FORESTS, CUMULANTS, MARTINGALES

BY PETER K. FRIZ[1,a], JIM GATHERAL[2,b] AND RADOŠ RADOIČIĆ[2,c]

[1]*Technische Universität Berlin and WIAS Berlin,* [a]*friz@math.tu-berlin.de*

[2]*Baruch College, CUNY,* [b]*jim.gatheral@baruch.cuny.edu,* [c]*rados.radoicic@baruch.cuny.edu*

This work is concerned with forest and cumulant type expansions of general random variables on a filtered probability space. We establish a "broken exponential martingale" expansion that generalizes and unifies the exponentiation result of Alòs, Gatheral, and Radoičić (SSRN'17; *Quant. Finance* **20** (2020) 13–27) and the cumulant recursion formula of Lacoin, Rhodes, and Vargas (arXiv; (2019)). Specifically, we exhibit the two previous results as lower dimensional projections of the same generalized forest expansion, subsequently related by forest reordering. Our approach also leads to sharp integrability conditions for validity of the cumulant formula, as required by many of our examples, including iterated stochastic integrals, Lévy area, Bessel processes, KPZ with smooth noise, Wiener–Itô chaos, and "rough" stochastic (forward) variance models.

1. Introduction.

1.1. *Statements of main results.* Consider a filtered probability space $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{0 \le t \le T}; \mathbb{P})$, on which all martingales admit a continuous version. (Itô's representation theorem, e.g., [26], Chapter V.3., states that this holds true for Brownian filtrations which covers all situations we have in mind.) Throughout, $T \in (0, \infty]$ should be thought of as a fixed parameter.

Let A_T be $\mathcal{F}_T$-measurable. Define, assuming sufficient integrability,[1]

$$X_t := \log \mathbb{E}_t e^{A_T}, \qquad Y_t := \mathbb{E}_t A_T.$$

By construction, X, Y have equal terminal value $X_T = Y_T = A_T$, and e^X, Y are martingales. Motivated by financial applications, in [2] an $\mathbb{F}$ *(forest) expansion* was given of the form[2]

$$\mathbb{E}_t e^{zA_T} = \mathbb{E}_t e^{zX_T} = e^{zX_t + \frac{1}{2}z(z-1)\mathbb{E}_t \langle X \rangle_{t,T} + \cdots}$$

$$\equiv \exp\left(zX_t + \frac{1}{2}z(z-1)(X \diamond X)_t(T) + \sum_{k \ge 2} \mathbb{F}_t^k(T; z)\right)$$

with quadratic recursion for the $\mathbb{F}$'s, homogenous in X but not in z, representable as forests. But A_T is also the terminal value of the martingale Y so that

$$\mathbb{E}_t e^{zA_T} = \mathbb{E}_t e^{zY_T} = e^{zY_t + \frac{1}{2}z^2\mathbb{E}_t Y_{t,T}^2 + \frac{1}{3!}z^3\mathbb{E}_t Y_{t,T}^3 + \frac{1}{4!}z^4(\mathbb{E}_t Y_{t,T}^4 - 3(\mathbb{E}_t Y_{t,T}^2)^2) + \cdots}$$

$$\equiv \exp\left(zY_t + \sum_{n \ge 2} \mathbb{K}_t^n(T; z)\right),$$

[1]Notation $\mathbb{E}_t \equiv \mathbb{E}(\cdot | \mathcal{F}_t)$.

[2]Corollary 3.1 in [2] is an expansion of the characteristic function, with $z = i\xi$.

the (time-t) conditional $\mathbb{K}$ *(cumulant, German: Kumulanten) expansion* of $A_T = Y_T$. A somewhat similar quadratic recursion for the $\mathbb{K}$'s, homogenous in z (equivalently: Y) was later obtained in [25], stated in the unconditional case, and motivated by applications in QFT, and independently in a first version of this paper, when revisiting the convergence properties of the $\mathbb{F}$-series. (Initially, the present authors were unaware of [25], whereas the authors of [25] were unaware of [2].[3]) We note that the $\mathbb{F}$-expansion was left as formal expansion in [2], whereas validity of the $\mathbb{K}$ recursion of [25] was only shown under a stringent integrability condition which rules out virtually all examples discussed later on. The main theorem of this paper is a $\mathbb{G}$ *(generalized forest) expansion*, which contains both $\mathbb{F}$- and $\mathbb{K}$-expansion as special cases, together with optimal integrability conditions for convergence. Our arguments are also well adapted to further localization, as seen in points (i) in Theorems 1.1 and 1.2 below.

DEFINITION 1.1. Given two continuous semimartingales A, B with integrable covariation process $\langle A, B \rangle$, the diamond product[4] of A and B is another continuous semimartingale given by, writing $\langle A, B \rangle_{t,T}$ for the difference of $\langle A, B \rangle_T$ and $\langle A, B \rangle_t$,

$$(A \diamond B)_t(T) := \mathbb{E}_t \big[\langle A, B \rangle_{t,T} \big] = \mathbb{E}_t \big[\langle A, B \rangle_T \big] - \langle A, B \rangle_t.$$

Here and below we say that A_T has exponential moments, if $\mathbb{E}e^{xA_T} < \infty$ for x in some neighbourhood of $0 \in \mathbb{R}$. This of course implies that A_T has moments of all orders: $A_T \in L^N$, for any $N \in \mathbb{N}$.

THEOREM 1.1 ($\mathbb{G}$ expansion, a.k.a. broken exponential martingale). *Let A_T be a real-valued, integrable and $\mathcal{F}_T$-measurable random variable and $Y_t = \mathbb{E}_t A_T, 0 \le t \le T$, the associated (continuous) martingale.*

(i) *Assume A_T has moments of all orders. Let $z_1, z_2 \in \mathbb{C}$ with $\mathrm{Re}(z_1) = 0, \mathrm{Re}(z_2) \le 0$. Then the following asymptotic expansion for the time-t conditional joint characteristic function / Laplace transform of $(Y_T, \langle Y \rangle_T)$:*

$$(1.1) \qquad \log \mathbb{E}_t \big[e^{z_1 Y_T + z_2 \langle Y \rangle_T} \big] \sim z_1 Y_t + z_2 \langle Y \rangle_t + \sum_{k \ge 2} \mathbb{G}_t^k(T; z) \quad as\ z \to 0,$$

$$\mathbb{G}_t^2(T, z) = \left(\frac{1}{2} z_1^2 + z_2 \right) (Y \diamond Y)_t(T) \quad and$$

$$(1.2)$$

$$\forall k > 2 : \mathbb{G}^k = \frac{1}{2} \sum_{j=2}^{k-2} \mathbb{G}^{k-j} \diamond \mathbb{G}^j + \big(z_1 Y \diamond \mathbb{G}^{k-1} \big).$$

(ii) *Assume $\langle Y \rangle_T$ (or Y_T in case $z_2 = 0$) has exponential moments, then (1.1) can be strengthened to equality, with a.s. absolutely convergent sum $\Lambda := \Lambda_t^T := \Lambda_t^T(z) := \sum_{k \ge 2} \mathbb{G}_t^k(T; z)$ on the right-hand side, for $z \in \mathbb{C}^2$ with $|z| < \rho_t(\omega)$, a.s. strictly positive.*

(iii) *For the multivariate case, with $Y_t^i = \mathbb{E}_\bullet Y_T^i, i = 1, \dots, d$, it suffices to replace*

$$z_1 Y_T \rightsquigarrow z_{1;i} Y_T^i, \qquad z_2 \langle Y \rangle_T \rightsquigarrow z_{2;j,k} \langle Y^j, Y^k \rangle_T,$$

$$\mathbb{G}^2 \rightsquigarrow \left(\frac{1}{2} z_{1;i} z_{1;j} + z_{2;i,j} \right) (Y^i \diamond Y^j)_t(T),$$

with summation over repeated indices.

---

[3]Preprint of [2] posted on SSRN in 2017. We share with the authors of [25] our surprise that such recursions had not been discovered decades earlier.

[4]Warning: Our diamond product is (very) different from the Wick product, for example, Chapter III of [22].

With $z = (z_1, -z_1/2)$, the $\mathbb{G}$-recursion becomes precisely the $\mathbb{F}$-recursion, equation (3.1) in [2], whereas the case $z = (z_1, 0)$ yields the $\mathbb{K}$ (cumulant) recursion, equations (3.4), (3.9) in [25] as seen in (1.3) below. We should note that the change-of-measure based derivation of Lacoin et al. was given under stringent integrability assumption (a L^∞ bound on $\langle Y \rangle_T$, hence Gaussian concentration of $Y_T = A_T$), though the authors (correctly) suspect validity of the cumulant recursion in greater generality. Below we achieve this under optimal conditions, namely finite exponential moments which is required for the cumulant generating function to exist. We will also show that the cumulant recursion is valid, as a *finite recursion*, under a matching *finite moment* assumption. (Existence of the first N moments is equivalent to existence of the first N cumulants.) In part (iii), we give the multivariate formulation.

THEOREM 1.2.

(i) *Let A_T be $\mathcal{F}_T$-measurable with $N \in \mathbb{N}$ finite moments. Then the recursion*

$$(1.3) \quad \mathbb{K}_t^1(T) := \mathbb{E}_t[A_T] \quad and \quad \forall n > 0: \quad \mathbb{K}_t^{n+1}(T) = \frac{1}{2} \sum_{k=1}^n (\mathbb{K}^k \diamond \mathbb{K}^{n+1-k})_t(T),$$

with $t \in [0, T]$ is well defined up to $\mathbb{K}^N$ and, for $z \in i\mathbb{R}$,

$$\log \mathbb{E}_t[e^{zA_T}] = \sum_{n=1}^N z^n \mathbb{K}_t^n(T) + o(|z|^N) \quad as \ z \to 0,$$

which identifies $n! \times \mathbb{K}_t^n(T)$ as the (time t-conditional) nth cumulant of A_T. If A_T has moments of all orders, we have the asymptotic expansion

$$(1.4) \qquad \log \mathbb{E}_t[e^{zA_T}] \sim \sum_{n=1}^\infty z^n \mathbb{K}_t^n(T) \quad as \ z \to 0.$$

(ii) *If A_T has exponential moments, so that its (time t-conditional) mgf $\mathbb{E}_t[e^{xA_T}]$ is a.s. finite for $x \in \mathbb{R}$ in some neighbourhood of zero, then there exist a maximal convergence radius $\rho = \rho_t(\omega) \in (0, \infty]$ a.s. such that for all $z \in \mathbb{C}$ with $|z| < \rho$,*

$$(1.5) \qquad \log \mathbb{E}_t[e^{zA_T}] = \sum_{n=1}^\infty z^n \mathbb{K}_t^n.$$

(iii) *In the multivariate case, with $\mathbb{K}^{(1);i} = \mathbb{E}_\bullet A_T^i, i = 1, \ldots, d$ replace $\mathbb{K}^n \rightsquigarrow \mathbb{K}^{(n)}$, n-tensors (over $\mathbb{R}^d$), with tensorial interpretation of the diamond product in (1.3), and substitute in the expansion*

$$z^n \mathbb{K}_t^n \rightsquigarrow \langle z^{\otimes n}, \mathbb{K}^{(n)} \rangle = z_{i_1} \cdots z_{i_n} \mathbb{K}^{(n);i_1,\ldots,i_n}.$$

The multivariate (t-conditional) cumulants of A_T are then precisely given by $n! \mathbb{K}_t^{(n);i_1,\ldots,i_n}$.

1.2. *Tree, forests and reordering.* Both quadratic recursions (1.2) and (1.3) lead to binary trees, with the (commutative, nonassociated) diamond product represented by *root joining*,

$$\tau_1 \diamond \tau_2 = \tau_2 \diamond \tau_1 = \quad \overset{\displaystyle \overset{\tau_1 \ \ \tau_2}{\diagdown \diagup}}{\bullet} \quad ,$$

where we agree to set $Y \equiv \circ$, interpreted as single leaf. In the 2-variate case, we can write $(Y^1, Y^2) = (\circ, \circledast)$, for the d-variate case use labels or colors. For better readability, write

$$(z_1, z_2) \rightsquigarrow (a, b), \qquad z_1 Y_T + z_2 \langle Y \rangle_T \rightsquigarrow a Y_T + b \langle Y \rangle_T,$$

in (1.2). The tree formalism is extremely convenient when it comes to doing explicit computations and also explains an interesting relation between the $\mathbb{G}$-recursion and the $\mathbb{K}$-recursion. Specifically, we see that the $\mathbb{G}$-recursion is equivalent to the 2-variate $\mathbb{K}$-recursion applied to $A_T = (Y_T, \langle Y \rangle_T)$ *after forest reordering*. This procedure has moreover the important effect of resolving infinite cancellations present in the 2-variate $\mathbb{K}$-expansion, as may be seen by applying it to the exponential martingale case ($b = -a^2/2$). The first few $\mathbb{G}$-forests are then spelled out as follows:

$$\mathbb{G}^2 = \left(\frac{1}{2}a^2 + b\right) \, \text{\small⌣},$$

$$\mathbb{G}^3 = a\left(\frac{1}{2}a^2 + b\right) \, \text{\small⌣},$$

(1.6)

$$\mathbb{G}^4 = \frac{1}{2}\left(\frac{1}{2}a^2 + b\right)^2 \, \text{\small⌣} + a^2\left(\frac{1}{2}a^2 + b\right) \, \text{\small⌣},$$

$$\mathbb{G}^5 = a\left(\frac{1}{2}a^2 + b\right)^2 \, \text{\small⌣} + \frac{1}{2}a\left(\frac{1}{2}a^2 + b\right)^2 \, \text{\small⌣} + a^3\left(\frac{1}{2}a^2 + b\right) \, \text{\small⌣}.$$

Note that these $\mathbb{G}$-forests consist of trees which are homogenous in the number of leaves ($\leftrightarrow Y$) but not in a, b (unless powers of b are counted twice). Upon setting $b = 0$ we get $\mathbb{G}^n(T; (a, 0)) = \mathbb{K}^n(T, a) = a^n \mathbb{K}^n(T)$ and get the first few $\mathbb{K}$-forests: $\mathbb{K}^1 = Y \equiv \circ$ and

$$\mathbb{K}^2 = \frac{1}{2}Y \diamond Y \equiv \frac{1}{2} \, \text{\small⌣},$$

(1.7)

$$\mathbb{K}^3 = \frac{1}{2}(Y \diamond Y) \diamond Y \equiv \frac{1}{2} \, \text{\small⌣},$$

$$\mathbb{K}^4 = \frac{1}{8}(Y \diamond Y)^{\diamond 2} + \frac{1}{2}((Y \diamond Y) \diamond Y) \diamond Y \equiv \frac{1}{8} \, \text{\small⌣} + \frac{1}{2} \, \text{\small⌣},$$

(1.8)

$$\mathbb{K}^5 = \cdots = \frac{1}{4} \, \text{\small⌣} + \frac{1}{8} \, \text{\small⌣} + \frac{1}{2} \, \text{\small⌣}.$$

REMARK 1.1. The number of different tree shapes seen in both $\mathbb{G}^n$, $\mathbb{K}^n$ above is precisely the number of interpretations of an $(n-1)$-fold (commutative but not associative) diamond product, that is, the number of ways to insert parentheses. Starting from the empty tree, the resulting integers $\{0, 1, 1, 1, 2, 3, 6, \ldots\}$ are known as Wedderburn–Etherington numbers (OEIS A001190). The pre-factors in $2^n \mathbb{K}^{n+1}$ further display the symmetry factors, for example, $2^4 \mathbb{K}^5 = 4 \, \text{\small⌣} + 2 \, \text{\small⌣} + 8 \, \text{\small⌣}$ with $4 + 2 + 8 = 14 = C_4$ where we recall that the nth Catalan number (A000108), standard example in analytic combinatorics, gives the number of binary trees with $n + 1$ leaves. We note the combinatorial consistency of (1.3) with Segner's recursion $C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$ if rewritten for $2^n \mathbb{K}^{n+1}$.

The 2-variate $\mathbb{F}$-forests can be represented by all possible consistent ways of marking leaves with $\times$. This leads to, for example, 2×2-matrix valued $\mathbb{K}^{(2)}$ and $(\mathbb{R}^2)^{\otimes 4} \cong \mathbb{R}^{2^4}$ valued $\mathbb{K}^{(4)}$, with representative trees of the form

$$\text{\small⌣}_{\otimes} \qquad \text{\small⌣}_{\otimes}.$$

Tensor-contracting these 2-variate $\mathbb{K}$-trees with powers of a (number of ∘ leaves) and b (number of ⊛ leaves), and a ⊛ $\rightsquigarrow$ ⧂ substitution,[5]

$$\mathbb{K}^1 = a\circ + b\,\text{⧂},$$

$$\mathbb{K}^2 = \frac{1}{2}(a\circ + b\,\text{⧂})^{\diamond 2} = \frac{1}{2}a^2\,\text{⧂} + ab\,\text{⧂} + \frac{1}{2}b^2\,\text{⧂},$$

$$\mathbb{K}^3 = \frac{1}{2}a^3\,\text{⧂} + \frac{1}{2}a^2 b\,\text{⧂} + a^2 b\,\text{⧂} + ab^2\,\text{⧂} + \frac{1}{2}ab^2\,\text{⧂} + \cdots,$$

(1.9)
$$\mathbb{K}^4 = \frac{1}{2}a^4\,\text{⧂} + \frac{1}{2^3}a^4\,\text{⧂} + \frac{1}{2}a^3 b\,\text{⧂}$$

$$+ \frac{1}{2}a^3 b\,\text{⧂} + a^3 b\,\text{⧂} + \frac{1}{2}a^3 b\,\text{⧂} + \cdots,$$

$$\mathbb{K}^5 = \frac{1}{2}a^5\,\text{⧂} + \frac{1}{2^3}a^5\,\text{⧂} + \frac{1}{2^2}a^5\,\text{⧂} + \cdots.$$

COROLLARY 1.1 (Forest reordering). *The $\mathbb{G}$-expansion* (1.6) *is a reordering of the 2-variate $\mathbb{K}$-expansion applied to* $A_T = (Y_T, \langle Y\rangle_T)$, *as displayed in* (1.9), *based on the number of leaves.*

PROOF. From Theorem 1.1, we know that $\log \mathbb{E}_t \exp(aY_T + b\langle Y\rangle_T)$ admits an (absolutely convergent) $\mathbb{G}$ expansion, with terms homogenous in Y ($\leftrightarrow$ number of leaves), and similarly a 2-variate $\mathbb{K}$-expansion with terms homogenous in a, b. The statement follows. □

REMARK 1.2 (Generalized forests vs. cumulants). The $\mathbb{G}$ and $\mathbb{K}$ expansions coincide in absence of the term $b\langle Y\rangle_T$: just set $b = 0$ in (1.6), (1.9). In general, $b \neq 0$ matters in application (e.g., Section 4.7), the expansions are then different and the question arises about their qualitative difference. The $\mathbb{G}$ expansion has the advantage of preserving some structural properties, lost in the 2-variate $\mathbb{K}$-expansion. To wit, consider the exponential martingale case in which case

$$b = -a^2/2 \quad \Longrightarrow \quad \mathbb{E}_t \exp(aY_T + b\langle Y\rangle_T) = \exp(aY_t + b\langle Y\rangle_t),$$

so that both $\mathbb{G}$-sum, that is, $\sum_{k\geq 2}\mathbb{G}_t^k$, and $\mathbb{K}$-sum must vanish. However, while the zero $\mathbb{G}$-sum only consists of zero summands $\mathbb{G}^k$, as seen in (1.6) upon taking $b = -a^2/2$, this is not so for the $\mathbb{K}$-sum, and only an infinite cascade of cancelations among the $\mathbb{K}^k$'s causes the sum to vanish.

REMARK 1.3. Corollary 1.1 suggests an alternative proof of the $\mathbb{G}$-expansion from the $\mathbb{K}$-expansion, based on the combinatorics of forest reordering.

2. Proofs. Recall (e.g., [26], Chapter IV.4.) that a (continuous) semimartingale $A = M^A + V^A$, defined on $[0, T]$, is said to be in class $\mathcal{H}_T^p$ if both square-root function and total variation, on $[0, T]$, have finite pth moments, $p > 0$, so that

$$\|A\|_{\mathcal{H}_T^p} := \left\| \sqrt{\langle M^A\rangle_T} + \int_0^T |dV^A| \right\|_{L^p(\mathbb{P})} < \infty.$$

---

[5]Strictly speaking, the 2-variate case $A_T = (Y_T, \langle Y\rangle_T)$ gives $\mathbb{E}_t A_T = (Y_t, \mathbb{E}_t\langle Y\rangle_T)$ with $\mathbb{E}_t\langle Y\rangle_T = (Y \diamond Y)_t(T) + \langle Y\rangle_t = \text{⧂} + BV$, where the bounded variation (BV) in t component $\langle Y\rangle_t$ shows up as a multiple of z_2 in (1.1), but does not contribute to subsequent diamond products.

Recall the Burkholder–Davis–Gundy (BDG) estimates and Doob's maximal inequality. For a (continuous, local) martingale M, with $M_T^* = \sup_{t \in [0,T]} M_t$, one has

$$(2.1) \qquad \|M\|_{\mathcal{H}_T^p} = \left\|\sqrt{\langle M \rangle_T}\right\|_{L^p(\mathbb{P})} \asymp \|M_T^*\|_{L^p(\mathbb{P})} \asymp \|M_T\|_{L^p(\mathbb{P})}$$

with two-sides estimates valid for $p > 0$ (BDG) and $p > 1$ (Doob), respectively.

PROPOSITION 2.1. *Consider continuous semimartingales A, B with continuous martingale parts $M^A \in \mathcal{H}_T^a$, $M^B \in \mathcal{H}_T^b$ for $a, b \geq 1$. Assume $c = 1/(a^{-1} + b^{-1}) \geq 1$. Then*

$$C_t := (A \diamond B)_t(T)$$

defines a semimartingale C with uniformly integrable martingale part M^C and integrable total variation of V^C on $[0, T]$. If $c > 1$, then $C = (A \diamond B)(T) \in \mathcal{H}_T^c$ and

$$(2.2) \qquad \|C\|_{\mathcal{H}_T^c} \lesssim \|A\|_{\mathcal{H}_T^a} \|B\|_{\mathcal{H}_T^b}$$

with a multiplicative constant that depends on a, b. For $c = 1$ we still have the estimate

$$(2.3) \qquad \sup_{t \in [0,T]} \|M_t^C\|_{L^1} + \left\|\int_0^T |dV^C|\right\|_{L^1} \leq \|A\|_{\mathcal{H}_T^a} \|B\|_{\mathcal{H}_T^b}.$$

REMARK 2.1. (i) Estimate (2.2) implies that the diamond product is a continuous map

$$\diamond : \mathcal{H}_T^a \times \mathcal{H}_T^b \to \mathcal{H}_T^c,$$

provided $c = 1/(a^{-1} + b^{-1}) > 1$. When $c = 1$ we still have continuity, but now in the weaker sense of (2.3).

(ii) It further follows that the diamond product is a bona fide (commutative, nonassociative) product on the linear space of semimartingales

$$\mathcal{H}_T^{\infty-} := \bigcap_{p < \infty} \mathcal{H}_T^p.$$

Similar to $L^{\infty-} = \bigcap_{p < \infty} L^p$ over $(\Omega, \mathcal{F}, \mathbb{P})$, the family of $\mathcal{H}_T^p$-norms give rise to a Fréchet space, which then becomes a topological algebra under the diamond product.[6]

PROOF OF PROPOSITION 2.1. First verify integrability of the quadratic covariation process $(\langle A, B \rangle_t : 0 \leq t \leq T)$: Using $\langle M^A \rangle_T^{1/2} \in L^a(\mathbb{P})$, $\langle M^B \rangle_T^{1/2} \in L^b(\mathbb{P})$, Cauchy–Schwarz (in the form of, for example, [26], Chapter IV.1), and Hölder's inquality, we see that, whenever $0 \leq t \leq T$,

$$\langle A, B \rangle_t = \langle M^A, M^B \rangle_t \leq \sqrt{\langle M^A \rangle_t}\sqrt{\langle M^B \rangle_t} \leq \sqrt{\langle M^A \rangle_T}\sqrt{\langle M^B \rangle_T} \in L^c(\mathbb{P}).$$

Since $c \geq 1$ the desired integrability follows. Hence, $C = (A \diamond B)_\bullet(T)$ is well defined. With $\langle A, B \rangle_T \in L^1(\mathbb{P})$, it is furthermore clear that $M_t^C := \mathbb{E}_t \langle A, B \rangle_T$ defines a uniformly integrable martingale and

$$\sup_{t \in [0,T]} \|M_t^C\|_{L^1} \leq \|M^A\|_{\mathcal{H}_T^a} \|M^B\|_{\mathcal{H}_T^b}.$$

Concerning the bounded variation part, $V_t^C := -\langle A, B \rangle_t \in L^c(\mathbb{P})$ not only for fixed $t \in [0, T]$, but also in total variation sense

$$\int_0^T |dV^C| \leq \langle A \rangle_T^{1/2} \langle B \rangle_T^{1/2} \in L^c(\mathbb{P}),$$

---

[6]See, for example, [30] for the analogous statement for $L^{\infty-}$.

as follows, for example, from [26], Chapter IV.4, Proposition 1.1.5, applied with $H \equiv K \equiv 1$. When $c > 1$, we have $\sup_{0 \leq t \leq T} |M_t^C| \in L^c(\mathbb{P})$ by Doob's maximal inequality. Thanks to the Burkholder–Davis–Gundy (BDG) inequalities, this is equivalent to $\langle M^C \rangle_T^{1/2} \in L^c(\mathbb{P})$ which concludes the proof that $C \in \mathcal{H}_T^c$. Finally, appealing to Doob and BDG in their quantitative form leads to the stated estimate. $\square$

2.1. *Proof of Theorem* 1.1. Recall that we work on a fixed time horizon, with $t \in [0, T]$ throughout. We may observe right away that under the assumption that $A_T = Y_T$ has moments of all orders, we have $Y \in \mathcal{H}_T^{\infty-}$ so that, thanks to Proposition 2.1, the recursion (1.2) is well defined. (See also Lemma 2.1 below for a more precise statement, as needed in Theorem 1.2.)

We begin by proving the second part of the theorem.

2.1.1. *Part* (ii). By assumption, the mgf of $\langle Y \rangle_T$ is finite in some neighbourhood of zero.[7] The same holds true for any linear combination $z_1 Y_T + z_2 \langle Y \rangle_T$, $z = (z_1, z_2) \in \mathbb{R}^2$. To see this, it suffices to treat Y_T (thanks to Cauchy–Schwarz). The argument goes along the proof of Novikov's criterion. (Indeed, by scaling Y if necessary and a deterministic time-change $T \rightsquigarrow \infty$, we may assume $\mathbb{E}[e^{\langle Y \rangle_\infty/2}] < \infty$. Then copy the estimates of [26], Chapter VIII.1, Proposition 1.15, based on the a priori fact that $\mathbb{E}[\mathcal{E}(Y)_\infty] \leq 1$.) Obviously, we can skip this argument in the case $z_2 = 0$, provided we assume directly that the mgf of Y_T is finite in some neighbourhood of zero.

Fix $z = (z_1, z_2) \in \mathbb{R}^2$ and define a family of continuous semimartingales given by

$$Z_t(\epsilon) := \epsilon z_1 Y_t + \epsilon^2 z_2 \langle Y \rangle_t.$$

Thanks to finite exponentials moments of $Y_T, \langle Y \rangle_T$, it is clear that, for every $p < \infty$,

$$\sup_{0 \leq \epsilon \leq 1} \| Z(\epsilon) \|_{\mathcal{H}_T^p} < \infty. \tag{2.4}$$

For (any) fixed $p < \infty$, we further see that for all ϵ small enough, the mgf of $\pm p Z(\epsilon)$ is finite, hence $e^{|Z_T(\epsilon)|} \in L^p(\mathbb{P})$. For such ϵ, keeping also T fixed,

$$M_t := M_t^T(\epsilon) := \mathbb{E}_t e^{Z_T(\epsilon)} = e^{Z_t(\epsilon) + \Lambda_t^T(\epsilon)} \tag{2.5}$$

defines a positive (and, by standing assumption: continuous) L^p-martingale, where $\Lambda_t^T(\epsilon)$ is defined through the last equality. This plainly gives a semimartingale $\Lambda_\bullet^T(\epsilon)$ on $[0, T]$ with zero terminal value, $\Lambda_T^T(\epsilon) = 0$. Itô's formula shows that the martingale M can be written as M_0 times the stochastic exponential $\mathcal{E}(L) \equiv \exp(L - \frac{1}{2}\langle L \rangle)$ of a local martingale L, known as *stochastic logarithm*, given by

$$L_t := \int_0^t M^{-1} dM = \log(M_0^{-1} M_t) + \frac{1}{2}\langle \log M \rangle_t = Z_t + \Lambda_t^T - (Z_0 + \Lambda_0^T) + \frac{1}{2}\langle Z + \Lambda^T \rangle_t,$$

where we write indifferently $L_t = L_t(\epsilon) = L_t^T(\epsilon)$, such as to highlight dependence on ϵ or ϵ, T whenever useful. We show that $L = L(\epsilon)$ is a genuine (even L^p-)martingale (any $p < \infty$, for all sufficiently small ϵ, depending on p.) To this end, fix $p < \infty$ and take ϵ small enough so that $\|M\|_{\mathcal{H}^{4p}} \vee \|N\|_{\mathcal{H}^{4p}} < \infty$ where $N_t := \mathbb{E}_t(1/M_T) = \mathbb{E}_t e^{-Z_T(\epsilon)}$ has similar integrability properties to M. From (conditional) Jensen, $1/M_t = 1/\mathbb{E}_t(M_T) \leq \mathbb{E}_t(1/M_T) = N_t$, hence

$$\left\| \sqrt{\int_0^T \frac{1}{M_s^2} d\langle M \rangle_s} \right\|_{L^p} \leq \left\| \int_0^T N_s^2 d\langle M \rangle_s \right\|_{L^p} \leq \|N_T^*\|_{L^{4p}}^2 \|\langle M \rangle_T\|_{L^{2p}}$$

$$\asymp \|N\|_{\mathcal{H}^{4p}}^2 \|M\|_{\mathcal{H}^{4p}}^2 < \infty,$$

---

[7]In general, one cannot replace $\langle Y \rangle_T$ by Y_T here. See Remark 2.3.

which precisely shows that $\|L\|_{\mathcal{H}_T^p} \equiv \|\sqrt{\langle L \rangle_T}\|_{L^p(\mathbb{P})} < \infty$, as was claimed. At this stage, we should point out that $Z + \Lambda^T = \log M$ (classical logarithm) is a semimartingale with decomposition $L - \frac{1}{2}\langle L \rangle$; it then immediately follows from $L \in \mathcal{H}_T^p$ that $Z + \Lambda^T \in \mathcal{H}^p$, any $p < \infty$, for all sufficiently small ϵ, depending on p. Thanks to (2.4) the same statement then holds for $\Lambda^T = \Lambda^T(\epsilon)$; that is,

$$(2.6) \qquad \forall p < \infty: \quad \limsup_{\epsilon \to 0} \|\Lambda^T(\epsilon)\|_{\mathcal{H}^p} < \infty.$$

(Proposition 2.1 now guarantees that all the diamond product in (2.7) below is well defined and take values in $\mathcal{H}^p$-semimartingale spaces, any $p < \infty$, provided ϵ is small enough.)

In what follows, write $Z_{t,T} = Z_T - Z_t$, $\Lambda_{t,T}^T = \Lambda_T^T - \Lambda_t^T = -\Lambda_t^T$, any $t \in [0, T]$. Using martingality of L, and hence $Z + \Lambda^T + \frac{1}{2}\langle Z + \Lambda^T \rangle$, we see that

$$(2.7) \qquad \Lambda_t^T = \mathbb{E}_t Z_{t,T} + \frac{1}{2}((Z + \Lambda^T) \diamond (Z + \Lambda^T))_t(T).$$

We further make explicit

$$(2.8) \qquad \mathbb{E}_t Z_{t,T} = \mathbb{E}_t[Z_{t,T}(\epsilon)] = \mathbb{E}_t[\epsilon z_1 Y_{t,T} + \epsilon^2 z_2 \langle Y \rangle_{t,T}] = \epsilon^2 z_2 (Y \diamond Y)_t(T).$$

REMARK 2.2. We view (2.7) as abstract (backward in t) functional equation (FRE). The quadratic right-hand side is reminiscent of Ricatti equations which describe the characteristic exponent of affine processes; see, for example, [11], Theorem 2.2.

The remainder of this proof is essentially a power series expansion of $\epsilon \mapsto \Lambda^T(\epsilon)$ at $\epsilon = 0$, which—carefully plugged into (2.7)—yields the stated recursions. As noted in the very beginning of this proof, the bivariate mgf of $(Y_T, \langle Y \rangle_T)$ is finite in some real neighbourhood of $(0, 0) \in \mathbb{R}^2$. As is well known, mgf's are analytic in the transform variables near the origin, with identical proof in the (time-t) conditional case.[8] Using analyticity of $\log(x)$ near $x = 1$ it is then clear that

$$\varepsilon \mapsto \log \mathbb{E}_t e^{\varepsilon z_1 Y_T + \varepsilon^2 z_2 \langle Y \rangle_T} = Z_t(\epsilon) + \Lambda_t^T(\epsilon)$$

is (real-)analytic at $\epsilon = 0$, with a.s. strictly positive, $\mathcal{F}_t$-measurable radius of convergence. Upon substraction of $Z_t(\epsilon) = \epsilon z_1 Y_t + \epsilon^2 z_2 \langle Y \rangle_t$ the same is true for

$$(2.9) \qquad \Lambda_t^T = \Lambda_t^T(\varepsilon) = \sum_{m \geq 2} \epsilon^m g_t^{T;m}$$

with $\mathcal{F}_t$-measurable coefficients $g_t^{T;m}$ and absolute convergence for ϵ small enough. Since $Z(\epsilon)$ equals ϵ times $g^{T;1} := z_1 Y$ plus a bounded variation term, invisible to the diamond product, it is useful with regard to (2.7) to write

$$(2.10) \qquad (Z + \Lambda^T) \diamond (Z + \Lambda^T) = \left(\sum_{m \geq 1} \epsilon^m g^{T;m} \right) \diamond \left(\sum_{n \geq 1} \epsilon^n g^{T;n} \right).$$

In essence, the proof is finished by matching powers of ϵ in (2.7), upon substitution of (2.8), (2.9) and (2.10). As it stands, the a.s. convergence of (2.9) is insufficient to interchange with diamond products[9] but rather than strengthening the convergence to $\mathcal{H}^p$-sense

---

[8]Expand the exponential function and employ (classical resp. conditional) dominated convergence; see, for example, classical textbook references [28, 31].

[9]Counter example: consider a Brownian motion B with adapted BV approximations $B_t^n := h^{-1} \int_{t-h}^t B_s \, ds$. Then $B_t^n \to B_t$ a.s. but $B^n \diamond B^n \equiv 0$, whereas $(B \diamond B)_t(T) = T - t$.

(cf. Proposition 2.1) we observe that it suffices to work in $\mathbb{R}[[\epsilon]]$, the algebra of formal power series in one indeterminate denoted ϵ. We then view (2.5) as $\mathbb{R}[[\epsilon]]$-valued martingale, and more precisely with values in the invertible series started at 1. In particular, $M_t^{-1}, \log(M_t)$ etc. are well defined and obey the same rules of calculus, both classical and Itô's. Stochastic exponential and logarithm are then defined and used in exactly the same way as before, until we arrive at (2.7), but now as relation between $\mathbb{R}[[\epsilon]]$-valued semimartingales. We can now safely match powers of ϵ in (2.7) to see that $g^{T,2}$ agrees with $\mathbb{G}^2(T, .)$ and satisfies the same recursion. This concludes the proof.

REMARK 2.3 (Exponential integrability). (a) In general, one cannot replace $\langle Y \rangle_T$ by Y_T when making the correct exponential integrability assumption in part (ii) of Theorem 1.1. An explicit example, again taking $T = \infty$ without loss of generality, of a (continuous, uniformly integrable) martingale Y such that Y_∞ has exponential moments, while $\langle Y \rangle_\infty$ does not, was kindly provided to us by Johannes Ruf and is here reproduced: Pick a nonnegative r.v. R with $\mathbb{E}[e^{cR}] < \infty$, some $c > 0$, while $\mathbb{E}[e^{cR^2}] = +\infty$ for all $c > 0$ (e.g., a standard exponential). Let B be an independent standard Brownian motion, stopped at the first time τ when $B_\tau \in \{-R, R\}$ and note $\tau < \infty$ a.s. Then the stopped process $Y = B^\tau$ is a uniformly integrable martingale (indeed, $Y \in \mathcal{H}^1$ with its running supremum bounded by $R \in L^1$). Clearly $Y_\infty = R$ has exponential moments. On the other hand,

$$\mathbb{E}e^{c\langle Y \rangle_\infty} = \mathbb{E}e^{c\tau} = \mathbb{E}[\mathbb{E}[e^{c\tau}|R]] \geq \mathbb{E}[e^{c\mathbb{E}[\tau|R]}] = \mathbb{E}e^{cR^2} = +\infty.$$

using the standard fact that the expected time for Brownian motion to hit $\pm r$ is r^2.

(b) One the positive side, one can employ sharp martingale estimates ([5], Theorem 3.1), which state (recall $O(p)$-growth of moments reflects exponential moments)

$$\left\| \sqrt{\langle Y \rangle_T} \right\|_{L^p(\mathbb{P})} \leq 2p^{1/2} \|Y_T\|_{L^p(\mathbb{P})}, \quad p \geq 2.$$

It follows that $A_T = Y_T \in L^\infty$ is sufficient to guarantee exponential integrability of $\langle Y \rangle_T$ and then, trivially, also of any linear combinations with Y_T.

2.1.2. *Part* (i). Assume now Y_T has moments of all orders, by BDG it is clear that the same is true for $(Y_T, \langle Y \rangle_T)$. As is well known, the bivariate, c.f. it is then a smooth (but not necessarily analytic) function of its transform variables. This entails an asymptotic expansion, as $(z_1, z_2) = (0 + iy_1, x_2 + iy_2) \in \mathbb{C}^2$, with $x_2 \leq 0$, tends to zero,

$$(2.11) \qquad \log \mathbb{E}[e^{z_1 Y_T + z_2 \langle Y \rangle_T}] \sim \sum_{m_1 \geq 1, m_2 \geq 1} \kappa^{(m_1, m_2)} \frac{z_1^{m_1} z_2^{m_2}}{m_1! m_2!},$$

which can be seen as defining equation of the (bivariate) cumulants, computable by taking derivatives of the left-hand side at $\xi = 0$. Nothing changes upon working with (time-t) conditional expectations, so that the existence of an expansion as asserted in (1.1) is a priori clear, with initial terms of the form

$$\log \mathbb{E}_t[e^{z_1 Y_T + z_2 \langle Y \rangle_T}] \sim z_1 \mathbb{E}_t Y_T + z_2 \mathbb{E}_t \langle Y \rangle_T + \frac{z_1^2}{2} \mathbb{V}_t Y_T \cdots$$

$$(2.12)$$

$$= z_1 Y_t + z_2 Y_t + \left(\frac{z_1^2}{2} + z_2 \right)(Y \diamond Y)_t(T)) + \cdots$$

For $k > 2$ then, $\mathbb{G}_t^k(T; z)$ are precisely given by the sum of all conditional (m_1, m_2)-cumulants with $k = m_1 + 2m_2$. This is precisely captured by replacing Y by ϵY, hence $z_1 Y_T + z_2 \langle Y \rangle_T$ by $\epsilon z_1 Y_T + z_2 \epsilon^2 \langle Y \rangle_T$, followed by extracting the $[\epsilon^k]$-power of the resulting formal power series, exactly as was done at the final stage of part (ii), above.

2.2. *Proof of Theorem* 1.2.   Part (ii) is an immediate corollary of Theorem 1.1, applied with $(z, 0) \in \mathbb{C}^2$, that is, $z_2 = 0$.

We give a proof of part (i) by localization, but also comment on a direct "Hermite" proof below.

LEMMA 2.1.   *Assume A has N moments, $N \in \mathbb{N}$. Then the recursion* (1.3) *is well defined for $k \leq N$ and yields $(\mathbb{K}_t^k(T) : 0 \leq t \leq T)$ as a semimartingale with a $L^{N/k}$-integrable martingale part and a $L^{N/k}$-integrable bounded variation* (BV) *component.*

PROOF.   Obviously $\mathbb{K}^1 = \mathbb{E}_\bullet A_T \in L^N$ so the statement is true for $k = 1$. By (finite) induction, assume the statement holds true for all $k = 1, \ldots, n$, so that $\mathbb{K}^k \in L^{N/k}$, $\mathbb{K}^{n+1-k} \in L^{N/(n+1-k)}$. As long as $n < N$, equivalently $N/(n+1) \geq 1$, Proposition 2.1 applies and tells us $\mathbb{K}^k \diamond \mathbb{K}^{n+1-k} \in L^{N/(n+1)}$. By the very recursion formula (1.3) this entails $\mathbb{K}^{n+1} \in L^{N/(n+1)}$. □

Given A_T, $\mathcal{F}_T$-measurable with $N \in \mathbb{N}$ finite moments—but whose mgf is not necessarily finite—we work with its two-sided truncation $(-\ell \vee A_T \wedge \ell) =: A_T^\ell$, followed by careful passage $\ell \to \infty$.

Theorem 1.1 applies to A_T^ℓ (bounded!) and hence shows that the $\mathbb{K}^{\ell;n}$, $n = 1, 2, \ldots$ defined by the recursion (1.3), started with $\mathbb{K}^{\ell;1} = \mathbb{E}_\bullet A_T^\ell$, are well defined and yield (up to a factorial factor) the nth conditional cumulants of A_T^ℓ. It is clear from dominated-convergence that (first N) moments of A_T^ℓ converge to those of A_T, and the same is true for cumulants, which are polynomial expression of the moments. (Nothing changes when working time-t conditionally other than using conditional dominated-convergence, which gives a.s. convergence). That settles pointwise a.s. convergence, $\mathbb{K}_t^{\ell;n} \to \mathbb{K}_t^n$, any $n \leq N$.

It remains to see that the diamond recursion

$$(2.13) \qquad \mathbb{K}_t^{\ell;n+1}(T) = \frac{1}{2} \sum_{k=1}^n (\mathbb{K}^{\ell;k} \diamond \mathbb{K}^{\ell;n+1-k})_t(T).$$

is also stable under this passage to the limit, as long as $n < N$. For $N = 1$ there is nothing to show so fix an integer $N > 1$. For $A_T \in L^N$, it follows from $A_T^\ell \to A_T$ in L^N that $\mathbb{K}^{\ell;1} = \mathbb{E}_\bullet A_T^\ell \to \mathbb{K}^1 = \mathbb{E}_\bullet A_T$ in $\mathcal{H}_T^N$. In view of Proposition 2.1, it suffices to show that $\mathbb{K}^{\ell;k} \to \mathbb{K}^k$ in $\mathcal{H}_T^{N/k}$ for all $k < N$. But knowing this for $k = 1$, the desired convergence can be seen inductively from (2.13) and estimate (2.2) of Proposition 2.1; the limit is then necessarily equal to the (pointwise a.s.) limit $\mathbb{K}^k$.

The asymptotic expansion is then a straightforward consequence of validity of the expansion of part (i) for all integers N. This finishes the proof of (i).

REMARK 2.4 (Hermite).   A direct proof of part (i) that relates $\mathbb{K}^n$, $n = 1, 2, 3, \ldots$ with the corresponding cumulants is possible based on Hermite polynomials. To understand the argument, start with the proof of part (ii), specialized to the cumulant case, so that the key identity reads

$$\mathbb{E}_t e^{\epsilon Y_T} = e^{\epsilon Y_t + \Lambda_t^T(\epsilon)}.$$

Rewritten with $Q_t^T(\epsilon) := -2(\mathbb{K}_t^2(T) + \epsilon \mathbb{K}_t^3(T) + \epsilon^2 \mathbb{K}_t^4(T) + \cdots)$,

$$\mathbb{E}_t e^{\epsilon Y_T} = \mathbb{E}_t e^{\epsilon Y_T - \frac{\epsilon^2}{2} Q_T^T(\epsilon)} = e^{\epsilon Y_t - \frac{\epsilon^2}{2} Q_t^T(\epsilon)},$$

we can deduce, by definition of Hermite polynomials ([26], Chapter IV.3), martingality of

$$e^{\epsilon Y_t - \frac{\epsilon^2}{2} Q_t^T(\epsilon)} = \sum_{n \geq 0} \frac{\epsilon^n}{n!} H_n(Y_t, Q_t^T(\epsilon)).$$

By taking $(\partial/\partial\epsilon)^n|_{\epsilon=0}$ we obtain a graded family of martingales, starting with ($n = 2$)

$$t \mapsto H_2(Y_t, Q_t^T(0)) = Y_t^2 - \frac{1}{2}\mathbb{K}_t^2(T).$$

Applying Itô's formula over $[t, T]$ and taking t-conditional expectation then identifies $\mathbb{K}_t^2(T)$ correctly as $\mathbb{E}_t \langle M \rangle_{t,T} = (M \diamond M)_t(T)$. Using suitable relations between Hermite polynomials, see also [9], this argument extends to $n > 2$ and provides an alternative route to the $\mathbb{K}$-recursion.

3. Relation to other works. We already commented in detail on Alòs et al. [2] and Lacoin et al. [25].

Growing exponential expansions on trees are reminiscent of the Magnus expansion, a type of continuous Baker–Campbell–Hausdorff formula, with classical recursions based on rooted binary trees; Iserles–Nørsett [21]. And yet, the $\mathbb{F}, \mathbb{K}, \mathbb{G}$ expansions are of a fundamentally different nature, for noncommutative algebra plays no role: our setup is one of multivariate random variables, associated martingales and their quadratic variation processes. (But see the recent preprint [12] for expansions related to signature cumulants, a noncommutative generalisation of classical cumulants).

In a Markovian situation, our expansion can be related to perturbative expansion of a "KPZ" type equation, by which we here mean a nonlinear parabolic partial differential equation of HJB type. We make this explicit in the case when $A = f(B)$, for a Brownian motion B and suitable f, in which case the $\mathbb{K}$'s are described by a cascade of linear PDEs, detailed in Section 4.5, indexed by trees such as (1.7), in the exact same way as the "Wild expansion" used in Hairer's KPZ analysis [19]. (This link is restricted to the algebraic part of the expansions and rough paths, analytic renormalisation etc. play no role here.)

That said, computing $\log \mathbb{E}_t[e^{\varepsilon A_T}]$ may also be viewed as a (linear) backward SDE with "Markovian" terminal data given by $e^{\varepsilon A_T} = e^{\varepsilon f(B_T)}$; upon suitable exponential change of variables this becomes a quadratic BSDE as studied by Kobylanski [24], Briand–Hu [4] and many others, in the weakly nonlinear regime (BSDE driver of order ϵ). Yet another point of view comes from Dupire's functional Itô-calculus [8] which would lead to similar (at least formal) computations as conducted in Section 4.5, for general $\mathcal{F}_T$-measurable A_T. And yet another point of view comes from the Boué–Dupuis [3] formula which gives an exact variational representation of $\log \mathbb{E}[e^{A_T}]$ when A_T is a sufficiently integrable measurable function of Brownian motion up to time T; here (1.5) can be viewed as an asymptotic solution to the Boué–Dupuis variational problem in the weakly nonlinear regime.

In Section 4.6, we compare Theorem 1.2 to the work of Nourdin–Peccati [32] where the authors use Malliavin integration by parts to describe cumulants of certain Wiener functionals, and notably compute cumulants of elements in a fixed Wiener chaos. (The ability to work under exponential resp. (sub)exponential integrability assumptions is crucial to deal with second resp. higher order chaos.) An important element in the second chaos with explicit cumulants is provided by *Lévy's stochastic area*; our (short) proof of its cumulant generating function should be compared with the combinatorial tour de force of [27], based on (signature) moments. It is conceivable that the multivariate cumulant formula applied to multidimensional Brownian motion and Lévy area (a.k.a. the Brownian rough path) provides new input into classical problems of stochastic numerics [15].

In Section 4.7, we apply Theorem 1.1 to establish a formula for the joint mgf of a process X, its quadratic variation $\langle X \rangle$, and $\mathbb{E}_t[d\langle X \rangle_T / dT]$, quantities (a.k.a. log-price, total variance, forward variance) that play an important role in stochastic financial modeling. Our expansion is most convenient for models written in forward variance form, state of the art in quantitative finance. In particular, the full expansion is computable in affine forward variance models, which includes the popular rough Heston model [10].

At last, we mention two recent preprints, [14] (v2 from 2021) and [12], that (among others) relax the continuity assumption. In particular, the correct diamond product definition for general semimartingales A, B is given by $A^c \diamond B^c$ where A^c, B^c are the respective continuous local martingale parts of A, B. Leaving details to these works, we note that generalized cumulant recursion comes with extra terms that account for the jumps. While such generalizations are clearly of interest, they do not preserve the quadratic structures which one has in the continuous setting of the present article. In particular, the resulting binary diamond trees and forests are really a feature of the continuous semimartingale setting.

4. Examples.

4.1. *Brownian motion.*

EXAMPLE 4.1 (Brownian motion with drift).   Let $A_t = \sigma B_t + \mu t$. Then

$$\mathbb{K}_t^1(T) = \sigma B_t + \mu T = A_t + \mu(T - t), \qquad \mathbb{K}_t^2(T) = \frac{1}{2}(\mathbb{K}^1 \diamond \mathbb{K}^1)_t(T) = \frac{1}{2}\sigma^2(T - t).$$

and $\mathbb{K}^k \equiv 0$ for all $k \geq 3$. These are the cumulants of $A_T - A_t \sim N(\mu(T - t), \sigma^2(T - t))$, as predicted by Theorem 1.1 (or Theorem 1.2) and the $\mathbb{K}$-forest expansion of the cumulant generating function (1.5) is trivially convergent (with infinite convergence radius).

EXAMPLE 4.2 (Stopped Brownian motion).   Consider the martingale $A = B^\tau$, standard Brownian motion B stopped at reaching ± 1. We compute

$$\mathbb{K}_t^1(T) = \mathbb{E}_t B_T^\tau = B_t^\tau = B_{t \wedge \tau}, \qquad \mathbb{K}_t^2(T) = \frac{1}{2}\mathbb{E}_t\langle B^\tau \rangle_{t,T} = \frac{1}{2}(\mathbb{E}_t(\tau \wedge T) - \tau \wedge t), \dots.$$

The second quantity equals the conditional variance $\mathbb{V}_t(B_T^\tau) = \mathbb{E}_t(B_T^\tau)^2 - (B_t^\tau)^2$, and thus "contains" familiar identities from optional stopping. With $T = \infty$, $A_T = B_\tau$ takes values ± 1 with equal probability. This is a bounded random variable, with globally defined and real analytic time-t conditional cgf given by

$$\Lambda_t(x) = \log\left(\frac{1}{2}[(1 + B_t^\tau)e^x + (1 - B_t^\tau)e^{-x}]\right).$$

Its convergence radius is random through the value of $B_t^\tau = B_t^\tau(\omega) \in [-1, 1]$. For instance, when $t = 0$, so that $B_t^\tau = 0$, we have $\Lambda_0(x) = \log\cosh(x)$ with a $\mathbb{K}$-forest expansion (1.5) of finite convergence radius $\rho_0 = \pi/2$. On the other hand, on the event $E := \{B_t^\tau = \pm 1\}$, the cgf $\Lambda_t(x)$ trivially takes the value $\log e^{\pm x} = \pm x$ so that, on E, we have $\rho_t(\omega) = +\infty$.

4.2. *Lévy area.*   We give a new proof of P. Lévy's theorem, which compares favourably with other available proofs [20, 27].

THEOREM 4.1 (P. Lévy).   *Let $\{X, Y\}$ be 2-dimensional standard Brownian motion and stochastic ("Lévy") area be given by*

$$\mathcal{A}_t = \int_0^t (X_s \, dY_s - Y_s \, dX_s).$$

Then for $T \in (-\frac{\pi}{2}, \frac{\pi}{2})$,

$$\mathbb{E}_0[e^{\mathcal{A}_T}] = \frac{1}{\cos T} = \exp\left(\int_0^T \tan s \, ds\right).$$

As a warmup, we compute the first few cumulants, using the $\mathbb{K}$-recursion from Theorem 1.2. (We note $\langle \mathcal{A}_T \rangle_T \notin L^\infty$, so that, strictly speaking, the result in [25] is not applicable.) By a direct computation (or a very special case of Theorem 4.2),

$$\mathbb{K}^2 = \frac{1}{2} \diamond\!\!\!\! \diamond = \frac{1}{2}(\mathcal{A} \diamond \mathcal{A})_t(T) = \frac{1}{2} \int_t^T \left(\mathbb{E}_t[X_s^2] + \mathbb{E}_t[Y_s^2]\right) ds$$

$$= \frac{1}{2}(T-t)^2 + \frac{1}{2}(X_t^2 + Y_t^2)(T-t)$$

$$\implies \quad d\mathbb{K}_s^2 = (X_s \, dX_s + Y_s \, dY_s)(T-s) + \mathrm{BV}.$$

With $d\mathbb{K}_s^1 = X_s \, dY_s - Y_s \, dX_s$ we see that the third forest vanishes,

$$\mathbb{K}^3 = \mathbb{K}^1 \diamond \mathbb{K}^2 = \mathbb{E}_t\left[\int_t^T d\langle \mathbb{K}^1, \mathbb{K}^2 \rangle_s\right] = \mathbb{E}_t\left[\int_t^T [XY \, d\langle Y \rangle_s - YX \, d\langle X \rangle_s](T-s)\right] = 0.$$

LEMMA 4.1.   *Set $J_t^k(T) := \frac{(T-t)^k}{k} + \frac{1}{2}(X_t^2 + Y_t^2)(T-t)^{k-1}$. Then*

$$(J^j \diamond J^k)_t(T) = \frac{2}{j+k-1} J_t^{(j+k)}(T).$$

PROOF.   With $dJ_s^k(T) = (X_s \, dX_s + Y_s \, dY_s)(T-s)^{k-1} + \mathrm{BV}$, computation as above.   □

Note $\mathbb{K}_t^n(T) = \alpha_n J_t^n(T)$ for $n = 2, 3$ with $\alpha_2 = 1, \alpha_3 = 0$. Assume by induction that this holds true up to the even/odd pair $(n-2, n-1)$, with $\alpha_{n-1} = 0$. Then the cumulant recursion gives, with sum over even integers $j, k \geq 2$,

$$2\mathbb{K}^n = \sum_{j+k=n} \alpha_j \alpha_k J^j \diamond J^k = \sum_{j+k=n} \alpha_j \alpha_k \frac{2}{n-1} J^n =: 2\alpha_n J^n,$$

while $\mathbb{K}^{n+1} = \mathbb{K}^1 \diamond \mathbb{K}^n$ (use $\mathbb{K}^3, \ldots, \mathbb{K}^{n-1} = 0$), which vanishes for the same reason as $\mathbb{K}^3$. (This completes the induction.) Hence,

$$\alpha_n = \frac{1}{n-1}(\alpha_2 \alpha_{n-2} + \alpha_4 \alpha_{n-4} + \cdots + \alpha_{n-2} \alpha_2), \quad \alpha_2 = 1.$$

Evaluated at $t = 0$, using $J_0^n(T) = T^n/n$, we thus see the $\mathbb{K}$-expansion take the form

$$\alpha_2 T^2/2 + \alpha_4 T^4/4 + \alpha_5 T^6/6 + \cdots = T^2/2 + \frac{1}{3}T^4/4 + \frac{2}{15}T^6/6 + \cdots,$$

where one starts to see the integrated expansion of $\tan(T) = T + \frac{1}{3}T^3 + \frac{2}{15}T^5 + \cdots$, integrated in time. To see that this is really so, check that $f(T) := \sum_{j \in 2\mathbb{N}} \alpha_j T^{j-1}$ satisfies the ODE $f'(T) = f(T)^2 + 1$, with $f(0) = 0$, which identifies $f \equiv \tan$.

4.3. *Diamond products of iterated stochastic integrals.*   Lévy's area is a particularly important example of Brownian iterated integrals, for which we have given explicit diamond computations. We now present systematic diamond computations for iterated stochastic integrals, which play a fundamental role in stochastic numerics and rough path theory [23, 29]. They are defined as follows. For a word $a = i_1 \ldots i_m$ of length m, with letters in

$\mathbb{A} = \{i : 1 \leq i \leq d\}$, write ai for the word (of length $m + 1$) obtained by concatenation of a with the letter i. Given a d-dimensional Brownian motion (B^i), introduce the iterated Itô resp. Stratonovich integrals

$$B^{ai} = \int_0^{\bullet} B^a \, dB^i; \qquad \widehat{B}^{ai} = \int_0^{\bullet} \widehat{B}^a \circ dB^i;$$

set also $B^{\phi} = \hat{B}^{\phi} = 1$ when ϕ is the empty word. One extends these definitions by linearity to linear combination of words, which becomes a commutative algebra under the shuffle product. It is inductively defined by $a \,\shuffle\, \varnothing = \varnothing \,\shuffle\, a = a$, $ai \,\shuffle\, bj = (a \,\shuffle\, bj)i + (ai \,\shuffle\, b)j$, for any words a, b and letters i, j; here $\varnothing$ denotes the empty word. (For instance, $12 \,\shuffle\, 3 = 312 + 132 + 123$). Then the remarkable identity

$$\hat{B}_t^a \hat{B}_t^b = \hat{B}_t^c; \quad c = a \,\shuffle\, b,$$

holds true (and reflects the validity of the usual chain rule for Stratonovich integration). In contrast, resolving $B_t^a B_t^b$ requires quasi-shuffle (Itô formula) which we will not introduce here. Let us also recall Fawcett's formula, as found in [13], Theorem 3.9,

$$(4.1) \qquad \mathbb{E}_0 \hat{B}_{0,1}^a = \langle e_{\otimes}^{\mathbb{I}/2}, a \rangle =: \sigma_a,$$

where $\mathbb{I}$ is the identity matrix (seen as 2-tensor). In other words, $\mathbb{E}_0 \hat{B}_{0,1}^a$ equals $1/(2^n n!)$ whenever $a = i_1 i_1 i_2 i_2 \cdots i_n i_n$ for some letters $i_1, \ldots i_n$, and zero else.

THEOREM 4.2. *Consider two (possibly empty) words a, b with respective length $|a|, |b|$ and letters $i = j \in \mathbb{A}$. Then*
(Itô)

$$(B^{ai} \diamond B^{bj})_t(T) = B_t^a B_t^b (T - t) + \frac{T - t}{1 + (|a| + |b|)/2} (B^a \diamond B^b)_t(T)$$

(Stratonovich)

$$(\hat{B}^{ai} \diamond \hat{B}^{bj})_t(T) = \hat{B}_t^a \hat{B}_t^b (T - t) + \hat{B}_t^a \sigma_b \frac{(T - t)^{\frac{|b|}{2} + 1}}{\frac{|b|}{2} + 1}$$

$$+ \hat{B}_t^b \sigma_a \frac{(T - t)^{\frac{|a|}{2} + 1}}{\frac{|a|}{2} + 1} + \frac{T - t}{1 + \frac{|a| + |b|}{2}} (\hat{B}^a \diamond \hat{B}^b)_t(T).$$

In case $i \neq j$, both diamond products vanish.

PROOF (Itô). By Itô isometry, and the product rule $B_s^a B_s^b = B_t^a B_t^b + \cdots + \langle B^a, B^b \rangle_{t,s}$, with omitted martingale increment $\int_t^s (B^a \, dB^b + B^b \, dB^a)$,

$$(B^{ai} \diamond B^{bj})_t(T) = \mathbb{E}_t \langle B^{ai}, B^{bj} \rangle_{t,T} = \delta^{ij} \mathbb{E}_t \int_t^T B_s^a B_s^b \, ds$$

$$= \delta^{ij} \int_t^T (B_t^a B_t^b + (B^a \diamond B^b)_t(s)) \, ds.$$

From the scaling properties of Brownian motion, the time t-conditional law of $(B^a \diamond B^b)_t(s)$ is equal to the law of

$$\left(\frac{s - t}{T - t} \right)^{\frac{|a| + |b|}{2}} (B^a \diamond B^b)_t(T),$$

followed by an immediate integration over $s \in [t, T]$.

(Stratonovich.) Note that

$$\hat{B}^{ai} = \int \hat{B}^a \, dB^i + \mathrm{BV}$$

so that, as in the Itô case (but now with noncentered dots),

$$(\hat{B}^{ai} \diamond \hat{B}^{bj})_t(T) = \delta^{ij} \mathbb{E}_t \int_t^T \hat{B}_s^a \hat{B}_s^b \, ds$$

$$= \delta^{ij} \int_t^T ds \left(\hat{B}_t^a \hat{B}_t^b + \hat{B}_t^a \mathbb{E}_t \hat{B}_{t,s}^b + \hat{B}_t^b \mathbb{E}_t \hat{B}_{t,s}^a + (\hat{B}^a \diamond \hat{B}^b)_t(s) \right).$$

From Brownian scaling and Fawcett's formula (4.1), we have

$$\mathbb{E}_t \hat{B}_{t,s}^b = (s - t)^{|b|/2} \mathbb{E}_0 \hat{B}_{0,1}^b = (s - t)^{|b|/2} \langle e_\otimes^{\mathbf{I}/2}, b \rangle =: (s - t)^{|b|/2} \sigma_b$$

and so see, with $i = j$,

$$(\hat{B}^{ai} \diamond \hat{B}^{bi})_t(T) = \hat{B}_t^a \hat{B}_t^b(T - t) + \hat{B}_t^a \sigma_b \frac{(T - t)^{\frac{|b|}{2}+1}}{\frac{|b|}{2} + 1}$$

$$+ \hat{B}_t^b \sigma_a \frac{(T - t)^{\frac{|a|}{2}+1}}{\frac{|a|}{2} + 1} + \frac{T - t}{1 + \frac{|a|+|b|}{2}} (\hat{B}^a \diamond \hat{B}^b)_t(T). \qquad \square$$

EXAMPLE 4.3 (Cameron–Martin formula). Following [26], Chapter XI.1, the Laplace transform of $\int_0^1 B_s^2 \, ds$ is given by

$$(\cosh \sqrt{2\lambda})^{-1/2} = \exp\left(-\frac{1}{2}\lambda + \frac{1}{6}\lambda^2 - \frac{4}{45}\lambda^3 + \cdots \right).$$

We can elegantly obtain this from the $\mathbb{G}$-expansion applied to the iterated Itô integral $Y_t = \int_0^t B_s \, dB_s$, $\langle Y \rangle_1 = \int_0^1 B_s^2 \, ds$, so that $\mathbb{G}^2 = -\lambda(Y \diamond Y)_t(T)$, and for $k > 2$: $\mathbb{G}^k = \frac{1}{2} \sum_{j=2}^{k-2} \mathbb{G}^{k-j} \diamond \mathbb{G}^j$. A computation similar to the one given in the Lévy area example, shows that the nth cumulant is given by $q_n/(2n)$, with recursion

$$q_n = \frac{2}{2n - 1}(q_1 q_{n-1} + \cdots + q_{n-1} q_1), \quad q_1 = 1,$$

from which one can also obtain the explicit functional form.

4.4. *Bessel process.* We use the $\mathbb{G}$-expansion to establish some identities of the *Bessel square process* with (time-dependent) dimension $\delta = \delta(t) \geq 0$, given as solution to

$$dX_t = 2\sqrt{X_t} \, dB_t + \delta(t) \, dt.$$

As in the case of Lévy area, the $\mathbb{G}$-expansion with diamond calculus compares (very) favourably with other available proofs, cf. [26], Chapter XII. For nonnegative, bounded measurable $\mu = \mu(t)$, set $Y_T := \frac{1}{2} \int_0^T \sqrt{\mu} \, dX_s$, hence $dY_s = \frac{1}{2} \sqrt{\mu} \, dX_s$ so that the $\mathbb{G}$-expansion[10] gives the Laplace transform of the weighted Bessel average

$$\langle Y \rangle_T = \int_0^T X_s \mu(s) \, ds,$$

---

[10] Applied with $(z_1, z_2) = (0, \lambda)$, so that only even terms appear in the $\mathbb{G}$-expansion.

starting with

$$\mathbb{G}^2 = -\lambda Y \diamond Y = \cdots = -\lambda\Big(X_t \int_t^T \mu(s)\,ds + \int_t^T \int_t^s \delta(r)\mu(s)\,dr\,ds\Big),$$

followed by $\mathbb{G}^3 = 0$. (Use $\mathbb{E}_t X_s = X_t + \int_t^s \delta(r)\,dr, s \geq t$.) By Lemma 4.2 below,

$$\log \mathbb{E}_t \exp\Big(-\lambda \int_t^T X_s \mu(s)\,ds\Big) = \sum_{n \geq 2,\text{even}} \frac{(-\lambda)^{n/2}}{2}\Big(X_t \Gamma_n(t) + \int_t^T \delta(r)\Gamma_n(r)\,dr\Big),$$

with $\psi(t) := \sum_{n \geq 2,\text{even}}(-\lambda)^{n/2}\Gamma_n(t)$ rewritten as

$$\log \mathbb{E}_t \exp\Big(-\lambda \int_t^T X_s \mu(s)\,ds\Big) = \frac{1}{2}X_t \psi(t) + \frac{1}{2}\int_t^T \delta(r)\psi(r)\,dr.$$

Thanks to (4.2) ψ, is immediately identified as (unique) backward ODE solution to $-\dot{\psi} = \lambda\dot{\Gamma}_2 + \psi^2 = -2\lambda\mu + \psi^2$ with terminal data $\psi(T) = 0$. (This constitutes a new and elegant route to Corollary 1.3, Theorem 1.7 in [26], Chapter XI.1, therein only given for constant δ, in which case the ODE can be written as $\phi'' = 2\lambda\mu\phi, \phi(t) = \exp(-\int_t^T \psi(s)\,ds)$.)

We can be more specific when μ is explicit. For instance, the conditional Laplace transform of X_T is obtained by taking $\mu(s)\,ds = \delta_T(ds)$, justified by an easy approximation argument (e.g., $\mu^n = \frac{1}{n}1_{[T-1/n],T]}$), in which case the ODE becomes $-\dot{\psi} = \psi^2$ with terminal data $\psi(T) = -2\lambda$, with unique solution $\psi(t) = -2\lambda/(1 + 2\lambda(T - t))$. Specializing to constant Bessel dimension $\delta(r) \equiv \delta, t = 0$, and $X_0 = x$, we obtain

$$\mathbb{E}\big[\exp(-\lambda X_T)\big] = \exp\Big(-\lambda x/(1 + 2\lambda T) - \frac{\delta}{2}\log(1 + 2\lambda T)\Big)$$

$$= (1 + 2\lambda T)^{-\delta/2}\exp(-\lambda x/(1 + 2\lambda T)),$$

in agreement with [26], Chapter XI.1. (For what it's worth, the controlled ODE structure of $\dot{\psi}$, with additive noise $W := \Gamma_2$, makes sense for any deterministic càdlàg path W, and allows to compute the transform of any "rough" integral, $\int_0^T X_s(\omega)\,dW_s := X_{0,T}W_{0,T} - \int_0^T W_s^- \,dX_s$, with final integral in Itô sense.)

LEMMA 4.2. *The general even/odd pair in the $\mathbb{G}$-expansion is of the form*

$$\mathbb{G}_t^n = \frac{(-\lambda)^{n/2}}{2}\Big(X_t \Gamma_n(t) + \int_t^T \delta(r)\Gamma_n(r)\,dr\Big), \quad \mathbb{G}^{n+1} \equiv 0,$$

with $\Gamma_n(t)$ determined by $\Gamma_2(t) = 2\int_t^T \mu(s)\,ds$ and the recursion, for even $n \geq 4$,

$$(4.2) \qquad -\dot{\Gamma}_n = \Gamma_2\Gamma_{n-2} + \Gamma_4\Gamma_{n-4} + \cdots + \Gamma_{n-2}\Gamma_2, \quad \Gamma_n(T) = 0.$$

PROOF. The statement was seen to be correct for $n = 2$. Assume by induction that it holds true, for all even/odd pairs up to $(n - 2, n - 1)$. In particular then $d\mathbb{G}^k = (-\lambda)^{k/2}\Gamma_k\sqrt{X}\,dB + d(\text{BV})$, for even $k < n$, and by the $\mathbb{G}$ recursion, with sums below always over even integers $j, k \geq 2$,

$$2\mathbb{G}^n = \sum_{j+k=n} \mathbb{G}^j \diamond \mathbb{G}^k = \sum_{j+k=n} \mathbb{E}_t \int_t^T d\langle \mathbb{G}^j, \mathbb{G}^k \rangle = (-\lambda)^{n/2} \sum_{j+k=n} \mathbb{E}_t \int_t^T \Gamma_j(s)\Gamma_k(s)X_s\,ds.$$

Set $\Gamma_n(t) := \sum_{j+k=n} \int_t^T \Gamma_j(s)\Gamma_k(s)\,ds$ (j, k, n even) and use $\mathbb{E}_t X_s = X_t + \int_t^s \delta(r)\,dr$ to conclude.($\mathbb{G}^{n+1} = 0$ is clear.) The ODE statement is also immediate. $\quad\square$

4.5. *A Markovian perspective and smooth KPZ.* The previously encountered trees $\{\circ, \lor\!\!\!\circ, \ldots\}$ from Section 1.2 were famously used in [19] as a minimal choice in indexing a *finite* expansion of the $(1+1)$-dimensional KPZ equation, with additional analytical (rough path) arguments to deal with the remainder.[11] The appearance of the same trees is more than a coincidence, as we shall now see. Consider functions $h_T = h_T(x)$ and $\xi = \xi(t, x)$ on $\mathbb{R}^d$ and $[0, T] \times \mathbb{R}^d$ respectively, for simplicity taken bounded with bounded derivatives of all orders, and consider

$$(4.3) \qquad A_T := h_T(B_T) + \int_0^T \xi(s, B_s)\, ds$$

with a standard d-dimensional Brownian motion B. Then

$$\mathbb{E}_t e^{\lambda A_T(\omega) - \lambda \int_0^t \xi(s, B_s)\, ds} = \mathbb{E}_t e^{\lambda\{h_T(B_T) + \int_t^T \xi(s, B_s)\, ds\}} =: e^{\lambda h(t, B_t)} =: z(t, B_t)$$

and $z = z(t, x)$ satisfies the Kolmogorov backward equation $-\dot{z} = \frac{1}{2}\Delta z + \lambda z \xi$, with terminal data $e^{\lambda h_T}$. (Equivalently, the above is the Feynman–Kac representation formula for this backward PDE.) By change of variable, $h := \log(z)/\lambda$, a.k.a Cole–Hopf transform, we obtain the $(d+1)$-dimensional KPZ equation

$$(4.4) \qquad \left(-\partial_t - \frac{1}{2}\Delta\right)h = \frac{\lambda}{2}(\nabla h \cdot \nabla h) + \xi, \quad h(T, \cdot) = h_T,$$

with smooth noise $\xi = \xi(t, x)$ and written in backward form. Following Hairer [19], who attributes such expansions to Wild (1955), one has the (formal) tree indexed expansion[12]

$$(4.5) \qquad h = u^\circ + \lambda u^{\lor\!\!\circ} + 2\lambda^2 u^{\cdots} + \lambda^3 u^{\cdots} + 4\lambda^3 u^{\cdots} + \cdots = \sum_\tau \lambda^{|\tau|-1} u^\tau$$

with sum over all binary trees with $|\tau| \geq 1$ leaves. More specifically, u° is the unique (bounded) solution to the linear problem ($\lambda = 0$), and then, whenever $\tau = [\tau_1, \tau_2]$, the root joining of trees τ_1 and τ_2, we get $u^{[\tau_1, \tau_2]} = u^{[\tau_2, \tau_1]}$ from[13]

$$(4.6) \quad \left(-\partial_t - \frac{1}{2}\Delta\right)u^\tau = \frac{1}{2}(\nabla u^{\tau_1} \cdot \nabla u^{\tau_2}), \qquad 2u^\tau = \mathcal{K} \star (\nabla u^{\tau_1} \cdot \nabla u^{\tau_2}) =: u^{\tau_1} \diamond u^{\tau_2},$$

where $\mathcal{K} \star (\ldots)$ denotes space–time convolution with the heat kernel. (Thanks to our strong assumptions on forcing ξ and terminal data h_T, the recursion for the $u^\tau = u^\tau(t, x)$ is well defined and all u^τ smooth.) We can then rewrite (4.5) as

$$\log \mathbb{E}^{t, x} e^{\lambda A_T(\omega)} = \lambda h(t, x) = \sum_{|\tau| \geq 1} \lambda^{|\tau|} u^\tau = \sum_{n \geq 1} \lambda^n \sum_{\tau : |\tau| = n} u^\tau =: \sum_{n \geq 1} \lambda^n K^n(t, x; T).$$

with $K^1 = u^\circ$ and then recursively

$$(4.7) \qquad K^{n+1} = \sum_{\tau : |\tau| = n+1} u^\tau = \sum_{\cdots} u^{[\tau_1, \tau_2]} = \frac{1}{2}\sum_{\cdots} u^{\tau_1} \diamond u^{\tau_2} = \frac{1}{2}\sum_{i=1}^n K^i \diamond K^{n+1-i},$$

using the unique decomposition of a binary tree τ with $|\tau| = n + 1 \geq 2$ leaves into smaller trees τ_1, τ_2 with i (resp. $n + 1 - i$) leaves for some $i = 1, \ldots, n$. By the Markov property,

---

[11] See also [18] and [13], Chapter 15, for similar trees in the KPZ context.

[12] We use $|\tau|$ to denote the number of leaves, which differs by 1 from the number of inner nodes which is the counting convention used in [19], Equ (2.3).

[13] Cf. Remark 1.1 for related combinatorial comments, including symmetry factors.

$\lambda^{-1} \log \mathbb{E}_t e^{\lambda A_T(\omega)} = h(t, B_t) =: \bar{h}$, where $\overline{(\dots)}$ indicates composition of a function with time-space Brownian motion $\bar{B}_t = (t, B_t)$. Then the $\bar{u}^{\tau_1}$ are semimartingales and by Itô calculus

$$\left(\bar{u}^{\tau_1} \diamond \bar{u}^{\tau_2}\right)_t(T) = \mathbb{E}_t \langle \bar{u}^{\tau_1}, \bar{u}^{\tau_2} \rangle_{t,T} = \mathbb{E}_t \left(\left(\int_t^T \nabla u^{\tau_1} \cdot \nabla u^{\tau_2} \right)(s, B_s) \, ds \right)$$

$$= \bar{u}_t^{[\tau_1, \tau_2]} = (u^{\tau_1} \diamond u^{\tau_2}) \circ \bar{B}_t,$$

which could be expressed as a commutative diagram. (Note that respective diamonds used on the left and right are different, introduced in Definition 1.1 and (4.6), respectively.) It follows from (4.7) that $\bar{K}^n := K^n \circ \bar{B}$ satisfies the same diamond recursion (1.3) as $\mathbb{K}^n$. The—in view of (4.5) still formal—conclusion

$$\log \mathbb{E}_t e^{\lambda A_T(\omega)} = \lambda \int_0^t \xi(s, B_s) \, ds + \lambda u^\circ(t, B) + \sum_{n \geq 2} \lambda^n \bar{K}^n(t, x; T)$$

is then in exact agreement with the $\mathbb{K}$ expansion, since

$$\mathbb{K}_t^1(T) := \mathbb{E}_t(A_T(\omega)) = \mathbb{E}_t \left(h_T(B_T) + \int_0^T \xi(s, B_s) \, ds \right) = u^\circ(t, B_t) + \int_0^t \xi(s, B_s) \, ds$$

and subsequent terms in the recursion are not affected by the final BV term. Theorem 1.2 now settles convergence of (4.5), with the additional advantage of removing the stringent conditions on the data: exponential moments for terminal data $h_T(B_T)$ and integrated forcing $\int \xi_0^T(s, B)s) \, ds)$ are enough. We summarize this discussion as the following.

THEOREM 4.3. *For λ small enough, the perturbative expansion for the KPZ equation* (4.4)

$$(4.8) \qquad \lambda^{-1} \log \mathbb{E}^{t,x} e^{\lambda \{h_T(B_T) + \int_t^T \xi(s, B_s) \, ds\}} = h(t, x) = \sum_\tau \lambda^{|\tau| - 1} u^\tau(t, x)$$

converges. Moreover, terms of same homogeneity have the stochastic representation

$$\sum_{\tau: |\tau| = n} u^\tau(t, x) = \mathbb{E}(\mathbb{K}_t^n(T) | B_t = x),$$

where $\mathbb{K}_t^n(T)$ follows (1.3), $n > 1$, *and started with t-conditional expectation of*

$$\mathbb{K}_T^1(T) = A_T = h_T(B_T) + \int_0^T \xi(s, B_s) \, ds.$$

REMARK 4.1. Theorem 4.3 is really a Markovian perspective on the cumulant recursion and the above argument is readily repeated when Brownian motion (with generator $\Delta/2$) is replaced by a generic diffusion process (resp. its generator), in which case $(\nabla u^{\tau_1} \cdot \nabla u^{\tau_2})/2$ in (4.6) must be replaced by the corresponding *carré du champ* $\Gamma(u^{\tau_1}, u^{\tau_2})$, cf. [26], Prop. VIII.3.3. Sufficient conditions for the recursion (4.6) to be well defined, so that $u^\tau \in C^{1,2}$ for all τ, hence $\bar{u}^\tau$ semimartingales, are a delicate issue. The martingale based diamond expansion bypasses this issue entirely, with $\bar{u}^\tau$ as part of $\mathbb{K}^{|\tau|}$ constructed directly, and so applies immediately when B in (4.8) is replaced by a generic diffusion processes on $\mathbb{R}^d$.

4.6. *Cumulants on Wiener–Itô chaos.* On the classical Wiener space $C([0, T], \mathbb{R})$, with Brownian motion $B(\omega, t) = \omega_t$, consider an arbitrary element in the second Wiener Itô chaos, written in the form

$$A_T := I_2(f) := \int_0^T \int_0^v f(w, v) \, dB_w \, dB_v,$$

with $f = f_A \in L^2$ on the simplex $\Delta_T = \{(s,t) : 0 \le s \le t \le T\}$. Note martingality $A_t := \mathbb{E}_t A_T$ so that $\mathbb{E}_t A_{t,T} = \mathbb{E}_t A_T - A_t = 0$. Then

$$A_{t,T} = \int_t^T \int_0^v f(w,v)\, dB_w\, dB_v = \int_t^T \int_t^v f(w,v)\, dB_w\, dB_v + \int_t^T \int_0^t f(w,v)\, dB_w\, dB_v$$

and

$$\langle A \rangle_{t,T} = \int_t^T \left(\int_0^v f(w,v)\, dB_w \right)^2 dv = \int_t^T \left(\int_0^t f(w,v)\, dB_w + \int_t^v (\cdots) \right)^2 dv,$$

so that

$$(A \diamond A)_t(T) = \mathbb{E}_t \langle A \rangle_{t,T} = \int_t^T \left(\int_0^t f(w,v)\, dB_w \right)^2 dv + 0 + \int_t^T \int_t^v f^2(w,v)\, dw\, dv.$$

We have thus computed $\mathbb{K}_t^2(T) = \frac{1}{2}(A \diamond A)_t(T)$. By polarization, for $A = I_2(f_A)$, $C = I_2(f_C)$,

$$(A \diamond C)_t(T) = \int_t^T \left(\int_0^t f_A(r,v)\, dB_r \right)\left(\int_0^t f_C(r,v)\, dB_r \right) dv$$

$$+ \int_t^T \int_t^v f_A(w,v) f_C(w,v)\, dw\, dv.$$

To go further, we exhibit the martingale part of $A \diamond C$ by writing

$$\int_0^T \left(\int_0^t f_A(r,v)\, dB_r \right)\left(\int_0^t f_C(r,v)\, dB_r \right) dv - \int_0^t (\cdots)\, dv$$

$$+ \int_t^T \int_t^v f_A(w,v) f_C(w,v)\, dw\, dv.$$

From the product rule, with $BV_t = \int_0^t f_A(r,v) f_C(r,v)\, dr$, we have

$$\left(\int_0^t f_A(r,v)\, dB_r \right)\left(\int_0^t f_C(r,v)\, dB_r \right)$$

$$= \int_0^t \int_0^s \left[f_A(r,v) f_C(s,v) + f_C(r,v) f_A(s,v) \right] dB_r\, dB_s + BV_t$$

Letting $\otimes_1$ indicate integration in one (the right-sided) variable and tilde symmetrisation,

$$f_A \tilde{\otimes}_1 f_C := \int_0^T \left(\left[f_A(r,v) f_C(s,v) + f_C(r,v) f_A(s,v) \right] \right) dv$$

so that

$$(A \diamond C)_t(T) = \int_0^t \int_0^s (f_A \tilde{\otimes}_1 f_C)\, dB_r\, dB_s + BV_t(T)$$

with

$$(A \diamond C)_0(T) = BV_0(T) = \int_0^T \int_0^v f_A(r,v) f_C(r,v)\, dr\, dv = \langle f_A, f_C \rangle_{\Delta_T} =: f_A \otimes_2 f_C.$$

In particular, we see that from (1.7) that the third cumulant of $A_T = I_2(f_A)$ is given by

$$\kappa_3(A_T) = \kappa_3(I_2(f_A)) = 3(A \diamond (A \diamond A))_0(T) = \langle f_A, f_{A \diamond A} \rangle = \langle f_A, (f_A \tilde{\otimes}_1 f_A) \rangle.$$

Theorem 1.2 then provides, in the present setting, an alternative to the (Malliavin calculus based) approach of Nourdin–Peccati [32]: by (5.22) in that paper, the nth cumulant of $I_2(f)$ is given by some explicit formula which reduces to (in case $n = 3$) our formula. It is not difficult

to push this "diamond" computation to recover cumulants for general integer n. The diamond approach of course works just as well for higher Wiener–Itô chaos and d-dimensional Wiener space, as was already seen in Section 4.3. Note however that the exponential integrability assumed in part (ii) of Theorem 1.2, valid in the second chaos, does not hold for third and higher chaos. However, any fixed chaos random variable has moments of all orders so that part (i) of this theorem is applicable. Last not least, note that [32] deal with Gaussian fields, whereas we have been dealing with processes.

4.7. Stochastic volatility.

4.7.1. *Joint law of SPX, realized variance and VIX squared.* We return to the financial mathematics context that originally gave rise to diamond expansions result. Our framework permits the valuation and hedging of complex derivatives involving combinations of assets and their quadratic variations. To be specific, let S be a strictly positive continuous martingale. Then $X := \log S$ is a semimartingale, with $e^{X_T} \in L^1$, so that X_T has moments of all orders. If the quadratic variation process $\langle X \rangle$ is absolutely continuous, the *stochastic variance* and *forward variance* are given by

$$v_t := d\langle X \rangle_t / dt, \qquad \xi_t(T) = \mathbb{E}_t[v_T].$$

Upon integration in time, these quantities—realized and expected quadratic variation at a future time T—constitute the payoff of a variance swap and VIX^2, respectively. (This application entails the interpretation of e^X as the risk neutral price process of the SPX index on which the VIX index is built.) We now illustrate the use of Theorem 1.1 to determine the joint law of (log)-price, realized and expected quadratic variation at a future time T, the precise setting for consistent pricing of options on SPX, realized variance and VIX squared, with time-T payoff

$$\zeta_T(T) = \int_T^{T+\Delta} \xi_T(u)\,du = \mathbb{E}_T \int_T^{T+\Delta} v_u\,du = \mathbb{E}_T \langle X \rangle_{T,T+\Delta}.$$

THEOREM 4.4. *Assume X_T has exponential moments. Then for $a, b, c \in \mathbb{R}$ sufficiently small,*

$$(4.9) \quad \mathbb{E}_t\big[e^{aX_T + b\langle X \rangle_T + c\zeta_T(T)}\big] = \exp\bigg\{aX_t + b\langle X \rangle_t + c\zeta_t(T) + \sum_{k=2}^{\infty} \mathbb{G}_t^k(T; a, b, c)\bigg\},$$

where the $\mathbb{G}^k$'s are given recursively by (1.2), starting with

$$\mathbb{G}^2 = \Big(\frac{1}{2}a(a-1) + b\Big) X \diamond X + acX \diamond \zeta + \frac{1}{2}c^2\zeta \diamond \zeta.$$

PROOF. This is a direct consequence of the multivariate $\mathbb{G}$-expansion of Theorem 1.1, employed with time-T terminal data, re-expressed in terms of the martingale $Y = X + \frac{1}{2}\langle X \rangle$,

$$aX_T + b\langle X \rangle_T + c\zeta_T(T) = aY_T + \Big(b - \frac{1}{2}a\Big)\langle Y \rangle_T + c\zeta_T(T).$$

We note for later use that the convergent $\mathbb{G}$-sum is exactly equal to Λ, which satisfies the "abstract Riccati" equation (2.7),

$$\Lambda_t^T = \mathbb{E}_t Z_{t,T} + \frac{1}{2}\big((Z + \Lambda^T) \diamond (Z + \Lambda^T)\big)_t(T).$$

Then, since Y and ζ are martingales, and $\langle Y \rangle$ is BV hence annihilated by $\diamond$,

$$\Lambda = \left(b - \frac{1}{2}a\right)Y \diamond Y + \frac{1}{2}((aY + c\zeta + \Lambda) \diamond (aY + c\zeta + \Lambda))$$

$$= \left(\frac{1}{2}a(a-1) + b\right)Y \diamond Y + acY \diamond \zeta + aY \diamond \Lambda + c\zeta \diamond \Lambda + \frac{1}{2}c^2 \zeta \diamond \zeta + \frac{1}{2}\Lambda \diamond \Lambda,$$

which in terms of trees with $Y = \circ$ and $\zeta = \bullet$ gives

$$(4.10) \quad \Lambda = \left(\frac{1}{2}a(a-1) + b\right)\overset{\circ}{\underset{}{\vee}}{}^{\circ} + ac\overset{}{\underset{}{\vee}}{}^{\bullet} + a\circ \diamond \Lambda + c\bullet \diamond \Lambda + \frac{1}{2}c^2\overset{}{\underset{}{\vee}}{}^{\bullet} + \frac{1}{2}\Lambda \diamond \Lambda. \qquad \square$$

REMARK 4.2. Martingality of $S = e^X$ is seen in (4.9) upon setting $(a, b, c) = (1, 0, 0)$. It is a feature of the $\mathbb{G}$ expansion that each term $\mathbb{G}_t^k(T; 1, 0, 0)$ vanishes so that the martingality constraint is preserved at arbitrary truncation of the $\mathbb{G}$ expansion, reminiscent of (Lie algebra preserving) Magnus expansions for differential equations on Lie groups [21]. This is not the case for the multivariate $\mathbb{K}$ expansion, cf. Remark 1.2 for a related discussion.

REMARK 4.3. We insist that (4.9) is a model free result, with $\mathbb{G}$-expansion given naturally by diamond trees with two types of leaves corresponding to X and ζ.

4.7.2. *Forward and affine forward variance models.* After Black–Scholes (constant volatility), classical stochastic volatility models consider $v = v(t, \omega)$ as a stochastic process in its own right. "Third generation models" where one specifies directly forward variances—viewed as a family of martingales indexed by their individual time horizon—are nowadays ubiquitous in equity financial modeling. In full generality this reads

$$(4.11) \qquad\qquad d_t\xi_t(u) = d_t\mathbb{E}_t[v_u] =: \sigma_t(u)\,dW_t^u, \quad t \le u,$$

where $v_t \equiv \xi_t(t)$ and $dS_t/S_t = \sqrt{v_t}\,dZ_t$, where the correlation (covariation) structure of the Brownian family $\{Z, W^T : T \ge 0\}$ also needs to be specified. We can then immediately write (diamonds with ζ amount to average diamonds with $\xi(T')$ over $T' \in [T, T+\Delta]$)

$$\overset{\circ}{\underset{}{\vee}}{}^{\circ} := (X \diamond X)_t(T) = \mathbb{E}_t\left[\int_t^T d\langle X\rangle_s\right] = \int_t^T \xi_t(s)\,ds,$$

$$\overset{}{\underset{}{\vee}}{}^{\bullet} := (X \diamond \xi(T'))_t(T) = \mathbb{E}_t\left[\int_t^T d\langle X, \xi(T')\rangle_s\right] = \mathbb{E}_t\int_t^T \sqrt{v_s}\sigma_s(T')\,d\langle Z, W^{T'}\rangle_s,$$

$$\overset{\bullet}{\underset{}{\vee}}{}^{\bullet} := (\xi(T') \diamond \xi(T'))_t(T) = \mathbb{E}_t\left[\int_t^T d\langle \xi(T'), \xi(T')\rangle_s\right] = \int_t^T \mathbb{E}_t[\sigma_s(T')^2]\,ds.$$

At this stage, more structure is required for computations. A particularly simple choice is the affine specification $\sigma_t(u) = \kappa(u - t)\sqrt{v_t}$ of [16]:

$$\frac{dS_t}{S_t} = \sqrt{v_t}\,dZ_t,$$

$$(4.12) \qquad\qquad d_t\xi_t(u) = \kappa(u - t)\sqrt{v_t}\,dW_t, \quad t \le u,$$

where κ is some L^2-kernel and the Brownian drivers satisfy $d\langle W, Z\rangle_t/dt = \rho$. Note that

$$\xi_t(u) = \xi_0(u) + \int_0^t \kappa(t - s)\sqrt{v_s}\,dW_s, \qquad v_t \equiv \xi_t(t) = \xi_0(t) + \int_0^t \kappa(t - s)\sqrt{v_s}\,dW_s$$

so that stochastic variance solves a stochastic Volterra equation. Special cases are the Heston and rough Heston models with exponential and power-law kernels, respectively. We also note that

$$(4.13) \qquad d_t \zeta_t(T) = \left(\int_T^{T+\Delta} \kappa(u-t) \, du \right) \sqrt{v_t} \, dW_t =: \bar{\kappa}(T-t) \sqrt{v_t} \, dW_t$$

has the same form as (4.12); in computations, ζ can then effectively be replaced by ξ.

LEMMA 4.3. *In the affine forward variance model* (4.12) *all diamond trees* (*with leaves of two types* $X = \circ$ *and* $\zeta = \bullet$, *respectively*), *and hence all forests terms* $\mathbb{G}_t^k$ *in* (4.9) *are of the form*

$$(4.14) \qquad \int_t^T \xi_t(u) h(T-u) \, du$$

for some integrable function h.

PROOF. As above, $\multimap^\circ = \int_t^T \xi_t(s) \, ds$, but now with (4.12), also noting (4.13),

$$(4.15) \qquad \multimap^\bullet = \rho \int_t^T \xi_t(u) \bar{\kappa}(T-u) \, du, \qquad \bullet\!\!\multimap^\bullet = \int_t^T \xi_t(u) \bar{\kappa}(T-u)^2 \, du.$$

We thus see that the claim holds for all diamond trees with two leaves and proceed by induction. Consider two trees

$$\mathbb{T}_t^i = \int_t^T \xi_t(u) h^i(T-u) \, du, \quad i = 1, 2$$

of the supposed form. Then

$$(\mathbb{T}^1 \diamond \mathbb{T}^2)_t(T) = \mathbb{E}_t \left[\int_t^T d\langle \mathbb{T}^1, \mathbb{T}^2 \rangle_u \right]$$

$$= \mathbb{E}_t \left[\int_t^T \int_u^T \int_u^T h^1(T-s) h^2(T-r) \, ds \, dr \, d\langle \xi(s), \xi(r) \rangle_u \right]$$

$$= \mathbb{E}_t \left[\int_t^T v_u \kappa(s-u) \kappa(r-u) \, du \int_u^T h^1(T-s) \, ds \int_u^T h^2(T-s) \, dr \right]$$

$$= \int_t^T \xi_t(u) h^{12}(T-u) \, du,$$

and the induction step is completed upon setting

$$h^{12}(T-u) = \int_u^T h^1(T-s) \kappa(T-s) \, ds \int_u^T h^2(T-r) \kappa(T-r) \, dr. \qquad \square$$

REMARK 4.4. The statement and proof of Lemma 4.3 may be extended to the non time-homogeneous setting $d_t \xi_t(u) = \kappa(u, t) \sqrt{v_t} \, dW_t$ without much extra effort.

EXAMPLE 4.4 (Classical Heston). In this case,

$$d\xi_t(u) = v e^{-\lambda(u-t)} \sqrt{v_t} \, dW_t.$$

Then, for example,

$$\multimap^{\circ\circ} = (X \diamond (X \diamond X))_t(T) = \frac{\rho v}{\lambda} \int_t^T \xi_t(u) [1 - e^{-\lambda(T-u)}] \, du.$$

EXAMPLE 4.5 (Rough Heston). In this case, with $\alpha = H + 1/2 \in (1/2, 1)$,

$$d\xi_t(u) = \frac{\nu}{\Gamma(\alpha)}(u-t)^{\alpha-1}\sqrt{v_t}\, dW_t.$$

Then, for example,

$$\overset{\diamond}{\vee} = (X \diamond X)_t(T) = \int_t^T \xi_t(u)\, du,$$

$$\overset{\diamond}{\vee\!\vee} = ((X \diamond X) \diamond (X \diamond X))_t(T) = \frac{\nu^2}{\Gamma(\alpha)^2}\int_t^T \xi_t(u)\, du \left(\int_u^T (s-u)^{\alpha-1}\, ds\right)^2$$

$$= \frac{\nu^2}{\Gamma(1+\alpha)^2}\int_t^T \xi_t(u)(T-u)^{2\alpha}\, du.$$

For a bounded forward variance curve ξ one then sees that diamond trees with k leaves are of order $(T-t)^{1+(k-2)\alpha}$. In this case, the $\mathbb{F}$-expansion (forest reordering according to number of leaves) has the interpretation of a short-time expansion, the concrete powers of which depend on the roughness parameter $\alpha = H + 1/2 \in (1/2, 1)$. The resulting diamond expansions (which can be obtained by alternative methods in the rough Heston case) were seen to be numerically efficient in [6, 17].

At this stage, it is tempting to combine Lemma 4.3 with Theorem 4.4 to compute the triple-joint mgf of X_T, $\langle X \rangle_{t,T}$, and $\zeta_T(T)$ by summing the full $\mathbb{G}$-expansion for an affine forward variance model. We then see that the mgf is necessarily of the convolutional form

$$\log \mathbb{E}_t[e^{aX_T + b\langle X \rangle_T + c\zeta_T(T)}] = aX_t + b\langle X \rangle_t + c\zeta_t(T) + \int_t^T \xi_t(u)g(T-u; a, b, c, \Delta)\, du,$$

which amounts to an infinite-dimensional version of the classical affine ansatz. Inserting $\Lambda_t(T) = \int_t^T \xi_t(u)g(T-u; a, b, c, \Delta)\, du$ directly into the "abstract Riccati" equation (4.10), we readily obtain that the triple-joint MGF satisfies a convolution Riccati equation of the type considered in [1, 16]. We summarize this in the following theorem.

THEOREM 4.5. *Let*

$$dX_t = -\frac{1}{2}v_t\, dt + \sqrt{v_t}\, dZ_t,$$

$$d\xi_t(T) = \kappa(T-t)\sqrt{v_t}\, dW_t,$$

with $d\langle W, Z \rangle_t = \rho dt$ *and let* $\langle X \rangle_{t,T} = \langle X \rangle_T - \langle X \rangle_t$. *Further let* $\tau = T - t$, $\bar{\kappa}(\tau) = \int_\tau^{\tau+\Delta} \kappa(u)\, du$, *and define the convolution integral*

$$(\kappa \star g)(\tau) = \int_0^\tau \kappa(\tau - s)g(s)\, ds.$$

Then

$$\mathbb{E}_t[e^{aX_T + b\langle X \rangle_{t,T} + c\zeta_T(T)}] = \exp\{aX_t + c\zeta_t(T) + (\xi \star g)(T-t; a, b, c, \Delta)\},$$

where $g(\tau; a, b, c, \Delta)$ *satisfies the convolution Riccati integral equation*

$$(4.16) \quad \begin{aligned} g(\tau; a, b, c, \Delta) &= b - \frac{1}{2}a + \frac{1}{2}(1-\rho^2)a^2 \\ &\quad + \frac{1}{2}[\rho a + c\bar{\kappa}(\tau) + (\kappa \star g)(\tau; a, b, c, \Delta)]^2, \end{aligned}$$

with the boundary condition $g(0; a, b, c, \Delta) = b + \frac{1}{2}a(a-1) + \rho ac\bar{\kappa}(0) + \frac{1}{2}c^2\bar{\kappa}(0)^2$.

PROOF. From (4.13), $d\zeta_t(T) = \sqrt{v_t}\,dW_t\bar{\kappa}(T-t)$. As before, let $\Lambda_t = \int_t^T \xi_t(u)g(T-u; a, b, c, \Delta)\,du$. Then dropping the arguments a, b, c, Δ for ease of notation,

$$d\Lambda_t = -\xi_t(t)g(T-t)\,dt + \int_t^T d\xi_t(s)g(T-s)\,ds$$

$$= -v_t g(T-t)\,dt + \sqrt{v_t}\,dW_t \int_t^T \kappa(s-t)g(T-s)\,ds$$

$$= -v_t g(T-t)\,dt + \sqrt{v_t}\,dW_t(\kappa \star g)(T-t).$$

We compute

$$d\langle X \rangle_t = v_t\,dt,$$

$$d\langle X, \zeta \rangle_t = \rho v_t \bar{\kappa}(T-t)\,dt,$$

$$d\langle \zeta \rangle_t = v_t \bar{\kappa}(T-t)^2\,dt,$$

$$d\langle X, \Lambda \rangle_t = \rho v_t(\kappa \star g)(T-t)\,dt,$$

$$d\langle \Lambda \rangle_t = v_t\big[(\kappa \star g)(T-t)\big]^2\,dt,$$

$$d\langle \zeta, \Lambda \rangle_t = v_t \bar{\kappa}(T-t)(\kappa \star g)(T-t)\,dt.$$

Integrating these terms from t to T, followed by taking a time-t conditional expectation allows us to compute all diamond products in the "abstract Riccati" equation (4.10)

$$\Lambda = \left(\frac{1}{2}a(a-1) + b\right)^{\heartsuit\circ} + ac^{\heartsuit\bullet} + a\circ\diamond\Lambda + \frac{1}{2}[c\bullet + \Lambda]^{\diamond 2}$$

to yield

$$g(\tau) = \left(\frac{1}{2}a(a-1) + b\right) + \rho ac\bar{\kappa}(\tau) + a\rho(\kappa \star g)(\tau) + \frac{1}{2}[c\bar{\kappa}(\tau) + (\kappa \star g)(\tau)]^2,$$

which upon rearrangement gives (4.16). Finally, $(\kappa \star g)(0) = 0$ gives the boundary condition

$$g(0) = b + \frac{1}{2}a(a-1) + \rho ac\bar{\kappa}(0) + \frac{1}{2}c^2\bar{\kappa}(0)^2. \qquad \square$$

4.8. *Multifactor Heston.* In this section, we demonstrate the general applicability of our approach by computing diamond trees in the multifactor volatility Heston model of [7].

4.8.1. *Wishart variance dynamics.* Let W_t and Z_t be $d \times d$ Brownian motions and $\mathcal{S}_+^d$ be the set of real-valued $d \times d$ positive semidefinite matrices. The model involves the $\mathcal{S}_+^d$-valued *Wishart process* (Σ_t) and reads

(4.17)
$$\frac{dS_t}{S_t} = \text{Tr}[\sqrt{\Sigma_t}\,dZ_t],$$

$$d\Sigma_t = \left(\Omega\Omega^T + M\Sigma_t + \Sigma_t M^T\right)dt + \sqrt{\Sigma_t}\,dW_t Q + Q^T(dW_t)^T\sqrt{\Sigma_t},$$

where Ω, M and Q are in $\mathbb{R}^{d \times d}$, with Ω and M invertible. From Itô's isometry, instantaneous variance is given by the $v_t = \text{Tr}[\Sigma_t]$ and in terms of the forward variance matrix $\Xi_t(u) := \mathbb{E}_t[\Sigma_u] \in \mathcal{S}_+^d$ we have forward variance $\xi_t(u) := \mathbb{E}_t[v_u] = \text{Tr}[\Xi_t(u)]$, for $t < u$.

Since Theorem 4.4 is in terms of diamond trees, we need only to show how to compute diamond products. First, we recast the model in forward variance form. Taking conditional expectations of (4.17), the forward variance matrix satisfies

(4.18)
$$\frac{d\Xi_t(u)}{du} = \Omega\Omega^T + M\Xi_t(u) + \Xi_t(u)M^T,$$

which in particular gives the evolution of forward variance as

$$(4.19) \qquad \frac{d}{du}\xi_t(u) := \frac{d}{du}\mathrm{Tr}[\Xi_t(u)] = \mathrm{Tr}[\Omega\Omega^T + M\,\Xi_t(u) + \Xi_t(u)M^T].$$

LEMMA 4.4. *The ODE* (4.19) *has the solution*

$$(4.20) \qquad \xi_t(u) = \mathrm{Tr}[\Xi_t(u)] = \mathrm{Tr}\left[-\frac{1}{2}M^{-1}\Omega\Omega^T + e^{(u-t)2M}\left(\Sigma_t + \frac{1}{2}M^{-1}\Omega\Omega^T\right)\right].$$

PROOF.   Using standard properties of the trace, the ODE (4.19) can be rewritten as

$$\frac{d}{du}\mathrm{Tr}[\Xi_t(u)] = \mathrm{Tr}[\Omega\Omega^T + 2M\,\Xi_t(u)],$$

which has the stated solution.   □

For diamond computations, we need to make explicit the martingale structure of $\xi_t(u)$ in t $(\leq u)$. Since we know a priori that there are no bounded variation terms, we read off directly from (4.20) and (4.17) that

$$(4.21) \qquad d_t\xi_t(u) = \mathrm{Tr}[d\,\Xi_t(u)] = \mathrm{Tr}[e^{(u-t)2M}(\sqrt{\Sigma_t}\,dW_t\,Q + Q^T(dW_t)^T\sqrt{\Sigma_t})].$$

4.8.2. *Diamonds from the forward variance formulation.*   Now we have the model in forward variance form, we may compute diamond trees. First, we compute

$$\text{♥}\!\!\text{♥} = (X \diamond X)_t(T) = \int_t^T \mathbb{E}_t[\mathrm{Tr}[\Sigma_u]]\,du = \int_t^T \mathrm{Tr}[\Xi_t(u)]\,du = \int_t^T \xi_t(u)\,du.$$

Thus

$$d(X \diamond X)_t(T) = \int_t^T d\xi_t(u)\,du + BV = \int_t^T \mathrm{Tr}[d\,\Xi_t(u)]\,du + BV,$$

where, defining $K(\tau) = e^{\tau 2M}$,

$$\mathrm{Tr}[d\,\Xi_t(u)] = \mathrm{Tr}[K(u-t)(\sqrt{\Sigma_t}\,dW_t\,Q + Q^T(dW_t)^T\sqrt{\Sigma_t})].$$

In [7], the correlation structure of the model is simplified by specifying

$$Z_t = W_t R^T + B_t\sqrt{\mathbb{1} - RR^T},$$

where $\mathbb{1}$ is the identity matrix, B_t is a matrix Brownian motion independent of W, and R is a matrix of correlations. In order to make progress, we need another lemma.

LEMMA 4.5.   *Let* W, Z *be as above. Let* $A = A(t, \omega)$, $B = B(t, \omega)$ *be locally square-integrable adapted process with values in* $\mathbb{R}^{d \times d}$. *Then*

$$d_t\left\langle \int_0^{\cdot} \mathrm{Tr}[A\,dW_t], \int_0^{\cdot} \mathrm{Tr}[B\,dZ_t]\right\rangle_t = \mathrm{Tr}[AB^T R]\,dt.$$

PROOF.   Standard Itô calculus. To check the algebra, in abusive (but popular) notation,

$$dW_t{}^{ji}dZ_t{}^{m\ell} = dW_t{}^{ji}dW_t{}^{mn}R_{n\ell}^T = \delta^{jm}\delta^{in}R_{n\ell}^T = \delta^{jm}R_{i\ell}^T\,dt,$$

so that, always with Einstein summation convention,

$$A_{ij}dW_t{}^{ji}B_{\ell m}dZ_t{}^{m\ell} = A_{ij}B_{\ell m}\delta^{jm}R_{i\ell}^T\,dt = A_{ij}(B^T)_{j\ell}R_{\ell i}\,dt.   \qquad □$$

PROPOSITION 4.1. *In the context of* (4.17), *with matrix kernel* $K(\tau) = e^{\tau 2M}$, *for* $t \le u$,

$$d_t \left\langle \int_0^\cdot \mathrm{Tr}[K(u-t)\sqrt{\Sigma_t}\,dW_t Q], \int_0^\cdot \mathrm{Tr}[\sqrt{\Sigma_t}\,dZ_t] \right\rangle_t = \mathrm{Tr}[RQK(u-t)\Sigma_t]\,dt,$$

$$d_t \left\langle \int_0^\cdot \mathrm{Tr}[K(u-t)Q^T(dW_t)^T, \sqrt{\Sigma_t}], \int_0^\cdot \mathrm{Tr}[\sqrt{\Sigma_t}\,dZ_t] \right\rangle_t = \mathrm{Tr}[RQK(u-t)^T\Sigma_t]\,dt.$$

PROOF. For the first relation, using the cyclical property of the trace, $\mathrm{Tr}[K(u-t) \times \sqrt{\Sigma_t}\,dW_t Q] = \mathrm{Tr}[QK(u-t)\sqrt{\Sigma_t}\,dW_t]$. Then apply Lemma 4.5 with $A = QK(u-t)\sqrt{\Sigma_t}$, $B = \sqrt{\Sigma_t}$. By definition, $\Sigma_t = \sqrt{\Sigma_t}(\sqrt{\Sigma_t})^T$. For the second relation, use that $\mathrm{Tr}[K(u-t)Q^T(dW_t)^T\sqrt{\Sigma_t}] = \mathrm{Tr}[(\sqrt{\Sigma_t})^T\,dW_t QK(u-t)^T]$ and after using the cyclical property of the trace again, apply Lemma 4.5 with $A = QK(u-t)^T(\sqrt{\Sigma_t})^T$, $B = \sqrt{\Sigma_t}$. □

THEOREM 4.6. *We have the following explicit form of* $\overset{\circ\!\!\circ}{\diamondsuit} = (X \diamond X) \diamond X$. *More precisely,*

$$\begin{aligned}
(4.22) \qquad ((X \diamond X) \diamond X)_t(T) &= \mathbb{E}_t\left[\int_t^T d_t\langle(X \diamond X)(T), X\rangle_s\right] \\
&= \int_t^T \mathrm{Tr}[H(T-s)\Xi_t(s)]\,ds,
\end{aligned}$$

with

$$H(\tau) = \int_s^T RQ(K(u-s) + K(u-s)^T)\,du.$$

PROOF. The above proposition shows that

$$d_t\left\langle \int_0^\cdot \mathrm{Tr}[d\,\Xi_t(u)], \int_0^\cdot \mathrm{Tr}[\sqrt{\Sigma_t}\,dZ_t]\right\rangle = \mathrm{Tr}[RQ(K(u-t) + K(u-t)^T)\Sigma_t]\,dt.$$

We can then compute

$$\begin{aligned}
\overset{\circ\!\!\circ}{\diamondsuit} &= \mathbb{E}_t\left[\int_t^T d_t\langle(X \diamond X)(T), X\rangle_s\right] \\
&= \mathbb{E}_t\left[\int_t^T \left(\int_s^T \mathrm{Tr}[\Xi_s(u)]\,du\right)\mathrm{Tr}[\sqrt{\Sigma_s}\,dZ_s]\right] \\
&= \mathbb{E}_t\left[\int_t^T \int_s^T \mathrm{Tr}[RQ(K(u-s) + K(u-s)^T)\Sigma_s]\,du\,ds\right] \\
&= \int_t^T \int_s^T \mathrm{Tr}[RQ(K(u-s) + K(u-s)^T)\Xi_t(s)]\,du\,ds. \qquad \square
\end{aligned}$$

In view of (4.22), we see that this argument iterates so that all diamond trees will be convolutions of the forward variance matrix $\Xi_t(s)$ and some matrix function of time, in full analogy to the classical Heston model, Example 4.4. We further observe that, once in forward variance form, the exponential form of the kernel, $K(\tau) = e^{\tau 2M}$, plays no role. Explicit diamond expansions are then also possible in rough extensions of multifactor Heston for which K takes power-law form, again in in full analogy with the rough Heston model as given in Example 4.5. (A systematic study of rough Wishart and multifactor Heston models is left to future research.)

REFERENCES

[1] ABI JABER, E., LARSSON, M. and PULIDO, S. (2019). Affine Volterra processes. *Ann. Appl. Probab.* **29** 3155–3200. MR4019885 https://doi.org/10.1214/19-AAP1477

[2] ALÒS, E., GATHERAL, J. and RADOIČIĆ, R. (2020). Exponentiation of conditional expectations under stochastic volatility. *Quant. Finance* **20** 13–27. MR4040259 https://doi.org/10.1080/14697688.2019.1642506

[3] BOUÉ, M. and DUPUIS, P. (1998). A variational representation for certain functionals of Brownian motion. *Ann. Probab.* **26** 1641–1659. MR1675051 https://doi.org/10.1214/aop/1022855876

[4] BRIAND, P. and HU, Y. (2008). Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields* **141** 543–567. MR2391164 https://doi.org/10.1007/s00440-007-0093-y

[5] BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42. MR0365692 https://doi.org/10.1214/aop/1176997023

[6] CALLEGARO, G., GRASSELLI, M. and PAGÈS, G. (2021). Fast hybrid schemes for fractional Riccati equations (rough is not so tough). *Math. Oper. Res.* **46** 221–254. MR4224426 https://doi.org/10.1287/moor.2020.1054

[7] DA FONSECA, J., GRASSELLI, M. and TEBALDI, C. (2008). A multifactor volatility Heston model. *Quant. Finance* **8** 591–604. MR2457710 https://doi.org/10.1080/14697680701668418

[8] DUPIRE, B. (2019). Functional Itô calculus. *Quant. Finance* **19** 721–729. MR3939653 https://doi.org/10.1080/14697688.2019.1575974

[9] EBRAHIMI-FARD, K., PATRAS, F., TAPIA, N. and ZAMBOTTI, L. (2020). Hopf-algebraic deformations of products and Wick polynomials. *Int. Math. Res. Not. IMRN* **24** 10064–10099. MR4190397 https://doi.org/10.1093/imrn/rny269

[10] EL EUCH, O. and ROSENBAUM, M. (2019). The characteristic function of rough Heston models. *Math. Finance* **29** 3–38. MR3905737 https://doi.org/10.1111/mafi.12173

[11] FILIPOVIĆ, D. and MAYERHOFER, E. (2009). Affine diffusion processes: Theory and applications. In *Advanced Financial Modelling* (H. Albrecher, W. J. Runggaldier and W. Schachermayer, eds.). *Radon Ser. Comput. Appl. Math.* **8** 125–164. de Gruyter, Berlin. MR2648460 https://doi.org/10.1515/9783110213140.125

[12] FRIZ, P. K., HAGER, P. and TAPIA, N. (2021). Unified signature cumulants and generalized Magnus expansions. Available at arXiv:2102.03345.

[13] FRIZ, P. K. and HAIRER, M. (2020). *A Course on Rough Paths. With an Introduction to Regularity Structures*, 2nd ed. Springer, Cham. MR4174393 https://doi.org/10.1007/978-3-030-41556-3

[14] FUKASAWA, M. and MATSUSHITA, K. (2021). Realized cumulants for martingales. *Electron. Commun. Probab.* **26** Paper No. 12. MR4236682

[15] GAINES, J. G. and LYONS, T. J. (1997). Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.* **57** 1455–1484. MR1470933 https://doi.org/10.1137/S0036139995286515

[16] GATHERAL, J. and KELLER-RESSEL, M. (2019). Affine forward variance models. *Finance Stoch.* **23** 501–533. MR3968277 https://doi.org/10.1007/s00780-019-00392-5

[17] GATHERAL, J. and RADOIČIĆ, R. (2019). Rational approximation of the rough Heston solution. *Int. J. Theor. Appl. Finance* **22** 1950010. MR3951951 https://doi.org/10.1142/S0219024919500109

[18] GUBINELLI, M. and PERKOWSKI, N. (2017). KPZ reloaded. *Comm. Math. Phys.* **349** 165–269. MR3592748 https://doi.org/10.1007/s00220-016-2788-3

[19] HAIRER, M. (2013). Solving the KPZ equation. *Ann. of Math.* (2) **178** 559–664. MR3071506 https://doi.org/10.4007/annals.2013.178.2.4

[20] IKEDA, N. and WATANABE, S. (2014). *Stochastic Differential Equations and Diffusion Processes*. Elsevier.

[21] ISERLES, A. and NØRSETT, S. P. (1999). On the solution of linear differential equations in Lie groups. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **357** 983–1019. MR1694700 https://doi.org/10.1098/rsta.1999.0362

[22] JANSON, S. (1997). *Gaussian Hilbert Spaces. Cambridge Tracts in Mathematics* **129**. Cambridge Univ. Press, Cambridge. MR1474726 https://doi.org/10.1017/CBO9780511526169

[23] KLOEDEN, P. E. and PLATEN, E. (1992). *Numerical Solution of Stochastic Differential Equations. Applications of Mathematics* (*New York*) **23**. Springer, Berlin. MR1214374 https://doi.org/10.1007/978-3-662-12616-5

[24] KOBYLANSKI, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* **28** 558–602. MR1782267 https://doi.org/10.1214/aop/1019160253

[25] LACOIN, H., RHODES, R. and VARGAS, V. (2019). A probabilistic approach of ultraviolet renormalisation in the boundary Sine–Gordon model. Preprint. Available at arXiv:1903.01394.

[26] LE GALL, J.-F. (2013). *Mouvement Brownien, Martingales et Calcul Stochastique. Mathématiques & Applications* (*Berlin*) [*Mathematics & Applications*] **71**. Springer, Heidelberg. MR3184878 https://doi.org/10.1007/978-3-642-31898-6

[27] LEVIN, D. and WILDON, M. (2008). A combinatorial method for calculating the moments of Lévy area. *Trans. Amer. Math. Soc.* **360** 6695–6709. MR2434307 https://doi.org/10.1090/S0002-9947-08-04526-1

[28] LUKACS, E. (1970). *Characteristic Functions*. Hafner Publishing Co., New York. MR0346874

[29] LYONS, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** 215–310. MR1654527 https://doi.org/10.4171/RMI/240

[30] MALLIAVIN, P. (1997). *Stochastic Analysis. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **313**. Springer, Berlin. MR1450093 https://doi.org/10.1007/978-3-642-15074-6

[31] MORAN, P. A. P. (1984). *An Introduction to Probability Theory. Oxford Science Publications*. The Clarendon Press, Oxford University Press, New York. MR0753828

[32] NOURDIN, I. and PECCATI, G. (2010). Cumulants on the Wiener space. *J. Funct. Anal.* **258** 3775–3791. MR2606872 https://doi.org/10.1016/j.jfa.2009.10.024