

Realized cumulants for martingales

Masaaki Fukasawa* Kazuki Matsushita†

Abstract

Generalizing the realized variance, the realized skewness (Neuberger, 2012) and the realized kurtosis (Bae and Lee, 2020), we construct realized cumulants with the so-called aggregation property. They are unbiased statistics of the cumulants of a martingale marginal based on sub-period increments of the martingale and its lower-order conditional cumulant processes. Our key finding is a relation between the aggregation property and the complete Bell polynomials. For an application we give an alternative proof and an extension of a cumulant recursion formula recently obtained by Lacoin et al. (2019) and Friz et al. (2020).

Keywords: cumulant; martingale; aggregation property; high-frequency data.

MSC2020 subject classifications: 60G44.

Submitted to ECP on October 16, 2020, final version accepted on February 17, 2021.

1 Introduction

A square-integrable martingale $M = \{M_t\}_{t \in [0, T]}$ satisfies

$$E[(M_T - M_0)^n] = E \left[\sum_{j=1}^N (M_{t_j} - M_{t_{j-1}})^n \right] \quad (1.1)$$

for $n = 2$ with an arbitrary time partition $0 = t_0 < \dots < t_N = T$. On one hand this is a crucial property to develop a theory of martingale (see e.g., [4]) and on the other hand, after the seminal work [2], this has played an important role in financial econometrics, where the sum of the squared increments is called the realized variance and has been used as an accurate measure of the volatility of an asset price modelled by M . Unfortunately (1.1) does not hold for $n \geq 3$, which causes a difficulty in estimating higher order moments based on high frequency data; see [10] for the details.

For a Borel function g on \mathbb{R}^d and an \mathbb{R}^d -valued adapted process \mathbb{X} on a filtered probability space $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t \in [0, T]})$, the aggregation property introduced by Neuberger [10] refers to that

$$E[g(\mathbb{X}_u - \mathbb{X}_s) | \mathcal{F}_s] = E[g(\mathbb{X}_u - \mathbb{X}_t) | \mathcal{F}_s] + E[g(\mathbb{X}_t - \mathbb{X}_s) | \mathcal{F}_s] \quad (1.2)$$

for any $s \leq t \leq u \leq T$. The identity (1.1) with $n = 2$ can be seen as a consequence of the aggregation property of $g(x) = x^2$ and $\mathbb{X} = M$. An interesting finding by Neuberger [10]

*Graduate School of Engineering Science, Osaka University. E-mail: fukasawa@sigmath.es.osaka-u.ac.jp

†Graduate School of Engineering Science, Osaka University.

E-mail: matsushita@sigmath.es.osaka-u.ac.jp

is that the aggregation property is met by $g(x, y) = x^3 + 3xy$ and $\mathbb{X} = (M, M^{(2)})$ for any L^3 -martingale M , where $M_t^{(n)} = E[(M_T - M_t)^n | \mathcal{F}_t]$. Observing that this implies

$$E[(M_T - M_0)^3] = E[g(\mathbb{X}_T - \mathbb{X}_0)] = E \left[\sum_{j=1}^N g(\mathbb{X}_{t_j} - \mathbb{X}_{t_{j-1}}) \right],$$

Neuberger [10] named the high-frequency statistic

$$\sum_{j=1}^N g(\mathbb{X}_{t_j} - \mathbb{X}_{t_{j-1}}) = \sum_{j=1}^N \left\{ (M_{t_j} - M_{t_{j-1}})^3 + 3(M_{t_j} - M_{t_{j-1}})(M_{t_j}^{(2)} - M_{t_{j-1}}^{(2)}) \right\}$$

the realized skewness, and conducted an econometric analysis with it. Bae and Lee [3] further extended the idea to construct the realized kurtosis, by finding that the aggregation property is met by $g(x, y, z) = x^4 + 6x^2y + 3y^2 + 4xz$ and $\mathbb{X} = (M, M^{(2)}, M^{(3)})$. Neuberger [10] and Bae and Lee [3]’s approach to find those polynomials is rather brute force. Indeed, they showed also that there is no other analytic function (up to linear combinations) g satisfying the aggregation property with $\mathbb{X} = (M, M^{(2)}, M^{(3)})$. An extension to the higher order cumulants (or moments) has been left open.

The aim of this paper is to unveil this mysterious property. Our finding is that

$$g_n(x_1, \dots, x_n) = B_{n+1}(x_1, \dots, x_n, 0)$$

satisfies the aggregation property with

$$\mathbb{X}^{(n)} = (X^{(1)}, \dots, X^{(n)})$$

for any $n \in \mathbb{N}$, where B_{n+1} is the $(n + 1)$ -th complete Bell polynomial and $X^{(i)}$ is the i th conditional cumulant process of an L^{n+1} integrable random variable X (see Section 4 for the details). When $X = M_T$, then $\mathbb{X}^{(3)} = (M, M^{(2)}, M^{(3)})$ and we have

$$X_t^{(n+1)} = E \left[\sum_{j=1}^N g_n(\mathbb{X}_{t_j}^{(n)} - \mathbb{X}_{t_{j-1}}^{(n)}) \middle| \mathcal{F}_t \right] \tag{1.3}$$

for any partition $t = t_0 < \dots < t_N = T$ of $[t, T]$. The cases $n = 1, 2$ and 3 correspond to the realized variance, Neuberger’s realized skewness and Bae and Lee’s realized kurtosis respectively.

Taking high-frequency limit in (1.3), we have

$$X_t^{(n+1)} = E \left[\sum_{s \in (t, T]} g_n(\Delta \mathbb{X}_s^{(n)}) + \frac{1}{2} \sum_{j=1}^n \binom{n+1}{j} \int_t^T d\langle X^{(n+1-j),c}, X^{(j),c} \rangle \middle| \mathcal{F}_t \right],$$

where $X^{(j),c}$ is the continuous local martingale part of the semimartingale $X^{(j)}$. This extends a cumulant recursion formula recently obtained by Lacoïn et al. [9] and Friz et al. [5].

After recalling the complete Bell polynomials and their properties in Section 2, we introduce conditional cumulants in Section 3. Then, we give our main results in Section 4. The application to the cumulant recursion formula is given in Section 5.

2 The complete Bell polynomials

Here, we recall the complete Bell polynomials and some of their properties we use in this work. This section does not contain any new results but we include proofs for the readers’ convenience.

Definition 2.1 (The complete Bell polynomials). Let $n \in \mathbb{N}$. For $(x_1, \dots, x_n) \in \mathbb{R}^n$, the n th complete Bell polynomial $B_n(x_1, \dots, x_n)$ is defined by

$$B_n(x_1, \dots, x_n) = \frac{\partial^n}{\partial z^n} \exp \left(\sum_{i=1}^n x_i \frac{z^i}{i!} \right) \Big|_{z=0} \tag{2.1}$$

with $B_0 = 1$.

Example 2.2. Here are some examples of the complete Bell polynomials.

$$\begin{aligned} B_1(x_1) &= x_1, \\ B_2(x_1, x_2) &= x_1^2 + x_2, \\ B_3(x_1, x_2, x_3) &= x_1^3 + 3x_1x_2 + x_3, \\ B_4(x_1, x_2, x_3, x_4) &= x_1^4 + 6x_1^2x_2 + 4x_1x_3 + 3x_2^2 + x_4, \\ B_n(x_1, 0, \dots, 0) &= \frac{\partial^n}{\partial z^n} \sum_{k=0}^{\infty} \frac{1}{k!} (x_1 z)^k \Big|_{z=0} = \frac{\partial^n}{\partial z^n} \frac{1}{n!} (x_1 z)^n \Big|_{z=0} = (x_1)^n, \\ B_n(0, \dots, 0, x_n) &= \frac{\partial^n}{\partial z^n} \sum_{k=0}^{\infty} \frac{1}{k!} \left(x_n \frac{z^n}{n!} \right)^k \Big|_{z=0} = \frac{\partial^n}{\partial z^n} \left(x_n \frac{z^n}{n!} \right) \Big|_{z=0} = x_n. \end{aligned}$$

Proposition 2.3 (The generating function). If one of

$$\sum_{n=1}^{\infty} x_n \frac{z^n}{n!} \quad \text{and} \quad \sum_{n=0}^{\infty} B_n(x_1, \dots, x_n) \frac{z^n}{n!}$$

is absolutely convergent on a neighbourhood of $z = 0$, then so is the other and

$$\exp \left(\sum_{n=1}^{\infty} x_n \frac{z^n}{n!} \right) = \sum_{n=0}^{\infty} B_n(x_1, \dots, x_n) \frac{z^n}{n!} \tag{2.2}$$

on a neighborhood of $z = 0$.

Proof. If, for example, the first series is absolutely convergent on a neighborhood of $z = 0$, then the left hand side of (2.2) is analytic on the neighborhood and so, admits the Taylor expansion around $z = 0$. The coefficients are determined as

$$\frac{\partial^n}{\partial z^n} \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{i=1}^{\infty} x_i \frac{z^i}{i!} \right)^k \Big|_{z=0} = \frac{\partial^n}{\partial z^n} \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{i=1}^n x_i \frac{z^i}{i!} \right)^k \Big|_{z=0} = B_n(x_1, \dots, x_n). \quad \square$$

Proposition 2.4 (A binomial type relation). Let $n \in \mathbb{N}$ and $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$. Then,

$$B_n(x_1 + y_1, \dots, x_n + y_n) = \sum_{j=0}^n \binom{n}{j} B_{n-j}(x_1, \dots, x_{n-j}) B_j(y_1, \dots, y_j). \tag{2.3}$$

Proof. By the Leibniz rule, we have

$$\begin{aligned} B_n(x_1 + y_1, \dots, x_n + y_n) &= \frac{\partial^n}{\partial t^n} \exp \left(\sum_{i=1}^n (x_i + y_i) \frac{t^i}{i!} \right) \Big|_{t=0} \\ &= \left[\sum_{j=0}^n \binom{n}{j} \frac{\partial^{n-j}}{\partial t^{n-j}} \exp \left(\sum_{i=1}^n x_i \frac{t^i}{i!} \right) \frac{\partial^j}{\partial t^j} \exp \left(\sum_{i=1}^n y_i \frac{t^i}{i!} \right) \right] \Big|_{t=0} \\ &= \sum_{j=0}^n \binom{n}{j} B_{n-j}(x_1, \dots, x_{n-j}) B_j(y_1, \dots, y_j). \quad \square \end{aligned}$$

Example 2.5. By (2.3) and the last identity of Example 2.2, for $n \geq 2$,

$$\begin{aligned} B_n(x_1, \dots, x_{n-1}, x_n) &= B_n(x_1, \dots, x_{n-1}, 0) + B_n(0, \dots, 0, x_n) \\ &= B_n(x_1, \dots, x_{n-1}, 0) + x_n. \end{aligned} \tag{2.4}$$

Proposition 2.6. Let $n \in \mathbb{N}$ and $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{R}^n$. If

$$B_k(x_1, \dots, x_k) = B_k(y_1, \dots, y_k)$$

for all $k \leq n$. Then $(x_1, \dots, x_n) = (y_1, \dots, y_n)$.

Proof. The result follows by induction. Indeed, $x_1 = B_1(x_1) = B_1(y_1) = y_1$. If $x_i = y_i$ for all $i \leq k - 1$, then, by (2.4),

$$\begin{aligned} x_k &= B_k(x_1, \dots, x_{k-1}, x_k) - B_k(x_1, \dots, x_{k-1}, 0) \\ &= B_k(y_1, \dots, y_{k-1}, y_k) - B_k(y_1, \dots, y_{k-1}, 0) = y_k. \end{aligned} \quad \square$$

Proposition 2.7. Let $p \geq 1$ and X be an L^p integrable random variable on a probability space (Ω, \mathcal{F}, P) . For a positive integer $n \leq p$, define the n th cumulant κ_n of X by

$$\kappa_n = (-\sqrt{-1})^n \frac{\partial^n}{\partial z^n} \log E[e^{\sqrt{-1}zX}] \Big|_{z=0}. \tag{2.5}$$

Then,

$$E[X^n] = B_n(\kappa_1, \dots, \kappa_n). \tag{2.6}$$

Proof. First note that

$$E[X^n] = (-\sqrt{-1})^n \frac{\partial^n}{\partial z^n} E[e^{\sqrt{-1}zX}] \Big|_{z=0}.$$

From this and (2.5), it is clear that κ_n is a polynomial of $E[X^m]$, $m \leq n$. Note also that $L^n \ni X \mapsto E[X^m] \in \mathbb{R}$ is continuous for $m \leq n$. Therefore $L^n \ni X \mapsto \kappa_m \in \mathbb{R}$ is continuous for $m \leq n$. Since X can be approximated by bounded random variables in L^n , in order to show (2.6), it is then sufficient to consider bounded X . For a bounded random variable X , $z \mapsto E[e^{zX}]$ is analytic and so,

$$\sum_{n=0}^{\infty} E[X^n] \frac{z^n}{n!} = E[e^{zX}] = \exp\left(\sum_{n=1}^{\infty} \kappa_n \frac{z^n}{n!}\right)$$

on a neighborhood of $z = 0$, which implies (2.6) in the light of (2.2). □

See [11] for combinatorial aspects of the Bell polynomials and cumulants.

3 The conditional cumulants

Here we introduce conditional cumulants that play a key role in this work. Let $p \geq 2$ and X be an L^p integrable random variable on a probability space (Ω, \mathcal{F}, P) . Let \mathcal{G} be a sub σ -algebra of \mathcal{F} . If there exists a regular conditional probability measure given \mathcal{G} , then it is natural to define the n th conditional cumulant $X_{\mathcal{G}}^{(n)}$ of X given \mathcal{G} by

$$X_{\mathcal{G}}^{(n)} = (-\sqrt{-1})^n \frac{\partial^n}{\partial z^n} \log E_{\mathcal{G}}[e^{\sqrt{-1}zX}] \Big|_{z=0} \tag{3.1}$$

for a positive integer $n \leq p$, where $E_{\mathcal{G}}$ denotes the expectation with respect to the regular conditional probability measure. Then, it is clear from Proposition 2.7 that

$$E[X^n | \mathcal{G}] = B_n(X_{\mathcal{G}}^{(1)}, \dots, X_{\mathcal{G}}^{(n)}). \tag{3.2}$$

However in general, a regular conditional probability measure might not exist and there might be no event of probability one on which $z \mapsto E[e^{\sqrt{-1}zX}|\mathcal{G}]$ is differentiable, since the conditional expectation is defined only up to a null set for each z . We therefore take another route. In light of Proposition 2.6, we use (3.2) as the defining property of conditional cumulants.

Proposition 3.1. Define $X_{\mathcal{G}}^{(i)}$, $i = 1, \dots, [p]$ by

$$\begin{aligned} X_{\mathcal{G}}^{(1)} &= E[X|\mathcal{G}], \\ X_{\mathcal{G}}^{(n)} &= E[X^n|\mathcal{G}] - B_n(X_{\mathcal{G}}^{(1)}, \dots, X_{\mathcal{G}}^{(n-1)}, 0) \quad \text{for } n \geq 2. \end{aligned}$$

Then, $X_{\mathcal{G}}^{(i)} \in L^{p/i}$ and (3.2) holds for $n \leq p$. Further, $L^p \ni X \mapsto X_{\mathcal{G}}^{(i)} \in L^{p/i}$ is continuous.

Proof. It is clear that $X_{\mathcal{G}}^{(1)} \in L^p$ and that $L^p \ni X \mapsto X_{\mathcal{G}}^{(1)} \in L^p$ is continuous. By Lemma 3.2 below, it follows by induction that $X_{\mathcal{G}}^{(i)} \in L^{p/i}$ and that $L^p \ni X \mapsto X_{\mathcal{G}}^{(i)} \in L^{p/i}$ is continuous. (3.2) is clear from (2.4). \square

Lemma 3.2. Let $n \in \mathbb{N}$ and $p \geq n$. For $(X_1, X_2, \dots, X_n) \in L^p \times L^{p/2} \times \dots \times L^{p/n}$, $B_n(X_1, X_2, \dots, X_n) \in L^{p/n}$. Further, the map

$$L^p \times L^{p/2} \times \dots \times L^{p/n} \ni (X_1, X_2, \dots, X_n) \mapsto B_n(X_1, X_2, \dots, X_n) \in L^{p/n}$$

is continuous.

Proof. $B_n(x_1, \dots, x_n)$ is a linear combination of terms of the form $\prod_{j=1}^k x_{i_j}$, where $i_j \in \{1, \dots, n\}$ with $\sum_{j=1}^k i_j = n$. Therefore it suffices to show that $\prod_{j=1}^k X_j \in L^{p/n}$ for $(X_1, \dots, X_k) \in L^{p/i_1} \times \dots \times L^{p/i_k}$ and that

$$L^{p/i_1} \times \dots \times L^{p/i_k} \ni (X_1, \dots, X_k) \mapsto \prod_{j=1}^k X_j \in L^{p/n}$$

is continuous. By the Hölder inequality,

$$\left\| \prod_{j=1}^k |X_j|^{p/n} \right\|_1 \leq \prod_{j=1}^k \| |X_j|^{p/n} \|_{n/i_j} = \prod_{j=1}^k \|X_j\|_{p/i_j}^{p/n},$$

which implies the result. \square

The following property will be often used in the sequel.

Proposition 3.3. If X is \mathcal{G} -measurable, then $X_{\mathcal{G}}^{(1)} = X$ and $X_{\mathcal{G}}^{(n)} = 0$ for $n \geq 2$.

Proof. It is clear that $X_{\mathcal{G}}^{(1)} = X$. Since $B_n(x, 0, \dots, 0) = x^n$ as seen in Example 2.2, $X_{\mathcal{G}}^{(2)} = X^2 - B_2(X, 0) = 0$. If $X_{\mathcal{G}}^{(n)} = 0$, then $X_{\mathcal{G}}^{(n+1)} = X^{n+1} - B_{n+1}(X, 0, \dots, 0) = 0$. \square

4 The aggregation property of the conditional cumulant processes

Here we give our main results. Let $(\Omega, \mathcal{F}, P, \{\mathcal{F}_t\}_{t \geq 0})$ be a filtered probability space satisfying the usual conditions. For a process $F = \{F_t\}$ and $s \leq t$, $F_{s,t}$ denotes $F_t - F_s$. Let $p \geq 2$ and $X \in L^p$. For $n \leq p$, define the n th conditional cumulant process $X^{(n)} = \{X_t^{(n)}\}$, where $X_t^{(n)} = X_{\mathcal{F}_t}^{(n)}$ defined as in Proposition 3.1 with $\mathcal{G} = \mathcal{F}_t$. This construction allows to take a cadlag version. Let $\mathbb{X}^{(n)} = (X^{(1)}, \dots, X^{(n)})$.

Example 4.1 (Brownian motion). Let $X = W_T$ for $T > 0$ and an $\{\mathcal{F}_t\}$ -Brownian motion $\{W_t\}$. Then, $\log E[e^{\sqrt{-1}zX}|\mathcal{F}_t] = \sqrt{-1}zW_{t\wedge T} - z^2(T-t)_+/2$. Therefore (3.1) implies $X_t^{(1)} = W_{t\wedge T}$, $X_t^{(2)} = (T-t)_+$ and $X^{(n)} = 0$ for $n \geq 3$. In order to see this follows also from (3.2), recall first a defining equation for the Hermite polynomials $H_n(x)$:

$$\exp\left\{tx - \frac{t^2}{2}\right\} = \sum_{n=0}^{\infty} H_n(x) \frac{t^n}{n!}$$

and its two consequences

$$H_n(x) = \int_{\mathbb{R}} (x + \sqrt{-1}z)^n \phi(z) dz, \quad n = 0, 1, 2, \dots,$$

where ϕ is the standard normal density, and

$$B_n(x, y, 0, \dots, 0) = \frac{\sqrt{y}^n}{\sqrt{-1}^n} H_n\left(\frac{\sqrt{-1}x}{\sqrt{y}}\right)$$

for $x \in \mathbb{R}$ $y > 0$ and $n \geq 2$. Then, it suffices to observe

$$E[X^n|\mathcal{F}_t] = \int (W_{t\wedge T} + \sqrt{(T-t)_+}z)^n \phi(z) dz = \frac{\sqrt{(T-t)_+}^n}{\sqrt{-1}^n} H_n\left(\frac{\sqrt{-1}W_{t\wedge T}}{\sqrt{(T-t)_+}}\right).$$

Here we give a key lemma.

Lemma 4.2. For any positive integer $n \leq p$ and for any stopping times $\tau \leq v$,

$$E[B_n(\mathbb{X}_{\tau,v}^{(n)})|\mathcal{F}_\tau] = 0.$$

Proof. By Proposition 3.1 and Lemma 3.2,

$$L^p \ni X \mapsto E[B_n(\mathbb{X}_{\tau,v}^{(n)})|\mathcal{F}_\tau] \in L^{p/n}$$

is continuous. Therefore it suffices to consider $X \in L^\infty$ as L^∞ is dense in L^p . If $X \in L^\infty$, then by Proposition 2.3 and (3.2), we have

$$\frac{1}{E[e^{zX}|\mathcal{F}_\tau]} = \left(\sum_{n=0}^{\infty} E[X^n|\mathcal{F}_\tau] \frac{z^n}{n!}\right)^{-1} = \exp\left(-\sum_{n=1}^{\infty} X_\tau^{(n)} \frac{z^n}{n!}\right) = \sum_{n=0}^{\infty} B_n(-\mathbb{X}_\tau^{(n)}) \frac{z^n}{n!}$$

on a neighborhood of $z = 0$. This implies

$$1 = \left(\sum_{n=0}^{\infty} E[X^n|\mathcal{F}_\tau] \frac{z^n}{n!}\right) \left(\sum_{n=0}^{\infty} B_n(-\mathbb{X}_\tau^{(n)}) \frac{z^n}{n!}\right) = \sum_{n=0}^{\infty} \sum_{j=0}^n \binom{n}{j} E[X^{n-j}|\mathcal{F}_\tau] B_j(-\mathbb{X}_\tau^{(j)}) \frac{z^n}{n!}$$

and so for $n \geq 1$,

$$0 = \sum_{j=0}^n \binom{n}{j} E[X^{n-j}|\mathcal{F}_\tau] B_j(-\mathbb{X}_\tau^{(j)}) = E\left[\sum_{j=0}^n \binom{n}{j} E[X^{n-j}|\mathcal{F}_v] B_j(-\mathbb{X}_\tau^{(j)}) \Big| \mathcal{F}_\tau\right].$$

The right hand side coincides with $E[B_n(\mathbb{X}_v^{(n)} - \mathbb{X}_\tau^{(n)})|\mathcal{F}_\tau]$ by Proposition 2.4 and (3.2). \square

Now we give the main result of this paper.

Theorem 4.3. For a positive integer $n \leq p - 1$, define $g_n : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$g_n(x_1, \dots, x_n) = B_{n+1}(x_1, \dots, x_n, 0).$$

Then,

$$E[g_n(\mathbb{X}_{\tau,v}^{(n)})|\mathcal{F}_\tau] = -E[X_{\tau,v}^{(n+1)}|\mathcal{F}_\tau] \tag{4.1}$$

for any stopping times $\tau \leq v$. In particular, g_n satisfies the aggregation property with $\mathbb{X}^{(n)}$, that is,

$$E[g_n(\mathbb{X}_{\sigma,v}^{(n)})|\mathcal{F}_\sigma] = E[g_n(\mathbb{X}_{\sigma,\tau}^{(n)})|\mathcal{F}_\sigma] + E[g_n(\mathbb{X}_{\tau,v}^{(n)})|\mathcal{F}_\sigma] \tag{4.2}$$

for any stopping times $\sigma \leq \tau \leq v$. If X is \mathcal{F}_v measurable for a stopping time v , then

$$E[g_n(\mathbb{X}_{\tau,v}^{(n)})|\mathcal{F}_\tau] = X_\tau^{(n+1)} \tag{4.3}$$

for any stopping time $\tau \leq v$.

Proof. (4.1) is clear from (2.4) and Lemma 4.2. (4.2) is clear from (4.1). If X is \mathcal{F}_v measurable, then $E[X^m|\mathcal{F}_v] = X^m$ for any m and so $\mathbb{X}_v^{(n+1)} = (X, 0, \dots, 0)$ by the second last identity in Example 2.2, which implies (4.3) together with (4.1). \square

Actually we have a stronger property than (4.2):

Theorem 4.4. For any positive integer $n \leq p - 1$ and for any stopping times $\sigma \leq \tau \leq v$,

$$E[g_n(\mathbb{X}_{\sigma,v}^{(n)})|\mathcal{F}_\tau] = g_n(\mathbb{X}_{\sigma,\tau}^{(n)}) + E[g_n(\mathbb{X}_{\tau,v}^{(n)})|\mathcal{F}_\tau]$$

Proof. By (2.3),

$$B_{n+1}(\mathbb{X}_{\sigma,v}^{(n+1)}) = \sum_{j=0}^{n+1} \binom{n+1}{j} B_j(\mathbb{X}_{\sigma,\tau}^{(j)}) B_{n+1-j}(\mathbb{X}_{\tau,v}^{(n+1-j)}).$$

Lemma 4.2 then implies

$$E[B_{n+1}(\mathbb{X}_{\sigma,v}^{(n+1)})|\mathcal{F}_\tau] = B_{n+1}(\mathbb{X}_{\sigma,\tau}^{(n+1)}). \tag{4.4}$$

The result then follows from (2.4) and (4.1). \square

Remark 4.5. From (4.4), we have that $B_n(\mathbb{X}_{0,t}^{(n)})$ is a martingale. This should be distinguished from the fact that $B_n(\mathbb{X}_t^{(n)}) = E[X^n|\mathcal{F}_t]$ is a martingale (see (3.2)).

Example 4.6. For an application, here we give an alternative proof for a well-known fact (see e.g., Proposition IV.3.8 of [12]) that $H_n(M, \langle M \rangle)$ is a continuous local martingale for any continuous local martingale M with $M_0 = 0$ and for any $n = 0, 1, 2, \dots$, where

$$H_n(x, t) = \sqrt{t}^n H_n\left(\frac{x}{\sqrt{t}}\right)$$

and $H_n(x)$ is the n th Hermite polynomial (see Example 4.1). By localization, we can assume without loss of generality that $\langle M \rangle$ is bounded, say, by a constant $T > 0$. By the Dambis-Dubins-Schwarz theorem (Theorem 3.4.6 of [8]), $M = W_{\langle M \rangle}$ for a Brownian motion W (with respect to a time-changed filtration). As seen in Example 4.1, for $X = W_T$, we have for any stopping time (with respect to the time-changed filtration) $\tau \leq T$,

$$B_n(\mathbb{X}_{0,\tau}^{(n)}) = H_n(W_\tau, \tau).$$

Lemma 4.2 then implies $E[H_n(W_\tau, \tau)] = 0$ and so, $E[H_n(M_\tau, \langle M \rangle_\tau)] = 0$ for any stopping time (with respect to the original filtration) τ . This means that $H_n(M, \langle M \rangle)$ is a martingale; see e.g., Problem 1.3.26 of [8].

Example 4.7. Let $M = \{M_t\}_{t \in [0, T]}$ is an L^p martingale. Taking $X = M_T$, we have $X^{(1)} = M$ and $X_T^{(n)} = 0$ for $n \geq 2$. As is well-known, the second cumulant is the variance and so, we have $X^{(2)} = M^{(2)}$, where $M_t^{(n)} = E[(M_{t,T})^n | \mathcal{F}_t]$ as in Introduction. For $n = 1$, since $g_n(x_1) = B_2(x_1, 0) = x_1^2$, (4.1) means $E[(M_{t,u})^2 | \mathcal{F}_t] = E[(M_{t,T})^2 | \mathcal{F}_t] - E[(M_{u,T})^2 | \mathcal{F}_t]$. For a partition $0 = t_0 < \dots < t_N = T$,

$$\sum_{j=1}^N g_1(\mathbb{X}_{t_{j-1}, t_j}^{(1)}) = \sum_{j=1}^N (M_{t_{j-1}, t_j})^2$$

is the so-called realized variance. The aggregation property implies

$$E \left[\sum_{j=1}^N (M_{t_{j-1}, t_j})^2 \right] = E \left[\sum_{j=1}^N g_1(\mathbb{X}_{t_{j-1}, t_j}^{(1)}) \right] = E[g_1(\mathbb{X}_{0,T}^{(1)})] = E[(M_{0,T})^2].$$

The third cumulant is known to be the centered third moment, that is, $X^{(3)} = M^{(3)}$. For $n = 2$, $g_n(x, y) = B_3(x, y, 0) = x^3 + 3xy$. Therefore

$$\sum_{j=1}^N g_n(\mathbb{X}_{t_{j-1}, t_j}^{(n)}) \tag{4.5}$$

with $n = 2$ is Neuberger’s realized skewness. Its expectation is $-E[X_{0,T}^{(3)}] = E[X_0^{(3)}] = E[(M_{0,T})^3]$ by the aggregation property. For $n = 3$, $g_n(x, y, z) = B_4(x, y, z, 0) = x^4 + 6x^2y + 4xz + 3y^2$. Therefore (4.5) with $n = 3$ is Bae and Lee’s realized kurtosis. Its expectation is $-E[X_{0,T}^{(4)}] = E[X_0^{(4)}]$ by the aggregation property.

Based on Theorem 4.3 and Example 4.7, we suggest to call the high-frequency statistic of the form (4.5) the realized $(n + 1)$ -th cumulant of the martingale $X^{(1)}$. It is an unbiased estimator of the $(n + 1)$ -th cumulant $X_0^{(n+1)}$ of X when X is \mathcal{F}_T measurable and \mathcal{F}_0 consists of null sets and their complements.

We conclude this section with a brief explanation on how the realized cumulants can be used in financial econometrics. Suppose that X represents an asset price at time T and there is a market of call and put options with maturity T written on the asset. The market prices of the options determine a probability measure (called a risk neutral measure) under which each option price is expressed as the expectation of its option payoff. The moments and hence cumulants of X under the risk neutral measure at any time $t < T$ can be therefore computed from the option market prices at time t . See e.g., [7] for some formulas. The realized n th cumulant is an ex-post value based on time series of those conditional cumulant processes, whose expectation under the risk neutral measure at time 0 coincides with $X_0^{(n)}$. Any systematic deviation of the realized n th cumulant value from $X_0^{(n)}$ is due to the difference between the risk neutral measure and the physical probability measure, and is interpreted as a risk premium. See Neuberger [10] and the references therein for analyses of variance risk premium and skewness risk premium.

5 Application to a cumulant recursion formula

It is clear from definition that $X^{(n)}$ is an $L^{p/n}$ semimartingale. The high-frequency limit of the realized $(n + 1)$ -th cumulant on $[t, T]$ is

$$\sum_{s \in (t, T]} g_n(\Delta \mathbb{X}_s^{(n)}) + \frac{1}{2} \sum_{j=1}^n \binom{n+1}{j} \langle X^{(n+1-j), c}, X^{(j), c} \rangle_{t, T},$$

where $X^{(j),c}$ is the continuous local martingale part of $X^{(j)}$. Here, the convergence is in probability. The second term comes from the fact that the quadratic terms contained in $B_{n+1}(x_1, \dots, x_{n+1})$ are

$$\frac{1}{2} \sum_{j=1}^n \binom{n+1}{j} x_{n+1-j} x_j$$

since

$$\exp \left(\sum_{i=1}^{\infty} x_i \frac{z^i}{i!} \right) = 1 + \sum_{i=1}^{\infty} x_i \frac{z^i}{i!} + \frac{1}{2} \left(\sum_{i=1}^{\infty} x_i \frac{z^i}{i!} \right)^2 + \dots$$

If $X \in L^p$ for any $p > 1$, then so is $X^{(n)}$. Then, an application of Itô's formula to $g_n(\mathbb{X}_{s,t})$ shows that the above high-frequency convergence is also in L^1 . The aggregation property then implies

$$E[g_n(\mathbb{X}_{t,T}^{(n)}) | \mathcal{F}_t] = E \left[\sum_{s \in (t,T]} g_n(\Delta \mathbb{X}_s^{(n)}) + \frac{1}{2} \sum_{j=1}^n \binom{n+1}{j} \langle X^{(n+1-j),c}, X^{(j),c} \rangle_{t,T} \middle| \mathcal{F}_t \right].$$

In case X is \mathcal{F}_T measurable, by (4.3), we obtain a recursion formula

$$X_t^{(n+1)} = E \left[\sum_{s \in (t,T]} g_n(\Delta \mathbb{X}_s^{(n)}) + \frac{1}{2} \sum_{j=1}^n \binom{n+1}{j} \langle X^{(n+1-j),c}, X^{(j),c} \rangle_{t,T} \middle| \mathcal{F}_t \right]. \quad (5.1)$$

Moreover if the filtration is continuous, then $\Delta \mathbb{X}^{(n)} = 0$ and so,

$$X_t^{(n+1)} = \frac{1}{2} \sum_{j=1}^n \binom{n+1}{j} E \left[\langle X^{(n+1-j)}, X^{(j)} \rangle_{t,T} \middle| \mathcal{F}_t \right]$$

or equivalently, for $Y^{(j)} = X^{(j)}/j!$,

$$Y^{(n+1)} = \frac{1}{2} \sum_{j=1}^n Y^{(n+1-j)} \diamond Y^{(j)},$$

where \diamond is the diamond operation introduced by Alòs et al. [1]. This last formula has been recently obtained by Lacoïn et al. [9] in their study of Quantum Field Theory. The assumption of $X \in \bigcap_{p>1} L^p$ can be relaxed; see Friz et al. [5]. See [5] also for several nice applications of this recursion formula. More recently (after the initial submission of this paper, independently from this work), the extension (5.1) is also given in Friz et al. [6] by a different approach.

References

- [1] Alòs, E., Gatheral, J. and Radoičić, R.: Exponentiation of conditional expectations under stochastic volatility. *Quant. Finance* **20**, (2020), 13–27. MR-4040259
- [2] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P.: The distribution of realized exchange rate volatility. *J. Amer. Statist. Assoc.* **96**, (2001), 42–55. MR-1952727
- [3] Bae, K. and Lee, S.: Realized higher-order comoments. *Quant. Finance* forthcoming.
- [4] Cohen, S.N and Elliott, R.: Stochastic calculus and applications. Second edition. *Springer*, Cham, (2015). MR-3443468
- [5] Friz, P.K., Gatheral, J. and Radoičić, R.: Forests, cumulants, martingales. arXiv, (2020), arXiv:2002.01448v3.
- [6] Friz, P.K., Hager, P. and Tapia, N.: Unified signature cumulants and generalized Magnus expansions. arXiv, (2021), arXiv:2102.03345.

- [7] Fukasawa, M.: The normalizing transformation of the implied volatility smile, *Math. Finance* **22**, (2012), 753–762. MR-2968285
- [8] Karatzas, I. and Shreve, S.E.: Brownian motion and stochastic calculus. Second edition. *Springer-Verlag*, New York, (1991). MR-1121940
- [9] Lacoïn, H., Rhodes R., and Vargas, V.: A probabilistic approach of ultraviolet renormalisation in the boundary Sine-Gordon model. *arXiv*, (2019), arXiv:1903.01394.
- [10] Neuberger, A.: Realized skewness. *Rev. Financ. Stud.* 25, (2012), 3423–3455.
- [11] Peccati, G. and Taqqu, M.S.: Wiener chaos: moments, cumulants and diagrams. *Springer*, Milan, (2011). MR-2791919
- [12] Revuz, D. and Yor, M.: Continuous martingales and Brownian motion. Third edition. *Springer-Verlag*, Berlin, (1999). MR-1725357