof of Theorem 2.7. First, similar to (2.11) and (2.13), one can verify that

$$
\begin{aligned}
R_N^m=\sum_{n=1}^{N+1}\sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_n}&[\mathbf{1}_{\{\|\vec{j}\|=N+1\}}+\mathbf{1}_{\{\|\vec{j}_{-1}\|=N,j_1=0\}}]\\
&\times\int_{\mathbb{T}_n^m}\sum_{\vec{\kappa}\in\mathcal{S}_n}\mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta_{t_n}^{\vec{t}},\Theta_{t_n}^{[t_0^m,T]})\circ dB_{\vec{t}}^{\vec{j}}.
\end{aligned}\tag{8.5}
$$

Fix n, $\vec{i}\in\mathcal{I}_n$, $\vec{j}\in\mathcal{J}_n$, and denote $\varphi(\vec{t}):=\sum_{\vec{\kappa}\in\mathcal{S}_n}\mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta_{t_n}^{\vec{t}},\Theta_{t_n}^{[t_0^m,T]})$ for $\vec{t}\in\mathbb{T}_n^m$. By (2.17) and (2.18), we may prove by induction that, for $s\in[T_m,t_0^m]$, $l=1,\dots,n-1$, and for some constant C_n which may depend on n,

$$\|\varphi(\cdot)\|^2_{s,\vec{j}} \le C_n \delta^{\|\vec{j}_l\|} \operatorname*{ess\,sup}_{T_m \le s_l \le \cdots \le s_1 \le s} \|\varphi(s_1,\ldots,s_l,\cdot)\|^2_{\vec{j}_{-l},s_l}$$
$$+ C_n \delta^{\|\vec{j}_{l+1}\|} \operatorname*{ess\,sup}_{T_m \le s_{l+1} \le \cdots \le s_1 \le s} \|\varphi(s_1,\ldots,s_{l+1},\cdot)\|^2_{\vec{j}_{-l-1},s_{l+1}}.$$

In particular, by setting $l = n-1$ we have

$$\mathbb{E}_m\left[\left|\int_{\mathbb{T}^m_n} \varphi(\vec{t}) \circ dB^{\vec{j}}_{\vec{t}}\right|^2\right] \le C_n[\Lambda_{n-1} + \Lambda_n], \quad \text{where} \tag{8.6}$$
$$\Lambda_{n-1} := \delta^{\|\vec{j}_{n-1}\|} \operatorname*{ess\,sup}_{\vec{s}\in\mathbb{T}^m_{n-1}} \mathbb{E}_m\left[\left|\int_{T_m}^{s_{n-1}} \varphi(\vec{s},t_n) \circ dB^{j_n}_{t_n}\right|^2\right], \quad \Lambda_n := \delta^{\|\vec{j}\|} \operatorname*{ess\,sup}_{\vec{s}\in\mathbb{T}^m_n} \mathbb{E}_m\left[|\varphi(\vec{s})|^2\right].$$

By (2.18), we have $\Lambda_n \le |A^m_{\|\vec{j}\|}|^2 \delta^{\|\vec{j}\|}$. Moreover, fix $\vec{s} \in \mathbb{T}_{n-1}$. When $j_n = 0$, we have

$$\mathbb{E}_m\left[\left|\int_{T_m}^{s_{n-1}} \varphi(\vec{s},t_n) \circ dB^{j_n}_{t_n}\right|^2\right] = \mathbb{E}_m\left[\left|\int_{T_m}^{s_{n-1}} \varphi(\vec{s},t_n) dt_n\right|^2\right]$$
$$\le \delta \mathbb{E}_m\left[\int_{T_m}^{s_{n-1}} |\varphi(\vec{s},t_n)|^2 dt_n\right] \le \delta^2 \operatorname*{ess\,sup}_{T_m \le t_n \le s_{n-1}} \mathbb{E}_m\left[|\varphi(\vec{s},t_n)|^2\right] \le \delta^2 |A^m_{\|\vec{j}\|}|^2.$$

Then

$$\Lambda_{n-1} \le \delta^{\|\vec{j}_{n-1}\|} \delta^2 |A^m_{\|\vec{j}\|}|^2 = |A^m_{\|\vec{j}\|}|^2 \delta^{\|\vec{j}\|}.$$

When $j_n > 0$, recalling (2.12) and by (8.2) we have the following: denoting $s_n := t_n$, we have

$$\int_{T_m}^{s_{n-1}} \varphi(\vec{s},t_n) \circ dB^{j_n}_{t_n}$$
$$= \int_{T_m}^{s_{n-1}} \sum_{\vec{\kappa}\in\mathcal{S}_n} \prod_{\alpha=1}^n \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{s_\alpha}_{t_n}) \mathcal{K}_+(\vec{i},\vec{\kappa};(\vec{s},t_n)) \langle \partial^{\vec{\kappa}}_{\vec{i}} u(t^m_0, \Theta^{[t^m_0,T]}_{t_n}), \vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};(\vec{s},t_n))\rangle \circ dB^{j_n}_{t_n}$$
$$= \int_{T_m}^{s_{n-1}} \sum_{\vec{\kappa}\in\mathcal{S}_n} \prod_{\alpha=1}^n \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{s_\alpha}_{t_n}) \mathcal{K}_+(\vec{i},\vec{\kappa};(\vec{s},t_n)) \langle \partial^{\vec{\kappa}}_{\vec{i}} u(t^m_0, \Theta^{[t^m_0,T]}_{t_n}), \vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};(\vec{s},t_n))\rangle dB^{j_n}_{t_n}$$
$$+ \frac{1}{2}\int_{T_m}^{s_{n-1}} \sum_{\vec{\kappa}\in\mathcal{S}_n} \Bigg[\sum_{\tilde{\alpha}=1}^n \prod_{\alpha\ne\tilde{\alpha}} \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{s_\alpha}_{t_n}) \mathcal{K}_+(\vec{i},\vec{\kappa};(\vec{s},t_n)) \langle \partial^{\vec{\kappa}}_{\vec{i}} u(t^m_0, \Theta^{[t^m_0,T]}_{t_n}), \vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};(\vec{s},t_n))\rangle$$
$$\times \sum_{\tilde{i}=1}^{d_1} [\partial_{x_{\tilde{i}}} \partial^{\vec{\kappa},\tilde{\alpha}}_{\vec{i}} V^{i_{\tilde{\alpha}}}_{j_{\tilde{\alpha}}}(\Theta^{s_{\tilde{\alpha}}}_{t_n}) K_{\tilde{i}}(s_{\tilde{\alpha}},t_n) V^{\tilde{i}}_{j_n}(X_{t_n})]$$
$$+ \prod_{\alpha=1}^n \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{s_\alpha}_{t_n}) \mathcal{K}_+(\vec{i},\vec{\kappa};(\vec{s},t_n))$$
$$\times \sum_{\tilde{i}=1}^{d_1} \langle \partial_{\mathbf{x}_{\tilde{i}}} \partial^{\vec{\kappa}}_{\vec{i}} u(t^m_0, \Theta^{[t^m_0,T]}_{t_n}), (\vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};(\vec{s},t_n)), K^{[t^m_0,T]}_{\tilde{i},t_n})\rangle V^{\tilde{i}}_{j_n}(X_{t_n})\Bigg] dt_n$$
$$= \int_{T_m}^{s_{n-1}} \sum_{\vec{\kappa}\in\mathcal{S}_n} \prod_{\alpha=1}^n \partial^{\vec{\kappa},\alpha}_{\vec{i}} V^{i_\alpha}_{j_\alpha}(\Theta^{s_\alpha}_{t_n}) \mathcal{K}_+(\vec{i},\vec{\kappa};(\vec{s},t_n)) \langle \partial^{\vec{\kappa}}_{\vec{i}} u(t^m_0, \Theta^{[t^m_0,T]}_{t_n}), \vec{\mathcal{K}}_0(\vec{i},\vec{\kappa};(\vec{s},t_n))\rangle dB^{j_n}_{t_n}$$
$$+ \frac{1}{2}\sum_{\tilde{i}=1}^{d_1} \int_{T_m}^{s_{n-1}} \sum_{\vec{\kappa}\in\mathcal{S}_n} \sum_{\tilde{\alpha}=0}^n \mathcal{V}((\vec{i},\tilde{i}),(\vec{j},j_n),(\vec{\kappa},\tilde{\alpha});(\vec{s},t_n,t_n),\Theta^{(\vec{s},t_n,t_n)}_{t_n},\Theta^{[t^m_0,T]}_{t_n}) dt_n,$$

where, similarly to (8.3), the last equality can be verified straightforwardly. Then

$$\mathbb{E}_m\Big[\Big|\int_{T_m}^{s_{n-1}}\varphi(\vec{s},t_n)\circ dB_{t_n}^{j_n}\Big|^2\Big]\le C\delta|A^m_{\|\vec{j}\|}|^2+C\delta^2|A^m_{\|(\vec{j},j_n)\|}|^2.$$

Thus,

$$\Lambda_{n-1}\le C\delta^{\|\vec{j}_{n-1}\|}[\delta|A^m_{\|\vec{j}\|}|^2+\delta^2|A^m_{\|(\vec{j},j_n)\|}|^2]=C[|A^m_{\|\vec{j}\|}|^2\delta^{\|\vec{j}\|}+|A^m_{\|\vec{j}\|+1}|^2\delta^{\|\vec{j}\|+1}].$$

So in all cases, by (8.6) we have

$$\mathbb{E}_m\Big[\Big|\int_{\mathbb{T}_n^m}\varphi(\vec{t})\circ dB_{\vec{t}}^{\vec{j}}\Big|^2\Big]\le C_n\Big[|A^m_{\|\vec{j}\|}|^2\delta^{\|\vec{j}\|}+|A^m_{\|\vec{j}\|+1}|^2\delta^{\|\vec{j}\|+1}\Big].$$

Then, by (8.5), we obtain

$$\mathbb{E}_m[|R_N^m|^2]\le C_N\sum_{n=1}^{N+1}\sum_{\vec{i}\in\mathcal{I}_n,\vec{j}\in\mathcal{J}_n}[\mathbf{1}_{\{\|\vec{j}\|=N+1\}}+\mathbf{1}_{\{\|\vec{j}_{-1}\|=N,j_1=0\}}]\Big[|A^m_{\|\vec{j}\|}|^2\delta^{\|\vec{j}\|}+|A^m_{\|\vec{j}\|+1}|^2\delta^{\|\vec{j}\|+1}\Big].$$

This implies (2.20) immediately. ■

Proof of Theorem 3.3. It is clear that the functional Itô formula (2.7) holds true under Q as well; then (3.5) and (8.5) also hold true under Q. Thus, by (3.9),

$$|Y_0-Y_0^Q|=\Big|\mathbb{E}[R_N]-\mathbb{E}^Q[R_N]\Big|.$$

Therefore, by Theorem 2.7, it suffices to show that

$$|\mathbb{E}^Q[R_N]|\le C_N\Big[A_{N+1}T^{\frac{N+1}{2}}(1+C_Q^{N-1})+A_{N+2}T^{\frac{N+2}{2}}(1+C_Q^{N-2})\Big]. \tag{8.7}$$

Now, for each n and $\vec{i}\in\mathcal{I}_n$, $\vec{j}\in\mathcal{J}_n$ as in (8.5), note that

$$(\omega_{k,t}^0)'=1,\quad |\{l:j_l=0\}|=\|\vec{j}\|-n,\quad |\{l:j_l>0\}|=2n-\|\vec{j}\|,$$

where ω' denotes the time derivative of ω. Then

$$\begin{aligned}\Big|\mathbb{E}^Q\Big[\int_{\mathbb{T}_n}\sum_{\vec{\kappa}\in\mathcal{S}_n}\mathcal{V}(\vec{i},\vec{j},\vec{\kappa};\vec{t},\Theta_{t_n}^{\vec{t}},\Theta_{t_n}^{T})\circ dB_{\vec{t}}^{\vec{j}}\Big]\Big|&\le A_{\|\vec{j}\|}\sum_{k=1}^{2W}\lambda_k\int_{\mathbb{T}_n}\prod_{l=1}^{n}|(\omega_{k,t_l}^{j_l})'|dt_n\cdots dt_1\\&\le\Big|C_nA_{\|\vec{j}\|}\sum_{k=1}^{2W}\lambda_k\Big(\frac{C_Q}{\sqrt{T}}\Big)\Big|^{2n-\|\vec{j}\|}T^n=C_nA_{\|\vec{j}\|}C_Q^{2n-\|\vec{j}\|}T^{\|\vec{j}\|}.\end{aligned} \tag{8.8}$$

If $\|\vec{j}\|=N+1$, we have $2n\ge N+1$. Denote $k:=2n-(N+1)$; then $0\le k\le N-1$, and

$$A_{\|\vec{j}\|}|C_Q|^{2n-\|\vec{j}\|}T^{\|\vec{j}\|}=A_{N+1}T^{\frac{N+1}{2}}C_Q^k.$$

If $\|\vec{j}_{-1}\| = N$, $j_1 = 0$, then $2n \geq N+2$. Denote $k := 2n - (N+2)$; then $0 \leq k \leq N-2$, and

$$A_{\|\vec{j}\|} C_Q^{2n-\|\vec{j}\|} T^{\|\vec{j}\|} = A_{N+2} T^{\frac{N+2}{2}} C_Q^k.$$

Plugging these into (8.8), we obtain (8.7) immediately, and hence (3.12) holds true.

Finally, when (3.13) holds true, by (3.9) one can easily see that Q in (3.6) is an N-Volterra cubature formula on $[0,T]$ if and only if the following rescaled one $\tilde{Q}$ is an N-Volterra cubature formula on $[0,1]$:

$$\tilde{Q} = \sum_{k=1}^{2W} \lambda_k \delta_{\tilde{\omega}_t}, \qquad d\tilde{\omega}_{k,t} = a_{k,l} dt,\, t \in (\tilde{s}_{l-1}, \tilde{s}_l], \qquad \tilde{s}_l := \frac{l}{L}, \quad l = 0, \dots, L. \tag{8.9}$$

In particular, this implies that $C_Q = C_{\tilde{Q}}$ is independent of T, ■

Proof of Theorem 4.5. Note that

$$\begin{aligned}
&\left|\mathbb{E}_m^{Q_m}[u(T_{m+1}, \Theta_{T_{m+1}}^{[T_{m+1},T]}] - \mathbb{E}_m[u(T_{m+1}, \Theta_{T_{m+1}}^{[T_{m+1},T]}]\right| \\
&\leq \left|\mathbb{E}_m^{Q_m}[I_N^m] - \mathbb{E}_m[I_N^m]\right| + |\mathbb{E}_m^{Q_m}[R_N^m]| + |\mathbb{E}_m[R_N^m]|.
\end{aligned}$$

Then, following the same arguments as in Theorem 3.3, it suffices to provide the desired estimate for $\left|\mathbb{E}_m^{Q_m}[I_N^m] - \mathbb{E}_m[I_N^m]\right|$. Similar to Lemma 3.4, by the desired symmetric properties, for any $\vec{j} \in \mathcal{J}_{n,N} \backslash \bar{\mathcal{J}}_{n,N}$ we have

$$\mathbb{E}_m^{Q_m}\left[\int_{\mathbb{T}_n^m} \mathcal{V}(\vec{i}, \vec{j}, \vec{\kappa}; \vec{t}, \Theta_{T_m}^{\vec{t}}, \Theta_{T_m}^{[t_0^m,T]}) \circ dB_{\vec{t}}^{\vec{j}}\right] = 0 = \mathbb{E}_m\left[\int_{\mathbb{T}_n^m} \mathcal{V}(\vec{i}, \vec{j}, \vec{\kappa}; \vec{t}, \Theta_{T_m}^{\vec{t}}, \Theta_{T_m}^{[t_0^m,T]}) \circ dB_{\vec{t}}^{\vec{j}}\right].$$

Then, by (4.8), we have

$$\left|\mathbb{E}_m^{Q_m}[I_N^m] - \mathbb{E}_m[I_N^m]\right| = \left|\mathbb{E}_m^{Q_m}[\check{R}_N^m] - \mathbb{E}_m[\check{R}_N^m]\right| \leq \left|\mathbb{E}_m^{Q_m}[\check{R}_N^m]\right| + \left|\mathbb{E}_m[\check{R}_N^m]\right|.$$

The estimate for $\left|\mathbb{E}_m[\check{R}_N^m]\right|$ is implied by (4.6). Moreover, for each $n \leq N$ and $\vec{j} \in \bar{\mathcal{J}}_{n,N}$, again set $k := \frac{N-\|\vec{j}\|-1}{2}$ as in the proof of Theorem 4.2. Then, by (4.4) and similar to (8.8),

$$\left|\mathbb{E}_m^{Q_m}\left[\int_{\mathbb{T}_n^m} \check{R}_N^m(\vec{t}) \circ dB_{\vec{t}}^{\vec{j}}\right]\right| \leq C_N^m C_{Q_m}^{2n-\|\vec{j}\|} e^{C_N^m T} \delta^{\frac{N+1}{2}}.$$

Thus, since $0 \leq 2n - \|\vec{j}\| \leq N-1$,

$$\left|\mathbb{E}_m^{Q_m}[\check{R}_N^m]\right| \leq \sum_{n=1}^{N} \sum_{\vec{j} \in \bar{\mathcal{J}}_{n,N}} C_N^m C_{Q_m}^{2n-\|\vec{j}\|} e^{C_N^m T} \delta^{\frac{N+1}{2}} \leq C_N^m (1 + C_{Q_m}^{N-1}) e^{C_N^m T} \delta^{\frac{N+1}{2}}.$$

Recalling $C_{Q_m} = C_{Q_N^*}$, this is the desired estimate. ■

8.1. The one period cubature formula for (6.1) with $N = 5$. Denote

$$\gamma_1 := \frac{T^{H_+ + 1}}{H_+(H_+ + 1)}, \quad \gamma_2 := \frac{T^{2H_+}}{8HH_+}. \tag{8.10}$$

First, similar to (3.2) we have

$$\begin{aligned}\tilde{b}_1(\Theta^{t_1}_{t_1}) = \quad & R_{\neq 2} + \int_0^{t_1} \Big[\partial_0 \tilde{b}_1(\Theta^{t_1}_{t_2}) + \sum_{i=1}^{2} \partial_i \tilde{b}_1(\Theta^{t_1}_{t_2}) b_i(\Theta^{t_2}_{t_2}) K_i(t_1, t_2)\Big] dt_2 \\ & + \int_0^{t_1}\int_0^{t_2} \sum_{i,j=1}^{2} \Big[\partial_{ji} \tilde{b}_1(\Theta^{t_1}_{t_3}) \sigma_i(\Theta^{t_2}_{t_3}) K_j(t_1, t_3) \\ & + \partial_i \tilde{b}_1(\Theta^{t_1}_{t_3}) \partial_j \sigma_i(\Theta^{t_2}_{t_3}) K_j(t_2, t_3)\Big] \sigma_j(\Theta^{t_3}_{t_3}) K_i(t_1, t_2) \circ dB^j_{t_3} \circ dB^i_{t_2}.\end{aligned} \tag{8.11}$$

We remark that $\partial_i \tilde{b}_1 = G'' b_1 + G' \partial_i b_1$; however, thanks to the new convention in Remark 6.1, we do not need to consider $G'' b_1$ and $G' \partial_i b_1$ separately. Then we can easily see that the terms in $\Delta \tilde{\mathbb{V}}^0_4$ derived from $\tilde{b}_1$ consist of the following stochastic integrals:

$$\begin{aligned}&\Delta \tilde{\mathbb{V}}^0_{4,b} := \Big\{\tilde{\Gamma}^{(i,j,k)}_{4,b} : (i,j,k) = (1,1,1), (1,2,1), (1,2,2), (2,1,1), (2,2,1), (2,2,2)\Big\}, \\ &\qquad \text{where} \quad \tilde{\Gamma}^{(i,j,k)}_{4,b} := \int_{\mathbb{T}_3} K_i(t_1, t_2) K_j(t_k, t_3) \circ dB^j_{t_3} \circ dB^i_{t_2} dt_1.\end{aligned} \tag{8.12}$$

Note that, for $i, j, k = 1, 2$,

$$\mathbb{E}[\tilde{\Gamma}^{(i,j,k)}_{4,b}] = \frac{1}{2}\int_{\mathbb{T}_2} K_i(t_1, t_2) K_j(t_k, t_2) d\langle B^i, B^j\rangle_{t_2} dt_1 = \begin{cases} \frac{T^2}{4}, & (i,j,k) = (1,1,1), \\ \frac{\rho\gamma_1}{2}, & (i,j,k) = (1,2,1), (2,1,1), \\ \gamma_2, & (i,j,k) = (2,2,1), \\ 0, & (i,j,k) = (1,2,2), (2,2,2). \end{cases} \tag{8.13}$$

Similarly, we may have the expansion of $\tilde{\sigma}_1(\Theta^{t_1}_{t_1})$ as in (8.11):

$$\begin{aligned}\tilde{\sigma}_1(\Theta^{t_1}_{t_1}) &= R_{\neq 3} + \sum_{i_2=1}^{2} \int_0^{t_1}\int_0^{t_2} \partial_{i_2 0} \tilde{\sigma}_1(\Theta^{t_1}_{t_3}) \sigma_{i_2}(\Theta^{t_3}_{t_3}) K_{i_2}(t_1, t_3) \circ dB^{i_2}_{t_3} dt_2 \\ &+ \sum_{i_1,i_2=1}^{2} \int_0^{t_1}\int_0^{t_2} \Big[\partial_{i_2 i_1} \tilde{\sigma}_1(\Theta^{t_1}_{t_3}) b_{i_1}(\Theta^{t_2}_{t_3}) K_{i_2}(t_1, t_3) \\ &+ \partial_{i_1} \tilde{\sigma}_1(\Theta^{t_1}_{t_3}) \partial_{i_2} b_{i_1}(\Theta^{t_2}_{t_3}) K_{i_2}(t_2, t_3)\Big] \sigma_{i_2}(\Theta^{t_3}_{t_3}) K_{i_1}(t_1, t_2) \circ dB^{i_2}_{t_3} dt_2 \\ &+ \sum_{i_1=1}^{2} \int_0^{t_1}\int_0^{t_2} \Big[\partial_{0 i_1} \tilde{\sigma}_1(\Theta^{t_1}_{t_3}) \sigma_{i_1}(\Theta^{t_2}_{t_3}) + \partial_{i_1} \tilde{\sigma}_1(\Theta^{t_1}_{t_3}) \partial_0 \sigma_{i_1}(\Theta^{t_2}_{t_3})\Big] K_{i_1}(t_1, t_2) dt_3 \circ dB^{i_1}_{t_2} \\ &+ \sum_{i_1,i_2=1}^{2} \int_0^{t_1}\int_0^{t_2} \Big[\partial_{i_2 i_1} \tilde{\sigma}_1(\Theta^{t_1}_{t_3}) b_{i_2}(\Theta^{t_3}_{t_3}) \sigma_{i_1}(\Theta^{t_2}_{t_3}) K_{i_2}(t_1, t_3) \\ &+ \partial_{i_1} \tilde{\sigma}_1(\Theta^{t_1}_{t_3}) \partial_{i_2} \sigma_{i_1}(\Theta^{t_2}_{t_3}) b_{i_2}(\Theta^{t_3}_{t_3}) K_{i_2}(t_2, t_3)\Big] K_{i_1}(t_1, t_2) dt_3 \circ dB^{i_1}_{t_2} + \xi,\end{aligned} \tag{8.14}$$

where the terms presented above involve ∂_t, and ξ contains the terms without ∂_t:

$$
\begin{aligned}
\xi := \sum_{i_1,i_2,i_3=1}^{2} \int_0^{t_1}\int_0^{t_2}\int_0^{t_3} \Big[&\partial_{i_3 i_2 i_1}\tilde{\sigma}_1(\Theta_{t_4}^{t_1})\sigma_{i_1}(\Theta_{t_4}^{t_2})\sigma_{i_2}(\Theta_{t_4}^{t_3})K_{i_2}(t_1,t_3)K_{i_3}(t_1,t_4)\\
&+\partial_{i_2 i_1}\tilde{\sigma}_1(\Theta_{t_4}^{t_1})\partial_{i_3}\sigma_{i_1}(\Theta_{t_4}^{t_2})\sigma_{i_2}(\Theta_{t_4}^{t_3})K_{i_2}(t_1,t_3)K_{i_3}(t_2,t_4)\\
&+\partial_{i_2 i_1}\tilde{\sigma}_1(\Theta_{t_4}^{t_1})\sigma_{i_1}(\Theta_{t_4}^{t_2})\partial_{i_3}\sigma_{i_2}(\Theta_{t_4}^{t_3})K_{i_2}(t_1,t_3)K_{i_3}(t_3,t_4)\\
&+\partial_{i_3 i_1}\tilde{\sigma}_1(\Theta_{t_4}^{t_1})\partial_{i_2}\sigma_{i_1}(\Theta_{t_4}^{t_2})\sigma_{i_2}(\Theta_{t_4}^{t_3})K_{i_2}(t_2,t_3)K_{i_3}(t_1,t_4)\\
&+\partial_{i_1}\tilde{\sigma}_1(\Theta_{t_4}^{t_1})\partial_{i_3 i_2}\sigma_{i_1}(\Theta_{t_4}^{t_2})\sigma_{i_2}(\Theta_{t_4}^{t_3})K_{i_2}(t_2,t_3)K_{i_3}(t_2,t_4)\\
&+\partial_{i_1}\tilde{\sigma}_1(\Theta_{t_4}^{t_1})\partial_{i_2}\sigma_{i_1}(\Theta_{t_4}^{t_2})\partial_{i_3}\sigma_{i_2}(\Theta_{t_4}^{t_3})K_{i_2}(t_2,t_3)K_{i_3}(t_3,t_4)\Big]\\
&\times \sigma_{i_3}(\Theta_{t_4}^{t_4})K_{i_1}(t_1,t_2)\circ dB_{t_4}^{i_3}\circ dB_{t_3}^{i_2}\circ dB_{t_2}^{i_1}.
\end{aligned}
\tag{8.15}
$$

Then we see that the terms in $\Delta\tilde{\mathbb{V}}_4^0$ derived from $\tilde{\sigma}_1$ are

$$
\begin{aligned}
&\Delta\tilde{\mathbb{V}}_{4,\sigma}^0 = \Delta\tilde{\mathbb{V}}_{4,\sigma,0}^0 \cup \Delta\tilde{\mathbb{V}}_{4,\sigma,1}^0 \cup \Delta\tilde{\mathbb{V}}_{4,\sigma,2}^0 \cup \Delta\tilde{\mathbb{V}}_{4,\sigma,3}^0, \quad \text{where, for } l=1,2,\\
&\Delta\tilde{\mathbb{V}}_{4,\sigma,0}^0 := \left\{\tilde{\Gamma}_{4,\sigma,0}^{(1)}, \tilde{\Gamma}_{4,\sigma,0}^{(2)}\right\},\\
&\Delta\tilde{\mathbb{V}}_{4,\sigma,l}^0 := \left\{\tilde{\Gamma}_{4,\sigma,l}^{(i_1,i_2,\kappa_2)} : (i_1,i_2,\kappa_2) = (1,1,1),(1,2,1),(1,2,2),(2,1,1),(2,2,1),(2,2,2)\right\}
\end{aligned}
\tag{8.16}
$$

and

$$
\begin{aligned}
\Delta\tilde{\mathbb{V}}_{4,\sigma,3}^0 := \Big\{\tilde{\Gamma}_{4,\sigma,3}^{(i_1,i_2,i_3,\kappa_2,\kappa_3)} : (i_1,i_2,i_3,\kappa_2,\kappa_3) = &(1,1,1,1,1),(1,1,2,1,1),\\
&(1,1,2,1,2),(1,1,2,1,3),(1,2,1,1,1),(1,2,1,2,1),(1,2,2,1,1),(1,2,2,1,2),\\
&(1,2,2,1,3),(1,2,2,2,1),(1,2,2,2,2),(1,2,2,2,3),(2,1,1,1,1),(2,1,2,1,1),\\
&(2,1,2,1,2),(2,1,2,1,3),(2,2,1,1,1),(2,2,1,2,1),(2,2,2,1,1),(2,2,2,1,2),\\
&(2,2,2,1,3),(2,2,2,2,1),(2,2,2,2,2),(2,2,2,2,3)\Big\},
\end{aligned}
\tag{8.17}
$$

where

$$
\begin{aligned}
\tilde{\Gamma}_{4,\sigma,0}^{(i_2)} &:= \int_{\mathbb{T}_3} K_{i_2}(t_1,t_3)\circ dB_{t_3}^{i_2} dt_2 \circ dB_{t_1}^1,\\
\tilde{\Gamma}_{4,\sigma,1}^{(i_1,i_2,\kappa_2)} &:= \int_{\mathbb{T}_3} K_{i_1}(t_1,t_2)K_{i_2}(t_{\kappa_2},t_3)\circ dB_{t_3}^{i_2} dt_2 \circ dB_{t_1}^1,\\
\tilde{\Gamma}_{4,\sigma,2}^{(i_1,i_2,\kappa_2)} &:= \int_{\mathbb{T}_3} K_{i_1}(t_1,t_2)K_{i_2}(t_{\kappa_2},t_3) dt_3 \circ dB_{t_2}^{i_1} \circ dB_{t_1}^1,\\
\tilde{\Gamma}_{4,\sigma,3}^{(i_1,i_2,i_3,\kappa_2,\kappa_3)} &:= \int_{\mathbb{T}_4} K_{i_1}(t_1,t_2)K_{i_2}(t_{\kappa_2},t_3)K_{i_3}(t_{\kappa_3},t_4) dB_{t_4}^{i_3}\circ dB_{t_3}^{i_2}\circ dB_{t_2}^{i_1}\circ dB_{t_1}^1.
\end{aligned}
$$

Note that, for $\vec{i} = (i_1,i_2,i_3)$ and $\vec{\kappa} = (\kappa_1,\kappa_2,\kappa_3)$ with $\kappa_1 = 1$, $\kappa_2 = 1,2$ and $\kappa_3 = 1,2,3$, we have, recalling (8.10),

$$\mathbb{E}[\tilde{\Gamma}^{(i_2)}_{4,\sigma,0}] = \mathbb{E}[\tilde{\Gamma}^{(i_1,i_2,\kappa_2)}_{4,\sigma,1}] = 0,$$

$$\mathbb{E}[\tilde{\Gamma}^{(i_1,i_2,\kappa_2)}_{4,\sigma,2}] = \frac{1}{2}\int_{\mathbb{T}_1} K_{i_1}(t_1,t_1)\int_0^{t_1} K_{i_2}(t_1,t_3)dt_3 d\langle B^{i_1},B^1\rangle_{t_1}$$

$$= \begin{cases} \frac{T^2}{4}, & (i_1,i_2,\kappa_2)=(1,1,1),\\ \frac{\rho\gamma_1}{2}, & (i_1,i_2,\kappa_2)=(1,2,1),\\ 0, & (i_1,i_2,\kappa_2)=(2,1,1),(2,2,1),(2,2,2),(1,2,2),\end{cases}$$

(8.18) $$\mathbb{E}[\tilde{\Gamma}^{(i_1,i_2,i_3,\kappa_2,\kappa_3)}_{4,\sigma,3}]$$

$$= \frac{1}{4}\int_0^T\int_0^{t_1} K_{i_1}(t_1,t_1)K_{i_2}(t_{\kappa_2},t_3)K_{i_3}(t_{\kappa_3},t_3)|_{t_2=t_1} d\langle B^{i_3},B^{i_2}\rangle_{t_3} d\langle B^{i_1},B^1\rangle_{t_1}$$

$$= \begin{cases} \frac{T^2}{8},\ \vec{i}=(1,1,1),\\ \frac{\rho\gamma_1}{4},\ \vec{i}=(1,2,1),\ \text{or}\ \vec{i}=(1,1,2)\ \text{and}\ \kappa_3=1,2,\\ \frac{\gamma_2}{2},\ \vec{i}=(1,2,2),\ \kappa_3=1,2,\\ 0,\ i_1=2\ \text{or}\ (i_3,\kappa_3)=(2,3).\end{cases}$$

We note that, by (6.12), (8.12), and (8.16)–(8.17),

(8.19) $$|\Delta\tilde{\mathbb{V}}^0_2| = 2 \quad\text{and}\quad |\Delta\tilde{\mathbb{V}}^0_4| = |\Delta\tilde{\mathbb{V}}^0_{4,b}| + |\Delta\tilde{\mathbb{V}}^0_{4,\sigma}| = 6+2+6+6+24 = 44.$$

We now construct a desired Q as in (3.6). Recall $\omega^0_t = t$ and $\omega_{k,t} = (\omega^1_{k,t},\omega^2_{k,t})$. We remark that the correlation between B^1, B^2 affects only the expectations $\mathbb{E}[\tilde{\Gamma}]$ in (8.13) and (8.18), but the expectations under Q in (3.8) remain the same. The integrals against $d\omega$ corresponding to the stochastic integrals in (6.12), (8.12), and (8.16)–(8.17) are the following: for each $k=1,\dots,W$ and for appropriate functions φ,

$$\int_{\mathbb{T}_2}\varphi(\vec{t})d\omega^{i_1}_{k,t_2}d\omega^1_{k,t_1} = \sum_{1\le l_2\le l_1\le m}\frac{a^1_{k,l_1}a^{i_1}_{k,l_2}}{T}\int_{s_{l_1-1}}^{s_{l_1}}\int_{s_{l_2-1}}^{s_{l_2}\wedge t_1}\varphi(\vec{t})dt_2dt_1,$$

$$\int_{\mathbb{T}_3}\varphi(\vec{t})d\omega^{i_2}_{k,t_3}d\omega^{i_1}_{k,t_2}dt_1 = \sum_{1\le l_2\le l_1\le m}\frac{a^{i_1}_{k,l_1}a^{i_2}_{k,l_2}}{T}\int_{s_{l_1-1}}^{s_{l_1}}\int_{s_{l_2-1}}^{s_{l_2}\wedge t_2}\int_{t_2}^{T}\varphi(\vec{t})dt_1dt_3dt_2,$$

(8.20) $$\int_{\mathbb{T}_3}\varphi(\vec{t})d\omega^{i_2}_{k,t_3}dt_2d\omega^1_{k,t_1} = \sum_{1\le l_2\le l_1\le m}\frac{a^1_{k,l_1}a^{i_2}_{k,l_2}}{T}\int_{s_{l_1-1}}^{s_{l_1}}\int_{s_{l_2-1}}^{s_{l_2}\wedge t_1}\int_{t_3}^{t_1}\varphi(\vec{t})dt_2dt_3dt_1,$$

$$\int_{\mathbb{T}_3}\varphi(\vec{t})dt_3d\omega^{i_1}_{k,t_2}d\omega^1_{k,t_1} = \sum_{1\le l_2\le l_1\le m}\frac{a^1_{k,l_1}a^{i_1}_{k,l_2}}{T}\int_{s_{l_1-1}}^{s_{l_1}}\int_{s_{l_2-1}}^{s_{l_2}\wedge t_1}\int_0^{t_2}\varphi(\vec{t})dt_3dt_2dt_1,$$

$$\int_{\mathbb{T}_4}\varphi(\vec{t})d\omega^{i_3}_{k,t_4}d\omega^{i_2}_{k,t_3}d\omega^{i_1}_{k,t_2}d\omega^1_{k,t_1} = \sum_{1\le l_4\le l_3\le l_2\le l_1\le 4}\frac{a^{i_1}_{k,l_1}a^{i_2}_{k,l_2}a^{i_3}_{k,l_3}a^{i_4}_{k,l_3}}{T^2}$$
$$\times\int_{s_{l_1-1}}^{s_{l_1}}\int_{s_{l_2-1}}^{s_{l_2}\wedge t_1}\int_{s_{l_3-1}}^{s_{l_3}\wedge t_2}\int_{s_{l_4-1}}^{s_{l_4}\wedge t_3}\varphi(\vec{t})dt_4dt_3dt_2dt_1.$$

For $N=5$, by (8.19), together with the constraint on λ_k, there will be in total 47 equations. In our example (7.7), however, the coefficients are homogeneous, i.e., independent of t. Then in (8.14) the terms $\partial_{i_2 0}\tilde{\sigma}_1 = 0$ and thus the two terms $\tilde{\Gamma}^{(i_2)}_{4,\sigma,0}\in\Delta\tilde{\mathbb{V}}^0_{4,\sigma,0}$, $i_2=1,2$, in (8.16) are not needed (the terms $\int_{\mathbb{T}_3}K_{i_1}(t_1,t_2)dt_3\circ dB^{i_1}_{t_2}\circ dB^1_{t_1}$ are stilled needed, even though $\partial_{0i_1}\tilde{\sigma}_1 =$

$\partial_0\tilde{\sigma}_1 = 0$ in (8.14), because they appear in the last line of (8.14) as well when $i_2 = 1$). Therefore, we have a total of 45 equations. We thus set $W = 5$ and $L = 4$. Since each two dimensional path involves 8 parameters, plus the parameters λ_k, there will be 45 unknowns. For each $\tilde{\Gamma} \in \Delta\tilde{\mathbb{V}}_2^0 \cup (\Delta\tilde{\mathbb{V}}_4^0 \backslash \Delta\tilde{\mathbb{V}}_{4,\sigma,0}^0)$, by (6.11), (8.11), (8.14), (8.15), one can easily derive from (8.20) the right-hand side of (6.9) as second or fourth order polynomials of $a_{k,l}^i$. This, together with (6.12), (8.13), (8.18), as well as $\lambda_1 + \cdots + \lambda_5 = \frac{1}{2}$, leads to the required 45 equations in exactly the same manner as in (5.26). Again, we are not able to solve these equations explicitly, so we will solve them numerically.

Acknowledgment. The authors would like to thank the two anonymous referees for their constructive comments which have helped to improve the paper greatly.

REFERENCES

[1] A. Alfonsi and A. Kebaier, *Approximation of Stochastic Volterra Equations with Kernels of Completely Monotone Type*, preprint, arXiv:2102.13505, 2021.

[2] F. Baudoin and L. Coutin, *Operators associated with a stochastic differential equation driven by fractional Brownian motions*, Stochastic Process. Appl., 117 (2007), pp. 550–574.

[3] C. Bayer and S. Breneis, *Markovian approximations of stochastic Volterra equations with the fractional kernel*, Quant. Finance, 23 (2023), pp. 53–70.

[4] C. Bayer, M. Fukasawa, and S. Nakahara, *On the weak convergence rate in the discretization of rough volatility models*, SIAM J. Financial Math., 13 (2022), pp. SC66–SC73, https://doi.org/10.1137/22M1482871.

[5] C. Bayer, E. J. Hall, and R. Tempone, *Weak error rates for option pricing under the rough Bergomi model*, Int. J. Theor. Appl. Finance, 25 (2022), 2250029.

[6] J. Beran, *Statistics for Long-Memory Processes*, Chapman & Hall, London, 1994.

[7] T. Blu and M. Unser, *Self-similarity: Part* II*: Optimal estimation of fractal processes*, IEEE Trans. Signal Process., 55 (2007), pp. 1364–1378.

[8] O. Bonesini, A. Jacquier, and A. Pannier, *Rough Volatility, Path-Dependent PDEs and Weak Rates of Convergence*, preprint, arXiv:2304.03042, 2023.

[9] I. Chevyrev and H. Oberhauser, *Signature moments to characterize laws of stochastic processes*, J. Mach. Learn. Res., 23 (2022), 176.

[10] F. Comte and E. Renault, *Long memory in continuous-time stochastic volatility models*, Math. Finance, 8 (1998), pp. 291–323.

[11] D. Crisan and K. Manolarakis, *Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing*, SIAM J. Financial Math., 3 (2012), pp. 534–571, https://doi.org/10.1137/090765766.

[12] D. Crisan and K. Manolarakis, *Second order discretization of backward SDEs and simulation with the cubature method*, Ann. Appl. Probab., 24 (2014), pp. 652–678.

[13] D. Crisan and E. McMurray, *Cubature on Wiener space for McKean-Vlasov SDEs with smooth scalar interaction*, Ann. Appl. Probab., 29 (2019), pp. 130–177.

[14] P. C. de Raynal and C. G. Trillos, *A cubature based algorithm to solve decoupled Mckean-Vlasov forward-backward stochastic differential equations*, Stochastic Process. Appl., 125 (2015), pp. 2206–2255.

[15] M. El Omari, *Mixtures of higher-order fractional Brownian motions*, Comm. Statist. Theory Methods, 52 (2023), pp. 4200–4215.

[16] D. Filipović, M. Larsson, and S. Pulido, *Markov cubature rules for polynomial processes*, Stochastic Process. Appl., 130 (2020), pp. 1947–1971.

[17] J. Foster, T. Lyons, and H. Oberhauser, *An optimal polynomial approximation of Brownian motion*, SIAM J. Numer. Anal., 58 (2020), pp. 1393–1421, https://doi.org/10.1137/19M1261912.

[18] J. Foster, G. dos Reis, and C. Strange, *High Order Splitting Methods for SDEs Satisfying a Commutativity Condition*, preprint, https://arxiv.org/abs/2210.17543v1, 2022.

[19] J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna, *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*, Cambridge University Press, Cambridge, UK, 2011.
[20] P. K. Friz, W. Salkeld, and T. Wagenhofer, *Weak Error Estimates for Rough Volatility Models*, preprint, https://arxiv.org/abs/2212.01591, 2022.
[21] M. Fukasawa and T. Ugai, *Limit Distributions for the Discretization Error of Stochastic Volterra Equations*, preprint, arXiv:2112.06471, 2021.
[22] P. Gassiat, *Weak error rates of numerical schemes for rough volatility*, SIAM J. Financial Math., 14 (2023), pp. 475–496, https://doi.org/10.1137/22M1485760.
[23] J. Gatheral, T. Jaisson, and M. Rosenbaum, *Volatility is rough*, Quant. Finance, 18 (2018), pp. 933–949.
[24] A. Gulisashvili, F. Viens, and X. Zhang, *Extreme-strike asymptotics for general Gaussian stochastic volatility models*, Ann. Finance, 15 (2019), pp. 59–101.
[25] A. Gupta, P. Singh, and M. Karlekar, *A novel signal modeling approach for classification of seizure and seizure-free EEG signals*, IEEE Trans. Neural Syst. Rehabil. Eng., 26 (2018), pp. 925–935.
[26] L. G. Gyurkó and T. J. Lyons, *Efficient and practical implementations of cubature on Wiener space*, in Stochastic Analysis 2010, Springer, Heidelberg, 2011, pp. 73–111.
[27] F. A. Harang and S. Tindel, *Volterra equations driven by rough signals*, Stochastic Process. Appl., 142 (2021), pp. 34–78.
[28] F. A. Harang, S. Tindel, and X. Wang, *Volterra equations driven by rough signals 2: Higher order expansions*, Stoch. Dyn., 23 (2023), 2350002, https://doi.org/10.1142/S0219493723500028.
[29] P. Harms, *Strong convergence rates for Markovian representations of fractional processes*, Discrete Contin. Dyn. Syst. Ser. B, 26 (2021), pp. 5567–5579.
[30] S. Hayakawa, H. Oberhauser, and T. Lyons, *Hypercontractivity meets random convex hulls: Analysis of randomized multivariate cubatures*, Proc. A, 479 (2023), 20220725.
[31] A. J. Jacquier and M. Oumgari, *Deep PPDEs for rough local stochastic volatility*, SSRN 3400035, 2019.
[32] P. Kidger, P. Bonnier, I. P. Arribas, C. Salvi, and T. Lyons, *Deep signature transforms*, in Advances in Neural Information Processing Systems, Vol. 19, 2019, pp. 3105–3115.
[33] F. J. Király and H. Oberhauser, *Kernels for sequentially ordered data*, J. Mach. Learn. Res., 20 (2019), 31.
[34] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Appl. Math. (N. Y.) 23, Springer-Verlag, Berlin, 1992.
[35] M. Li, C. Huang, and Y. Hu, *Numerical methods for stochastic Volterra integral equations with weakly singular kernels*, IMA J. Numer. Anal., 42 (2022), pp. 2656–2683.
[36] S. Liao, T. Lyons, W. Yang, and H. Ni, *Learning Stochastic Differential Equations Using RNN with Log Signature Features*, preprint, arXiv:1908.08286, 2019.
[37] C. Litterer and T. Lyons, *Cubature on Wiener space continued*, in Stochastic Processes and Applications to Mathematical Finance, World Scientific, Hackensack, NJ, 2007, pp. 197–217.
[38] C. Litterer and T. Lyons, *High order recombination and an application to cubature on Wiener space*, Ann. Appl. Probab., 22 (2012), pp. 1301–1327.
[39] T. Loussot, R. Harba, G. Jacquet, C. L. Benhamou, E. Lespesailles, and A. Julien, *An oriented fractal analysis for the characterization of texture: Application to bone radiographs*, EUSIPCO Signal Process., 1 (1996), pp. 371–374.
[40] T. Lyons and N. Victoir, *Cubature on Wiener space*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 460 (2004), pp. 169–198.
[41] J. Ma, W. Yang, and Z. Cui, *Semimartingale and Continuous-Time Markov Chain Approximation for Rough Stochastic Local Volatility Models*, preprint, arXiv:2110.08320, 2021.
[42] S. Ninomiya and Y. Shinozaki, *On implementation of high-order recombination and its application to weak approximations of stochastic differential equations*, in Proceedings of the NFA 29th Annual Conference, 2021.
[43] S. Ninomiya and N. Victoir, *Weak approximation of stochastic differential equations and application to derivative pricing*, Appl. Math. Finance, 15 (2008), pp. 107–121.
[44] D. Nualart, *Malliavin Calculus and Related Topics*, 2nd ed., Springer, Berlin, 2006.

[45] D. Nualart and B. Saikia, *Error distribution of the Euler approximation scheme for stochastic Volterra equations*, J. Theoret. Probab., 36 (2023), pp. 1829–1876.
[46] R. Passeggeri, *On the signature and cubature of the fractional Brownian motion for* $H>1/2$, Stochastic Process. Appl., 130 (2020), pp. 1226–1257.
[47] E. Perrin, R. Harba, C. Berzin-Joseph, I. Iribarren, and A. Bonami, *nth-order fractional Brownian motion and fractional Gaussian noises*, IEEE Trans. Signal Process., 49 (2001), pp. 1049–1059.
[48] A. Richard, X. Tan, and F. Yang, *Discrete-time simulation of stochastic Volterra equations*, Stochastic Process. Appl., 141 (2021), pp. 109–138.
[49] A. Richard, X. Tan, and F. Yang, *On the discrete-time simulation of the rough Heston model*, SIAM J. Financial Math., 14 (2023), pp. 223–249, https://doi.org/10.1137/21M1443807.
[50] J. Ruan and J. Zhang, *Numerical Methods for High Dimensional Path-Dependent PDEs Driven by Stochastic Volterra Integral Equations*, manuscript, 2022.
[51] F. Viens and J. Zhang, *A martingale approach for fractional Brownian motions and related path dependent PDEs*, Ann. Appl. Probab., 29 (2019), pp. 3489–3540.
[52] T. Wang, *Linear quadratic control problem of stochastic Volterra integral equations*, ESAIM Control Optim. Calc. Var., 24 (2018), pp. 1849–1879.
[53] C. Wen and T. Zhang, *Improved rectangular method on stochastic Volterra equations*, J. Comput. Appl. Math., 235 (2011), pp. 2492–2501.
[54] J. Zhang, *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory,* Probab. Theory Stoch. Model. 86, Springer, New York, 2017.
[55] X. Zhang, *Euler schemes and large deviations for stochastic Volterra equations with singular kernels*, J. Differential Equations, 244 (2008), pp. 2226–2250.