Probab. Theory Relat. Fields 103, 475 – 492 (1995)

Probability Theory and Related Fields

Embedding and asymptotic expansions for martingales

P.A. Mykland

Department of Statistics, The University of Chicago, Chicago, IL 60637, USA
(Fax: +1-312/702 9810; e-mail: mykland@galton.uchicago.edu)

Received: 20 June 1994 / In revised form: 17 April 1995

Summary. The paper develops a way of embedding general martingales in continuous ones in such a way that the quadratic variation of the continuous martingale has conditional cumulants (given the original martingale) that are explicitly given in terms of optional and predictable variations of the original process. Bartlett identities for the conditional cumulants are also found. A main corollary to these results is the establishment of second (and in some cases higher) order asymptotic expansions for martingales.

Mathematics Subject Classification (1991): 60G42, 60G44, 60F99, 62E20

1 Beyond the black box

The embedding of martingales in Brownian motion is a powerful tool, particularly in asymptotic theory (see Hall and Heyde (1980) or Khoshnevisan (1993), for example). There are two types of results: (i) any martingale can be represented as a time changed Brownian motion, and (ii) any discrete time martingale can be interpolated with a continuous one. There is a substantial amount of literature in this area, including Dambis (1965), Dubins and Schwartz (1965), Skorokhod (1965), Sawyer (1967), Dubins (1968), Root (1969), Clark (1970), Monroe (1972), Chacon and Walsh (1976), Rost (1976), Heath (1977), Azéma and Yor (1979), Meilijson (1983) and Khoshnevisan (1993).

The papers cited use a variety of constructions of the embedding. Mainly (though see, for example, Rost, 1976; Khoshnevisan, 1993), the principle was

Research supported in part by National Science Foundation grant DMS 93-05601 and Army Research Office grant DAAH04-1-0105
This manuscript was prepared using computer facilities supported in part by the National Science Foundation grants DMS 89-05292, DMS 87-03942, and DMS 86-01732 awarded to the Department of Statistics at The University of Chicago, and by The University of Chicago Block Fund

that any construction would do, one would just use it as a black box in subsequent manipulations. On the whole, this strategy has functioned quite well in central limit theory and for the law of the iterated logarithm.

Our reason for raising a 10 year old topic, however, is that the above approach runs into problems in connection with higher order asymptotics. It works, but it does not work very well. For example, in Mykland (1993), a key step in the development is to show that, in an asymptotic sense,

$$E(\tau_t|M_s, 0 \leqq s \leqq t) \approx \tfrac{1}{3}[M,M]_t + \tfrac{2}{3}\langle M,M\rangle_t , \tag{1.1}$$

where M_t is a discrete time martingale which can be embedded in Brownian motion as $M_t = W_{\tau_t}$. In other words, one is interested in the conditional distribution of the embedding, and one has to go through longish arguments to show that a black box embedding satisfies (1.1). This gets worse for higher order expansions. A similar need to have more precise properties of the embedding motivates the work of Khoshnevisan (1993).

This leads to the question of whether one can construct embeddings with nicely controllable properties. The purpose of this paper is to partially answer this by constructing a specific embedding for which (1.1) holds exactly. We also give a formula for the conditional variance of τ_t. In the quasi-left continuous case, we give all conditional cumulants.

As a corollary to these results, we derive a two step asymptotic expansion for martingales. For quasi-left continuous martingales, higher order expansions also follow.

2 Explicit structure

We shall consider embeddings for which

$$E(\tau_t|\mathscr{F}) = \tfrac{1}{3}[M,M]_t + \tfrac{2}{3}\langle M,M\rangle_t ,$$

where $\mathscr{F}$ is a σ-field with respect to which the martingale (M_t) is measurable, and similarly for higher order conditional cumulants. To characterize such embeddings, let $[Y,\dots,Y]_t, \langle Y,\dots,Y\rangle_t$ and $\kappa_p(Y)_t = \kappa(Y,\dots,Y)_t$ be the *pth* order optional, predictable and cumulant variations, respectively, of the *càdlàg* process (Y_t) (see Sections 2 and 6 of Mykland, 1994). In the discrete time case, if

$$Y_t = \sum_{n=1}^{t} X_n ,$$

the variations are given by

$$[Y,\dots,Y]_t = \sum_{n=1}^{t} X_n^p ,$$

$$\langle Y,\dots,Y\rangle_t = \sum_{n=1}^{t} E(X_n^p|\mathscr{F}_{n-1})$$

and

$$\kappa(Y,\ldots,Y)_t = \sum_{n-1}^{t} \mathrm{cum}_p(X_n|\mathscr{F}_{n-1}),$$

where $(\mathscr{F}_t)$ is the relevant filtration. Analogous definitions apply in the multivariate case. We now define the family of desired embeddings.

Definition. Let (M_t) be a martingale with finite $2p$th moment, and let $\mathscr{F}$ be a σ-field with respect to which (M_t) is measurable. An embedding of this martingale in Brownian motion will be said to have *explicit structure* up to order p if $E|\tau_t|^p$ is finite and (for $q \leqq p$) $\mathrm{cum}_q(\tau_t|\mathscr{F})$ can be expressed as a linear combination of variations

$$[M,\ldots,M,\kappa(M,\ldots,M),\ldots,\kappa(M,\ldots,M)]_t,$$

where "M" appears exactly $2q$ times, or, equivalently, as a linear combination of variations

$$[M,\ldots,M,\langle M,\ldots,M\rangle,\ldots,\langle M,\ldots,M\rangle]_t.$$

The same terminology applies by analogy to interpolation of discrete time martingales.

For example, one should be able to express $\mathrm{var}(\tau_t|\mathscr{F})$ as a linear combination of $[M,M,M,M]_t$, $[M,\kappa(M,M,M)]_t = [M,\langle M,M,M\rangle]_t$, $[\langle M,M\rangle,\langle M,M\rangle]_t$ and $\kappa(M,M,M,M)_t$ (or $\langle M,M,M,M\rangle_t$).

The reason for focusing on cumulants rather than moments is that in the discrete time case, embeddings are (typically) constructed to be conditionally independent given $\mathscr{F}$ in each interval, whence

$$\mathrm{cum}_p(\tau_t|\mathscr{F}) = \sum_{n=1}^{t} \mathrm{cum}_p(\tau_n - \tau_{n-1}|\mathscr{F}). \tag{2.1}$$

This means that constructing an embedding with explicit structure reduces to constructing it in each time interval. Obviously, this will not work for moments.

Note that in the continuous time case, the analogous formula to (2.1) is that, for $p \geqq 2$,

$$\mathrm{cum}_p(\tau_t|\mathscr{F}) = \sum_{0\leqq s\leqq t} \mathrm{cum}_p(\Delta\tau_s|\mathscr{F}). \tag{2.2}$$

This follows from τ_t being an increasing process with conditionally (given the data) independent increments. It is easy to see that the process (2.2) is càdlàg: $[\tau,\ldots,\tau]_t$ can clearly be taken to be càdlàg, and hence the same applies to $E([\tau,\ldots,\tau]_t|\mathscr{F})$. The cumulants are then defined from moments in the usual way (see McCullagh, 1987, Chap. 2), and the càdlàg property follows.

3 Bartlett identities, and an existence theorem

If $M_t = W_{\tau_t}$, it follows (subject to integrability conditions) that, for small θ,

$$\exp\left(\theta M_t + \sum_{p=1}^{\infty} \theta^{2p}\left(-\tfrac{1}{2}\right)^p \frac{1}{p!}\operatorname{cum}_p(\tau_t|\mathscr{F})\right) = E\left[\exp\left(\theta M_t - \tfrac{1}{2}\theta^2\tau_t\right)|\mathscr{F}\right] \tag{3.1}$$

is a likelihood. This is because it integrates to 1, since $\exp(\theta W_t - \frac{1}{2}\theta^2 t)$ is a martingale. Hence $\ell_t^{(p)}$ given by

$$\begin{aligned} \ell_t^{(1)} &= M_t\,, \\ \ell_t^{(2p)} &= \left(-\tfrac{1}{2}\right)^p \frac{2p!}{p!}\operatorname{cum}_p(\tau_t|\mathscr{F}), \quad p \geqq 1\,, \\ \ell_t^{(2p+1)} &= 0\,, \quad p \geqq 1 \end{aligned} \tag{3.2}$$

can be seen as derivatives (at 0) of a log likelihood, and consequently they satisfy Bartlett identities (Bartlett 1953a, b; see also McCullagh, 1987, Chap. 7). There are a number of variations over these identities, see, for example, the discussion in Section 2 of Mykland and Ye (1992). If one assumes that the ℓ_t's are adapted, *càdlàg*, and satisfy (2.2), there also exists a conditional cumulant version, where, e.g., the third identity takes the form

$$\kappa(\ell^{(1)},\ell^{(1)},\ell^{(1)})_t + 3\kappa(\ell^{(1)},\ell^{(2)})_t + \kappa(\ell^{(3)})_t = 0\,. \tag{3.3}$$

This is of course heuristic. A formal result is stated below as Theorem 2. If we go back to the variation notation, the first five identities have the following form:

$$\begin{aligned} &\kappa(M)_t = 0\,, \\ &\kappa_2(M)_t - \kappa(E(\tau_.|\mathscr{F}))_t = 0\,, \\ &\kappa_3(M)_t - 3\kappa(M,E(\tau_.|\mathscr{F}))_t = 0\,, \\ &\kappa_4(M)_t - 6\kappa(M,M,E(\tau_.|\mathscr{F}))_t \\ &\quad + 3\kappa_2(E(\tau_.|\mathscr{F}))_t + 3\kappa(\operatorname{var}(\tau_.|\mathscr{F}))_t = 0\,, \\ &\kappa_5(M)_t - 10\kappa(M,M,M,E(\tau_.|\mathscr{F}))_t \\ &\quad + 15\kappa(M,E(\tau_.|\mathscr{F}),E(\tau_.|\mathscr{F}))_t + 15\kappa(M,\operatorname{var}(\tau_.|\mathscr{F}))_t = 0\,. \end{aligned} \tag{3.4}$$

This holds, obviously, without reference to explicit structure. However, if the embedding does have such structure of order (at least) one, and if M_t has finite third absolute moment, then the second and third identities imply that

$$E(\tau_t|\mathscr{F}) = \tfrac{1}{3}[M,M]_t + \tfrac{2}{3}\langle M,M\rangle_t \tag{3.5}$$

unless $\kappa(M,M,M)_t$ is zero a.s. (The derivation is given in Section 7.) Similarly, the fourth and fifth identities yield

$$\begin{aligned} \operatorname{var}(\tau_t|\mathscr{F}) = {} & \tfrac{8}{45}\kappa(M,M,M,M)_t + \tfrac{2}{45}[M,M,M,M]_t \\ & + \left(\tfrac{44}{45} - c\right)[\langle M,M\rangle,\langle M,M\rangle]_t \\ & + c[M,M,\langle M,M\rangle]_t - c[M,\kappa(M,M,M)]_t\,, \end{aligned} \tag{3.6}$$

where c is nonrandom. This is provided M_t has fifth absolute moment and $\kappa_5(M)_t$ and $[\kappa_2(M), \kappa_3(M)]_t$ are not zero a.s. It should be emphasized that if there is a constant nonrandom linear relation between any of the variations appearing in (3.6), other relations may obviously also hold, and similarly for (3.5). The extreme case (in discrete time) is where the increment of M_t has a symmetric binary distribution. In that case, $\langle M,M,M,M\rangle_t$, $[M,M,M,M]_t$, $[\langle M,M\rangle, \langle M,M\rangle]_t, [M,M,\langle M,M\rangle]_t$ and $[M, \kappa(M,M,M)]_t$ are all the same. In this case, one can for example, write $\mathrm{var}(\tau_t|\mathscr{F}) = \frac{2}{3}[M,M,M,M]_t$.

On the other hand, it is worth noting that if M is quasi-left continuous (see Jacod and Shiryaev, 1987, p. 22), the terms in (3.6) involving c all vanish: there is only one possible expression for the conditional variance (when there is explicit structure). An argument similar to the one leading to (3.6) easily shows that the same is the case for the higher order conditional cumulants. $\mathrm{cum}_p(\tau_t|\mathscr{F})$ is a unique linear combination of $\langle M,\ldots,M\rangle_t$ and $[M,\ldots,M]_t$, whence, in view of the Bartlett identities,

$$\mathrm{cum}_p(\tau_t|\mathscr{F}) = (-1)^{p+1}2^{3p}\tfrac{p!}{2p!}B_{2p}\{\underbrace{\langle M,\ldots,M\rangle_t}_{2p \text{ times}} + \tfrac{1}{2p}\underbrace{[M,\ldots,M]_t}_{2p \text{ times}}\}\,, \tag{3.7}$$

where B_{2k} is the $2k$th Bernoulli number ($B_2 = \frac{1}{6}$, $B_4 = -\frac{1}{30}$, $B_6 = \frac{1}{42}$, etc.), see pp. 804–810 of Abramowitz and Stegun (1972).

Any c is compatible with the Bartlett identities. We do not know, however, whether there is an embedding for every value of c. It is clear, for example, that unless the predictable jumps (or increments) of (M_t) are bounded, the nonnegativity of (3.6) implies that

$$c \geqq 0\,.$$

The embeddings we do know of are the following:

Theorem 1 (Embedding theorem) *Let M_t, $t \geqq 0$, be a càdlàg martingale for which $[M,\ldots,M]_t$ (the order $2p$ optional variation, p integer) is integrable. Then there is an extension of the original probability space with a filtration $(\mathscr{G}_t)$, a $(\mathscr{G}_t)$ Wiener process (W_t) and a family $\tau(t)$ of $(\mathscr{G}_t)$ stopping times, with conditionally independent increments given the data, so that the process $W(\tau(t))$ is indistinguishable from M_t, and so that the embedding has explicit structure up to order p. In particular,* (3.5) *is satisfied. If $p \geqq 2$ and c is a nonrandom number satisfying*

$$\tfrac{16}{45} < c \leqq \tfrac{136}{315}\,, \tag{3.8}$$

the embedding can be chosen to satisfy (3.6).

Whether embeddings exist for c's not satisfying (3.8) is not known to the author in the general case. We have also not investigated the formulas for higher order embeddings for other than quasi-left continuous martingales.

The results on Bartlett identities can be formalized as follows.

Theorem 2 (Bartlett identities) *Let (M_t) be a martingale with at least pth absolute moment ($p \geqq 1$, integer), and assume that it is embedded in a Brownian motion in such a way that $E\tau_t^{p/2} < \infty$. Let $\ell_t^{(k)}$, $k = 1,\ldots,p$, be given by* (3.2)*. Then the first p ordinary Bartlett identities hold, in the sense that, for $q \leqq p$,*

$$\sum \frac{q!}{\Pi(k!)^{q_k} q_k!} E\ell_t^{(1)q_1} \ldots \ell_t^{(k)q_k} \ldots = 0 \tag{3.9}$$

and

$$\sum \frac{q!}{\Pi(k!)^{q_k} q_k!} \operatorname{cum}(\ell_t^{(1)q_1}, \ldots, \ell_t^{(k)q_k}, \ldots) = 0\,, \tag{3.10}$$

where the sum is over all $q_1, q_2, \ldots$ so that $q_1 + 2q_2 + \cdots + kq_k + \cdots = q$, and where each combination only occurs once. If, in addition, τ_t has independent increments given the data, and if the $\ell_t^{(k)}$ are adapted and càdlàg, then the p first moment and cumulant Bartlett identities of type (3.3) *also hold, in other words*

$$\sum \frac{q!}{\Pi(k!)^{q_k} q_k!} \langle \underbrace{\ell^{(1)}, \ldots, \ell^{(1)}}_{q_1 \text{ times}}, \ldots, \underbrace{\ell^{(k)}, \ldots, \ell^{(k)}}_{q_k \text{ times}}, \ldots \rangle_t = 0 \tag{3.11}$$

and

$$\sum \frac{q!}{\Pi(k!)^{q_k} q_k!} \kappa(\underbrace{\ell^{(1)}, \ldots, \ell^{(1)}}_{q_1 \text{ times}}, \ldots, \underbrace{\ell^{(k)}, \ldots, \ell^{(k)}}_{q_k \text{ times}}, \ldots)_t = 0\,. \tag{3.12}$$

Similar results hold for interpolation of martingales, which can also be done for multivariate martingales. If $(M_t^1, \ldots, M_t^p)$ is a martingale, embedded in $(\overline{M}_t^1, \ldots, \overline{M}_t^p)$, the likelihood would be such that $\partial\ell_t/\partial\theta^\alpha = M_t^\alpha$, $\partial^{2p+1}\ell_t/\partial\theta^{\alpha_1} \ldots \partial\theta^{\alpha_{2p+1}} = 0$, and

$$\frac{\partial^{2p}\ell_t(\theta)}{\partial\theta^{\alpha_1} \ldots \partial\theta^{\alpha_{2p}}} = (-1)^p \sum \operatorname{cum}_{2p}(\langle \overline{M}^{\beta_1}, \overline{M}^{\beta_2} \rangle, \ldots, \langle \overline{M}^{\beta_{2p-1}}, \overline{M}^{\beta_{2p}} \rangle_t | \mathscr{F})\,,$$

where the sum is over all ways of partitioning $\{\alpha_1, \ldots, \alpha_{2p}\}$ into pairs $\{\{\beta_1, \beta_2\}, \ldots, \{\beta_{2p-1}, \beta_{2p}\}\}$. The formula assumes that the α's are distinct. (If they are not, one should pretend that they are when carrying out calculations).

The generalization of Theorem 2 would be in the form indicated in Chap. 7 of McCullagh (1987). In particular, note that they imply that for martingales (M_t) and (N_t),

$$E(\langle \overline{M}, \overline{N} \rangle_t | \mathscr{F}) = \tfrac{1}{3}[M,N]_t + \tfrac{2}{3}\langle M,N \rangle_t\,,$$

and

$$\begin{aligned}
&\operatorname{cov}(\langle \overline{M}^{\alpha}, \overline{M}^{\beta}\rangle_t, \langle \overline{M}^{\gamma}, \overline{M}^{\delta}\rangle_t | \mathscr{F})\,[3] \\
&\quad = \tfrac{8}{15}\kappa(M^{\alpha}, M^{\beta}, M^{\gamma}, M^{\delta})_t + \tfrac{2}{15}[M^{\alpha}, M^{\beta}, M^{\gamma}, M^{\delta}]_t \\
&\quad \times \left(\tfrac{44}{45} - c\right)[\langle M^{\alpha}, M^{\beta}\rangle, \langle M^{\gamma}, M^{\delta}\rangle]_t\,[3] + \frac{c}{2}[M^{\alpha}, M^{\beta}, \langle M^{\gamma}, M^{\delta}\rangle]_t\,[6] \\
&\quad - \frac{3c}{4}[M^{\alpha}, \kappa(M^{\beta}, M^{\gamma}, M^{\delta})]_t\,[4]\,, \qquad (3.13)
\end{aligned}$$

where [3], [4] and [6] indicates summation over all relevant combinations, see, for example, McCullagh (1987). By expanding the conditional variance given the data of $\langle \overline{M} + \overline{N}, \overline{M} + \overline{N}\rangle_t = \langle \overline{M}, \overline{M}\rangle_t + 2\langle \overline{M}, \overline{N}\rangle_t + \langle \overline{N}, \overline{N}_t\rangle$, it is easily seen that when embedding more than one martingale, c must be the same in all the possible formulas (3.13) ($\operatorname{var}(\langle \overline{M}, \overline{M}\rangle_t | \mathscr{F})$, $\operatorname{cov}(\langle \overline{M}, \overline{N}\rangle_t, \langle \overline{M}, \overline{M}\rangle_t | \mathscr{F})$, $\operatorname{cov}(\langle \overline{M}, \overline{N}\rangle_t | \mathscr{F})$, etc.). Hence, *however many martingales are embedded (on the same space), there is only one degree of freedom for the conditional covariances.*

The formal interpolation result is as follows.

Theorem 3 (Interpolation theorem) *Let M_t, $t = 0, 1, 2, \ldots$ be a (vector) martingale whose increments have moments of order at least $2s$ ($s \geqq 1$). Then there is a continuous martingale $\overline{M}_t$, $t \geq 0$, on an extension of the probability space, so that $\overline{M}_t = M_t$ for all integer $t \geqq 0$, and so that the embedding has explicit structure up to order s. Furthermore, $\langle \overline{M}, \overline{M}\rangle_t$, $t = 0, 1, 2, \ldots$ has conditionally independent increments given the data.*

4 Relationship to some other embeddings

A number of the constructions in the literature use dichotomous transitions. Martingale increments are either embedded as a conditionally binary random variable (Skorokhod, 1965), or as a succession of such variables (Sawyer, 1967; Chacon and Walsh, 1976; Azéma and Yor, 1979; cf. Meilijson, 1983). Our procedure also uses multistage dichotomous increments, but unlike the Chacon–Walsh family, the two values are random.

It is possible to use (3.5)–(3.6) to say something about the (non-sequential and sequential) dichotomous procedures. If M_1 is the object to be embedded, and the transitions happen at countable times $t \in (0, 1)$ with associated σ-fields $\mathscr{F}_t$, then the mean and variance given $\mathscr{F}_1$ is on the form (3.5)–(3.6). Since binary martingales are extremal, a simple argument involving Bartlett identities (à la the one that led to (3.6)) shows that the conditional cumulants are unique, so one can set $c = 0$ in (3.6). This yields a formula for the conditional cumulants.

Obviously, however, the predictable variations are now no longer formed with respect to the original filtration, so the sequentially dichotomous procedure does not yield explicit structure in general. We have not investigated whether exceptions can occur.

Finally, it is worth observing that embeddings with explicit structure do not, in general, have the minimum variance property of Rost (1976). It is easy to see that (3.5)–(3.6) imply that $\mathrm{var}(\tau_t)$ is independent of c, and that (in the general case) there are several embeddings corresponding to this variance (cf. Theorem 1). This result then follows from the Corollary on p. 203 of Rost (1976).

5 Asymptotic expansions

The existing results on asymptotic expansions for martingales (Mykland, 1992, 1993, 1995) are only concerned with one-term expansions. With the help of the current embedding, however, these results can be generalized substantially.

The results concern the asymptotic behavior of $Eg(M_t)$, where M_t is a martingale and g is a sufficiently smooth function. Consider first the time- and samplepath-continuous case.

Theorem 4 (Expansion for continuous martingales) *Let W_t be a (triangular array of) Brownian motion(s), and let τ be stopping times. Then, if $\tau = \mu_2 + o_p(1)$, for nonrandom μ_2,*

$$Eg(W_\tau) = Eg(N(0,\mu_2)) + \sum_{j=1}^{k}(-1)^{j-1}\frac{1}{j!2^j}E[\tau-\mu_2]^j g^{(2j)}(W_\tau) + \Delta\,, \quad (5.1)$$

where

$$|\Delta| \leqq E\left[\sup_s |g^{(2k)}(W_s) - g^{(2k)}(W_\tau)||\tau-\mu_2|^k\right],$$

the supremum being over s between μ_2 and τ.

The proof is quite straightforward.

Proof of Theorem 4. Write

$$u_k = E\int_{\mu_2}^{\tau}(s-\mu_2)^{k-1}g^{(2k)}(W_s)\,ds\,,$$

$$v_k = E(\tau-\mu_2)^k g^{(2k)}(W_\tau)$$

and

$$f(t,x) = (t-\mu_2)^k g^{(2k)}(x)\,.$$

As in Lemma 5.1 of Mykland (1992), one can use Ito's formula on $f(\tau, W_\tau) - f(\mu_2, W_{\mu_2})$ to get that

$$v_0 = Eg(N(0,b^2)) + \tfrac{1}{2}u_1$$

and

$$v_k = ku_k + \tfrac{1}{2}u_{k+1} \quad \text{for } k > 0\,.$$

Solving this recurrence relation gives the result of the lemma. □

Clearly, if one has an embedding with explicit structure, one can use Theorem 4 to find an expansion of $Eg(M_t)$ in terms of optional and cumulant variations of (M_t). For example, if the martingale is quasi-left continuous, one can now use (3.7) in conjunction with (5.1).

For general martingales, with our knowledge of the first two conditional moments, we can get a second order expansion. Set

$$(M,N)_t = \tfrac{1}{3}[M,N]_t + \tfrac{2}{3}\langle M,N\rangle_t\,. \tag{5.2}$$

Using (3.5)–(3.6) yields that, for a general martingale, and for c satisfying (3.8),

$$\begin{aligned} Eg(M_t) &= Eg(N(0,\mu_2)) + \tfrac{1}{2}E((M,M)_t - \mu_2)g''(M_t) \\ &\quad - \tfrac{1}{8}E\left\{((M,M)_t - \mu_2)^2 + \tfrac{8}{45}\kappa_4(M)_t \right. \\ &\qquad - c[M,\kappa_3(M)]_t + \left(\tfrac{44}{45} - c\right)[\langle M,M\rangle, \langle M,M\rangle]_t \\ &\qquad \left. + c[M,M,\langle M,M\rangle]_t + \tfrac{2}{45}[M,M,M,M]_t\right\} g^{(iv)}(M_t) \\ &\quad + O(EB([M,M]_t - \mu_2)^2 + EB[M,M,M,M]_t) \end{aligned} \tag{5.3}$$

where B is bounded and $o_p(1)$. This is provided g has four continuous derivatives, with $g^{(iv)}$ bounded, and

$$E([M,M]_t - \mu_2)^2 + E[M,M,M,M]_t = o(1)\,. \tag{5.4}$$

We have here used Lenglart's inequality (see, e.g., p. 35 of Jacod and Shiryaev (1987)). Also, in (5.4) and the error term in (5.3), we have used that $\mathrm{var}([M,M]_t - \langle M,M\rangle_t) = E[M,M,M,M]_t - E[\langle M,M\rangle, \langle M,M\rangle]_t$ and that $[M,\kappa_3(M)]_t$, $[\langle M,M\rangle, \langle M,M\rangle]_t$ and $[M,M,\langle M,M\rangle_t]$ are all bounded by $[M,M,M,M]_t + \langle M,M,M,M\rangle_t$.

By combining (5.3) for two different c's, the c-terms can be made to vanish. This yields the following result.

Theorem 5 (Second order expansion for general martingales) *Let M_t be a (triangular array of) martingale(s), and suppose that* (5.4) *holds. Then, for 4 times continuously differentiable gs, with $g^{(iv)}$ bounded,*

$$\begin{aligned} Eg(M_t) &= Eg(N(0,\mu_2)) + \tfrac{1}{2}E((M,M)_t - \mu_2)g''(M_t) \\ &\quad - \tfrac{1}{8}E\left\{((M,M)_t - \mu_2)^2 + \tfrac{8}{45}\kappa_4(M)_t \right. \\ &\qquad \left. + \tfrac{2}{45}[M,M,M,M]_t + \tfrac{4}{9}[\langle M,M\rangle, \langle M,M\rangle]_t\right\} g^{(iv)}(M_t) \\ &\quad + O(EB([M,M]_t - \mu_2)^2 + EB[M,M,M,M]_t)\,. \end{aligned} \tag{5.5}$$

To turn this into an ordinary Edgeworth type expansion, one can now use a suitable central limit theorem on all terms except $E((M,M)_t - \mu_2)g''(M_t)$. A martingale approach to this would be to approximate $(M,M)_t - \mu_2$ by a (triangular array of) martingale(s). One can then use the following theorem. The proof is given in Sect. 9, and has some independent interest.

Theorem 6 (Expansion for $EN_t g(M_t)$) *Let (M_t) and (N_t) be triangular arrays of martingales. Assume* (5.4) *and also that $E\sup|\Delta N_t| = o(1)$ and $[N,M]_t =$*

$v + o_p(1)$, where v is nonrandom. Then, for thrice continuously differentiable gs, with $g^{(v)}$ bounded for $v = 0, \ldots, 3$,

$$EN_t g(M_t) = Eg'(M_t)(N,M)_t - \tfrac{1}{2}Eg''(M_t)((M,M)_t - \mu_2)E(N,M)_t + \tfrac{1}{2}Eg''(M_t)N_t((M,M)_t - \mu_2) + O((E[N,N]_t)^{1/2}(EB((M,M)_t - \mu_2)^2)^{1/2}), \tag{5.6}$$

where B is bounded and $o_p(1)$. □

6 Using the expansion results

The expansion results will mostly be relevant in cases where sums are of $O_p(t)$, so we shall in the following make the normalization explicit: $M_t/\sqrt{t}$ is asymptotically normal, and so on. This also makes it possible to state the order of convergence in terms of t.

As indicated above, our main assumption is that

$$[M,M]_t - \mu_2 = N_t + R, \tag{6.1}$$

where (N_t) is a martingale or a triangular array of martingales, and R is a remainder term of order $O_p(1)$.

Theorem 7 (Standard edgeworth expansion) *Assume* (6.1). *Also assume that*

$$t^{-1/2}([\ldots]_t - E[\ldots]_t)$$

is uniformly integrable and converges to a normal limit jointly with $M_t/\sqrt{t}$, where $[\ldots]_t$ is $(N,M)_t$, $[M,M,M,M]_t$, $[\langle M,M\rangle, \langle M,M\rangle]_t$ and $\kappa(M,M,M,M)_t$ (or $\langle M,M,M,M\rangle_t$). Also suppose that $[N,N]_t/t$ is uniformly integrable, and that $E\sup|\Delta N_t| = o(t^{1/2})$. Finally, let R be uniformly integrable and asymptotically independent of $M_t/\sqrt{t}$. Set $\sigma_t^2 = E[M,M]_t$, and suppose that σ_t^2 is not $o(t)$. Then

$$Eg(M_t/\sigma_t) = \int g(x)\phi(x)\lambda(x)\,dx + o(t^{-1})$$

for all functions g that are bounded and have five bounded continuous derivatives, where

$$\lambda(z) = 1 - \tfrac{1}{6}\rho_3 h_3(z) + \tfrac{1}{24}\rho_4 h_4(z) + \tfrac{1}{72}\rho_3^2 h_6(z).$$

h_3, h_4 and h_6 are the relevant Hermite polynomials, and ρ_3 and ρ_4 are the standardized cumulants, given by $\rho_k = \mathrm{cum}_k(M_t/\sigma_t)$.

We emphasize that the theorem is valid for triangular arrays. If one replaces $[M,M]_t$ by $(M,M)_t$ in (6.1), the result is a straightforward consequence of Theorems 5 and 6 above, and of the Bartlett identities for martingales (Mykland, 1994). Since $[M,M]_t - \langle M,M\rangle_t$ is a martingale satisfying the requirements imposed on N_t in the theorem, the stated result follows.

The condition (6.1) is satisfied in a number of standard situations. For example, if we are in discrete time and ΔM_n is a stationary and mixing sequence of martingale increments, we can set

$$\Delta N_n = \sum_{k=n}^{\infty} [E(\Delta M_k^2 | \mathscr{F}_n) - E(\Delta M_k^2 | \mathscr{F}_{n-1})], \tag{6.2}$$

cf. Chap. 5 of Hall and Heyde (1980). The same approach works with Markov processes, cf. pp. 447–448 of Jacod and Shiryaev (1987).

7 Bartlett identities proofs

Here we show formulas (3.5) and (3.6), and then Theorem 2. Theorem 1 is shown along with the proof of Theorem 3 in the next section.
To see (3.5), assume explicit structure:

$$E(\tau_t | \mathscr{F}) = c_1 [M,M]_t + c_2 \langle M,M \rangle_t . \tag{7.1}$$

The second identity in (3.4) implies $c_1 + c_2 = 1$. The third identity yields

$$\kappa_3(M)_t - 3c_1 \kappa(M,[M,M])_t - 3c_2 \kappa(M, \langle M,M \rangle)_t = 0 . \tag{7.2}$$

Since $\kappa(M, \langle M,M \rangle)_t = 0$ and $\kappa(M,[M,M])_t = \kappa(M,M,M)_t$, the result follows provided $\kappa_3(M)_t$ is not zero a.s.

Turning to (3.6), write

$$\begin{aligned} \operatorname{var}(\tau_t | \mathscr{F}) = & c_1 \kappa(M,M,M,M)_t + c_2 [M,M,M,M]_t \\ & + c_3 [\langle M,M \rangle, \langle M,M \rangle]_t + c_4 [M,M, \langle M,M \rangle]_t \\ & + c_5 [M, \kappa(M,M,M)]_t . \end{aligned} \tag{7.3}$$

Assuming (3.5), the fourth identity in (3.4) becomes

$$\kappa_4(M) - 2\kappa(M,M,[M,M])_t + \tfrac{1}{3}\kappa_2([M,M])_t + 3\kappa(\operatorname{var}(\tau.(\mathscr{F}))_t = 0 , \tag{7.4}$$

since $\kappa(M,M, \langle M,M \rangle)_t = 0$ and since $\kappa_2(E(\tau. | \mathscr{F}))_t = \kappa_2([M,M])_t / 9$. Furthermore,

$$\begin{aligned} \kappa(M,M,[M,M])_t &= \kappa_2([M,M])_t \\ &= \kappa_4(M)_t + 2[\langle M,M \rangle, \langle M,M \rangle]_t . \end{aligned} \tag{7.5}$$

Substituting (7.5) in (7.4) and comparing to (7.3) yields

$$c_1 + c_2 = \tfrac{2}{9} \quad \text{and} \quad -3c_1 + c_3 + c_4 = \tfrac{4}{9} , \tag{7.6}$$

unless $\kappa_4(M)_t$ and $[\langle M,M \rangle, \langle M,M \rangle]_t$ are linearly dependent. Similarly, inserting (3.5) into the fifth identity in (3.4) yields

$$\begin{aligned} \kappa_5(M)_t - \tfrac{10}{3}\kappa(M,M,M,[M,M])_t + \tfrac{15}{9}\kappa(M,[M,M],[M,M])_t \\ + 15\kappa(M, \operatorname{var}(\tau. | \mathscr{F}))_t = 0 . \end{aligned} \tag{7.7}$$

Since

$$\kappa_5(M)_t = \langle M,M,M,M,M\rangle_t - 10[\kappa_2(M),\kappa_3(M)]_t\,,$$
$$\kappa(M,M,M,[M,M])_t = \langle M,M,M,M,M\rangle_t - 4[\kappa_2(M),\kappa_3(M)]_t$$

and

$$\kappa(M,[M,M],[M,M])_t = \langle M,M,M,M,M\rangle_t - 2[\kappa_2(M),\kappa_3(M)]_t\,,$$

and since, by (7.3),

$$\kappa(M,\mathrm{var}(\tau_.|\mathscr{F}))_t = c_2\langle M,M,M,M,M\rangle_t + (c_4+c_5)[\kappa_2(M),\kappa_3(M)]_t\,, \quad (7.8)$$

(7.7) yields

$$c_2 = \tfrac{2}{45} \quad \text{and} \quad c_4 + c_5 = 0\,, \quad (7.9)$$

provided $\langle M,M,M,M,M\rangle_t$ and $[\kappa_2(M),\kappa_3(M)]_t$ are not zero a.s. Setting $c_4 = c$ and solving (7.6) and (7.9) yields (3.6).

The above assumes that there is no (nonrandom) linear relation between the variations used. If there are linear relations, the above shows only that (3.6) is one of several possible representations. □

We now turn to the proof of the Bartlett identities.

Proof of Theorem 2. We begin with the unconditional identities. Set $\tau_t^N = \tau_t \wedge N$, $M_t^{(N)} = W_{\tau_t^N}$ and $\mathscr{F}^{(N)} = \sigma(\mathscr{F}, M^{(n)}, n \geqq N)$. Obviously, the identities hold for finite N. As $N \to \infty$, for $q \leqq p$, $\mathrm{cum}_q(\tau_t^N|\mathscr{F}^{(N)}) \xrightarrow{P} \mathrm{cum}_q(\tau_t|\mathscr{F})$ by the reverse martingale limit theorem. Also, if $q_1 + 2q_2 + \cdots + kq_k = q$, then $|M_t^{(N)}|^{q_1}\prod_{i=2}^k|\mathrm{cum}_{q_i}(\tau_t^N|\mathscr{F}^{(N)})|$ is dominated by a constant times $|M_t^{(N)}|^{q_1}E((\tau_t^N)^{(q-q_1)/2}|\mathscr{F}^{(N)})$. Hence, by letting $N \to \infty$, the unconditional identities remain valid.

In the conditional case ((3.3) and so on), the rigorous proof is more convoluted, and the likelihoodization is more of a heuristic (though it does give the right answer). The rigorous argument begins with the "local" likelihood

$$\sum \exp(\theta(W_{\tau_{t_{i+1}}} - W_{\tau_{t_i}}) - \tfrac{1}{2}\sigma^2(\tau_{t_{i+1}} - \tau_{t_i})) \quad (7.10)$$

by appropriate stopping and then differentiating, and then letting $\max(t_{i+1} - t_i)$ go to zero, one gets Bartlett identities on the form that for $q \leqq p$,

$$\sum \frac{q!(-1)^{q_2}}{q_1!2^{q_2}q_2!}[\underbrace{M,\ldots,M}_{q_1 \text{ times}},\underbrace{\tau,\ldots,\tau}_{q_2 \text{ times}}]_t \quad (7.11)$$

is a martingale. Hence,

$$\sum \frac{q!(-1)^{q_2}}{q_1!2^{q_2}q_2!}[M,\ldots,M,\, E([\tau,\ldots,\tau]_.|\mathscr{F})]_t \quad (7.12)$$

is a martingale with respect to the original filtration ($E([\tau,\dots,\tau]_t|\mathscr{F})$ being càdlàg and adapted by (7.14) below). It follows that

$$\sum \frac{q!(-1)^{q_2}}{q_1!2^{q_2}q_2!}\langle M,\dots,M, E([\tau,\dots,\tau].|\mathscr{F})\rangle_t = 0\,. \tag{7.13}$$

On the other hand, by the conditional independence assumption and the definition of cumulants,

$$E(\underbrace{[\tau,\dots,\tau]_t}_{r\ \text{times}}|\mathscr{F}) = \sum \frac{r!}{\Pi(k!)^{r_k}r_k!}[\dots,\underbrace{\mathrm{cum}_k(\tau.|\mathscr{F}),\dots,\mathrm{cum}_k(\tau.|\mathscr{F})}_{r_k\ \text{times}},\dots]_t\,. \tag{7.14}$$

Combining (7.13), (7.14) and (3.2) now yields (3.11), keeping in mind that $r = q$ and that $r_k = q_{2k}$ in the notation of (3.11). (3.12) then follows as outlined in Exercise 7.1 (p. 222) in McCullagh (1987). □

8 Proof of the existence of embedding with explicit structure

Proof of Theorems 1 and 3. With the exception of (3.8), Theorem 1 is an obvious consequence of Theorem 3. This can be seen by inspecting the proof of Theorem 11 (pp. 1300– 1301) of Monroe (1972). For Theorem 3, it is enough to show that the result holds for $t = 1$ and with $\mathscr{F}_0$ as the trivial σ-field.

The construction is as follows. Set $\widetilde{M}_1 = M_1$, and then iteratively, for $n = 0,1,2,\dots$, let $\widetilde{M}'_{2^{-n}}$ be an independent copy of $\widetilde{M}_{2^{-n}}$, let $(\widetilde{M}^{(1)}_{2^{-n}}, \widetilde{M}^{(2)}_{2^{-n}})$ be drawn at random from $(\widetilde{M}_{2^{-n}},\widetilde{M}'_{2^{-n}})$ and $(\widetilde{M}'_{2^{-n}}, \widetilde{M}_{2^{-n}})$ with probability p and $1-p$, respectively $(0 < p < 1)$, independently of everything else in sight, and set

$$\widetilde{M}_{2^{-(n+1)}} = p\widetilde{M}^{(1)}_{2^{-n}} + (1-p)\widetilde{M}^{(2)}_{2^{-n}}\,.$$

The accompanying filtration is given by $\widetilde{\mathscr{F}}_{2^{-(n+1)}} = \sigma(\widetilde{M}^{(1)}_{2^{-n}},\widetilde{M}^{(2)}_{2^{-n}},\widetilde{M}_{2^{-k}},\widetilde{M}'_{2^{-k}}, k \geqq n+1)$ and $\widetilde{\mathscr{F}}_1 = \sigma(\widetilde{M}_{2^{-k}},\widetilde{M}'_{2^{-k}}, k \geqq 0)$.

Since

$$\begin{aligned} E(\widetilde{M}_{2^{-n}}|\widetilde{\mathscr{F}}_{2(-n+1)}) &= \sum_{i=1}^{2} E(\widetilde{M}_{2^{-n}}I\{\widetilde{M}_{2^{-n}} = \widetilde{M}^{(i)}_{2^{-n}}\}|\widetilde{\mathscr{F}}_{2^{-(n+1)}}) \\ &= \widetilde{M}_{2^{-(n+1)}}\,, \end{aligned}$$

$\widetilde{M}_{2^{-n}}$ is a martingale with respect to the filtration $\widetilde{\mathscr{F}}_{2^{-n}}$. Since

$$\widetilde{M}^{(1)}_{2^{-n}} - \widetilde{M}_{2^{-(n+1)}} = (1-p)(\widetilde{M}^{(1)}_{2^{-n}} - \widetilde{M}^{(2)}_{2^{-n}})$$

and

$$\widetilde{M}^{(2)}_{2^{-n}} - \widetilde{M}_{2^{-(n+1)}} = -p(\widetilde{M}^{(1)}_{2^{-n}} - \widetilde{M}^{(2)}_{2^{-n}})\,,$$

it follows that a continuous martingale $\overline{M}_t$ can be created by setting, for $2^{-(n+1)} \leqq t < 2^{-n}$,

$$\overline{M}_t = \widetilde{M}_{2^{-(n+1)}} + (\widetilde{M}^{(1)}_{2^{-n}} - \widetilde{M}^{(2)}_{2^{-n}})L^{(n)}_{2^{n+1}t-1}\,,$$

where the $L_1^{(n)}$ are embeddings of the two point martingale which takes values $(1-p)$ and $-p$. Note that $L_1^{(n)}$ is dependent on whether $\widetilde{M}_{2^{-n}}^{(1)}$ or $\widetilde{M}_{2^{-n}}^{(2)}$ is, indeed, $\widetilde{M}_{2^{-n}}$.

Clearly, $\overline{M}_t$ is a continuous martingale, with $\overline{M}_1 = M_1$. Also, $\overline{M}_0 = 0$ since

$$\begin{aligned}\operatorname{var}(\widetilde{M}^{\alpha}_{2^{-(n+1)}}) &= (p^2+(1-p)^2)\operatorname{var}(\widetilde{M}^{\alpha}_{2^{-n}})\\ &= (p^2+(1-p)^2)^{n+1}\operatorname{var}(M_1^{\alpha})\\ &\to 0\end{aligned}$$

as $n \to \infty$, if $0 < p < 1$.

Since

$$\begin{aligned}\langle \overline{M}^{\alpha}, \overline{M}^{\beta}\rangle_1 &= \sum_{n=0}^{\infty} (\widetilde{M}^{(1)\alpha}_{2^{-n}} - \widetilde{M}^{(2)\alpha}_{2^{-n}})(\widetilde{M}^{(1)\alpha}_{2^{-n}} - \widetilde{M}^{(2)\alpha}_{2^{-n}})\langle L^{(n)}\rangle_1\\ &= \sum_{n=0}^{\infty} (\widetilde{M}^{\alpha}_{2^{-n}} - \widetilde{M}'^{\alpha}_{2^{-n}})(\widetilde{M}^{\beta}_{2^{-n}} - \widetilde{M}'^{\beta}_{2^{-n}})\langle L^{(n)}\rangle_1\,,\end{aligned}$$

by embedding of two point martingales, it follows that $\operatorname{cum}(\langle \overline{M}^{\alpha_{11}}, \overline{M}^{\alpha_{12}}\rangle_1, \ldots, \langle \overline{M}^{\alpha_{q1}}, \overline{M}^{\alpha_{q2}}\rangle_1 | \widetilde{\mathscr{F}}_1)$ has explicit structure with respect to $\widetilde{M}_t$. It further follows by the iterated conditioning formula for cumulants (cf. Brillinger, 1969; Speed, 1983) that the embedding has explicit structure. It remains to pin down the coefficients of the conditional variance to show (3.8). Let M_t be scalar, and set

$$\begin{aligned}\operatorname{var}(\langle \overline{M}\rangle_1 | M_1) = {} & a_0 \operatorname{cum}_4(M_1) + b_0 M_1^4\\ & + c_0 \operatorname{var}(M_1)^2 + d_0 M_1^2 \operatorname{var}(M_1) + e_0 M_1 \operatorname{cum}_3(M_1)\,.\end{aligned}$$

The way the construction is carried out, however, it is also true for $n \geqq 1$ that

$$\begin{aligned}\operatorname{var}(\langle \overline{M}\rangle_{2^{-n}} | \mathscr{G}_n) = {} & a_n \operatorname{cum}_4(\widetilde{M}_{2^{-n}}) + b_n \widetilde{M}^4_{2^{-n}}\\ & + c_n \operatorname{var}(\widetilde{M}_{2^{-n}})^2 + d_n \widetilde{M}_{2^{-n}} \operatorname{var}(\widetilde{M}_{2^{-n}})\\ & + e_n \widetilde{M}_{2^{-n}} \operatorname{cum}_3(\widetilde{M}_{2^{-n}})\,,\end{aligned}$$

where $\mathscr{G}_n = \sigma(\widetilde{M}_{2^{-k}}, \widetilde{M}^{(1)}_{2^{-k}}, \widetilde{M}^{(2)}_{2^{-k}}, \widetilde{M}_{2^{-k-1}}, k \leqq n-1)$. A tedious but straightforward calculation then shows that $\operatorname{var}(\langle \overline{M}\rangle_{2^{-n+1}} | \mathscr{G}_{n-1})$ can be represented as the sum of terms given by Table 1,

Table 1. $\operatorname{var}(\langle \overline{M}\rangle_{2^{-n+1}} | \mathscr{G}_{n-1})$

	1	$p(1-p)$	$p^2(1-p)^2$
$\operatorname{cum}_4(\widetilde{M}_{2^{-n+1}})$	a_n	$-4a_n + b_n + \frac{2}{3}$	0
$\widetilde{M}^4_{2^{-n+1}}$	b_n	$-5b_n + \frac{2}{9}$	$5b_n - \frac{2}{9}$
$\operatorname{var}(\widetilde{M}_{2^{-n+1}})^2$	c_n	$3b_n + 4c_n + d_n + 2$	$-9b_n + 4c_n - 2d_n - \frac{22}{9}$
$\widetilde{M}^2_{2^{-n+1}} \operatorname{var}(\widetilde{M}_{2^{-n+1}})$	d_n	$-5d_n + \frac{16}{9}$	$6b_n + 6d_n - \frac{4}{3}$
$\widetilde{M}_{2^{-n+1}} \operatorname{cum}_3(\widetilde{M}_{2^{-n+1}})$	e_n	$-5e_n - \frac{16}{9}$	$4b_n + 6e_n + \frac{8}{9}$

where, for example, entry (2,3) says that the coefficient in front of $\widetilde{M}^4_{2^{-n+1}} p^2 (1-p)^2$ is $5b_n - \frac{2}{9}$.

Assume first that $E|M_1|^5 < \infty$. Then we know that the representation (3.6) must hold, i.e.

$$\begin{aligned}\mathrm{var}(\langle \widetilde{M}\rangle_{2^{-n}}|\mathscr{G}_n) = {} & \tfrac{8}{45}\mathrm{cum}_4(\widetilde{M}_{2^{-n}}) + \tfrac{2}{45}\widetilde{M}^4_{2^{-n}} + \tfrac{44}{45}\mathrm{var}(\widetilde{M}_{2^{-n}})^2 \\ & + d_n(-\mathrm{var}(\widetilde{M}_{2^{-n}})^2 + \widetilde{M}^2_{2^{-n}}\mathrm{var}(\widetilde{M}_{2^{-n}}) - M_{2^{-n}}\mathrm{cum}_3(\widetilde{M}_{2^{-n}})).\end{aligned}$$

By the central limit theorem, $\widetilde{M}_{2^{-n}}/\sigma(p^2+(1-p)^2)^{n/2} \xrightarrow{\mathscr{L}} U$, where U is $N(0,1)$ and $\sigma^2 = \mathrm{var}(M_1)$. Also, obviously, $E\widetilde{M}^4_{2^{-n}}/(p^2+(1-p)^2)^{2n} \to 3\sigma^4$, and hence $\mathrm{var}(\langle \widetilde{M}\rangle_{2^{-n}}|\mathscr{G}_n)/(p^2+(1-p)^2)^{2n}$ is $O_p(1)$ in view of Burkholder's inequality. Hence, one can write $\alpha_n = d_n\beta_n$, where $\alpha_n = O_p(1)$ and $\beta_n \xrightarrow{\mathscr{L}} \sigma^4(U^2-1)$. Hence $d_n = O(1)$.

One can now use Table 1 and the fact that $b_n = \frac{2}{45}$ to set up the recursion formulas (d_n is the same as c in formula (3.6))

$$d_{n-1} = d_n(1-5p(1-p)+6p^2(1-p)^2) + \tfrac{16}{9}p(1-p) - \tfrac{16}{15}p^2(1-p)^2$$

from which one gets

$$d_n = \frac{d_0-\gamma}{\alpha^n} + \gamma\,,$$

where $\gamma = (\frac{16}{9}p(1-p) - \frac{16}{15}p^2(1-p)^2)/(1-\alpha)$ and $\alpha = 1-5p(1-p)+6p^2(1-p)^2$. Since $0<p<1$, it follows that $|\alpha|<1$. Since $d_n = O(1)$, one must have $d_0 = \gamma$, i.e.

$$d_0 = \frac{16}{45}\frac{5-3p(1-p)}{5-6p(1-p)}\,.$$

This specifies the formula for $\mathrm{var}(\langle \widetilde{M}\rangle_1|M_1)$, and a simple limiting argument shows that one can reduce the moment requirements to $EM_1^4 < \infty$. By letting p vary in $(0,1)$, one gets (3.8). □

9 The $EN_tg(M_t)$ results

These are the results which permit us to turn the expansions (5.1) and (5.5) into ordinary Edgeworth expansions. We first present an exact result. Define

$$A = \inf\{t\colon \langle M,M\rangle_t = \mu_2\}\,. \tag{9.1}$$

Theorem 8 *Let (M_t) and (N_t) be continuous martingales. Provided the relevant expectations exist, the following is true: For all $g \in \mathscr{C}_3$,*

$$\begin{aligned}EN_tg(M_t) = {} & Eg'(M_t)\langle N,M\rangle_t - \tfrac{1}{2}E\int_A^t g'''(M_s)\langle M,N\rangle_s\,d\langle M,M\rangle_s \\ & + \tfrac{1}{2}E\int_A^t g''(M_s)N_s d\langle M,M\rangle_s\end{aligned} \tag{9.2}$$

and for all $g \in \mathscr{C}_1$,

$$EN_A g(M_A) = Eg'(M_A)\langle N,M\rangle_A\,, \tag{9.3}$$

$$EN_{t\wedge A} g(M_A) = Eg'(M_A)\langle N,M\rangle_{A\wedge t} \tag{9.4}$$

and

$$EN_t g(M_A) = Eg'(M_A)\langle N,M\rangle_{A\wedge t}\,. \tag{9.5}$$

Proof of Theorem 8. Let $h \geqq 0$, and consider the martingale

$$M_t^h = M_t + hN_t\,. \tag{9.6}$$

Set

$$A(h) = \inf\{t\colon \langle M^h, M^h\rangle_t = \mu_2\}\,. \tag{9.7}$$

Suppose g is a function with four bounded and continuous derivatives. By Ito's lemma,

$$Eg\left(M_t^h\right) = Eg(N(0,\mu_2)) + \tfrac{1}{2}\int_{A(h)}^{t} g''(M_s^h)d\langle M^h, M^h\rangle_s\,. \tag{9.8}$$

Differentiating with respect to h and setting $h = 0$ now yields

$$EN_t g'(M_t) = Eg''(M_A)\langle M,N\rangle_A + E\int_A^t g''(M_s)\,d\langle M,N\rangle_s$$

$$+ \tfrac{1}{2}E\int_A^t g'''(M_s)N_s\,d\langle M,M\rangle_s\,. \tag{9.9}$$

Note that this differentiation presupposes that $\langle M,M\rangle_t$, $\langle M,N\rangle_t$ and $\langle N,N\rangle_t$ are continuously differentiable. If this is not the case, the result is shown first for M_t's and N_t's satisfying this condition, and the result for general M_t and N_t then follows by taking limits. Using Ito's formula again yields that

$$Eg''(M_t)\langle M,N\rangle_t = Eg''(M_A)\langle M,N\rangle_A + E\int_A^t g''(M_s)\,d\langle M,N\rangle_s$$

$$+ \tfrac{1}{2}E\int_A^t g^{(iv)}(M_s)\langle M,N\rangle_s\,d\langle M,M\rangle_s\,. \tag{9.10}$$

Merging (9.9) and (9.10) yields that

$$EN_t g'(M_t) = Eg''(M_t)\langle M,N\rangle_t - \tfrac{1}{2}E\int_A^t g^{(iv)}(M_s)\langle M,N\rangle_s\,d\langle M,M\rangle_s$$

$$+ \tfrac{1}{2}E\int_A^t g'''(M_s)\,N_s\,d\langle M,M\rangle_s\,. \tag{9.11}$$

Replacing g' by g and taking limits yields the result (9.2).

On the other hand, another application of Ito's formula yields

$$EN_t g'(M_t) = EN_A g'(M_A) + E\int_A^t g''(M_s)\, d\langle M,N\rangle_s + \tfrac{1}{2}E\int_A^t g'''(M_s) N_s\, d\langle M,M\rangle_s \,. \tag{9.12}$$

Merging (9.9) and (9.12) now yields

$$EN_A g'(M_A) = Eg''(M_A)\langle M,N\rangle_A \,. \tag{9.13}$$

Again, replacing g' by g and taking limits yields (9.3).

Finally, let $N_s = N_t$ for $s \geqq t$. Then, (9.3) reduces to (9.4). Since (9.4) only uses properties of N_s up to $t \wedge A$, this then holds for general (N_s). The martingale property of N_t then yields (9.5). □

Proof of Theorem 6. Assume first that (M_t) and (N_t) are continuous, and that $\langle M,M\rangle_t \leqq \mu_2 + 1$. Then, in view of (5.4), (9.2) and the Mean Value Theorem, (5.6) holds. If $\langle M,M\rangle_t$ is not bounded, note that, for example,

$$\begin{aligned} &|Eg'(M_t)(\langle N,M\rangle_t - \langle N,M\rangle_{t\wedge\tau})I(\tau < t)| \\ &\leqq \sup_x |g'(x)|(E\langle N,N\rangle_t)^{1/2}(E(\langle M,M\rangle_t - \mu_2)^2 I(\tau < t))^{1/2} \end{aligned}$$

by the Kunita–Watanabe and Hölder inequalities, where $\tau = \inf\{s\colon \langle M,M\rangle_s > \mu_2 + 1\}$. By using two interpolations and making a linear combination to get $c = 0$, the result now follows for discrete time and hence (by taking limits, since the order of convergence is really a small sample bound) general martingales. This is provided we can show that the assumption of the theorem implies that (when embedding in continuous time martingales $\overline{M}$ and $\overline{N}$), $\langle \overline{N},\overline{M}\rangle_t = v + o_p(1)$. However, for discrete time martingales

$$[N,M]_t - \langle \overline{N},\overline{M}\rangle_t = \sum_i \int_{t_i}^{t_{i+1}} (\overline{M}_s - \overline{M}_{t_i})\, d\overline{N}_s + \sum_i \int_{t_i}^{t_{i+1}} (\overline{N}_s - \overline{N}_{t_i})\, d\overline{M}_s$$

which, by Lenglart's inequality (see Jacod and Shiryaev, 1987, p. 35) is $o_p(1)$ provided

$$\sum_i \int_{t_i}^{t_{i+1}} (\overline{M}_s - \overline{M}_{t_i})^2\, d\langle \overline{N}\rangle_s + \sum_i \int_{t_i}^{t_{i+1}} (\overline{N}_s - \overline{N}_{t_i})^2\, d\langle \overline{M}\rangle_s$$

is $o_p(1)$. This is true by the same reasoning as in Mykland (1994), since $E\sup|\Delta M_s| = o(1)$ and $E\sup|\Delta N_s| = o(1)$. The latter is assumed, the former follows from (5.4) since

$$\left(\sup_{0\leqq s\leqq t} |\Delta M_s|\right)^4 \leqq [M,M,M,M]_t \,. \qquad \square$$

Acknowledgements. The author wishes to thank Peter Hall, Peter McCullagh, Isaac Meilijson, Jim Pitman and an anonymous referee for useful discussions, and Mitzi Nakatsuka for typing the manuscript.

References

Abramowitz, M., Stegun, I.A.: Handbook of mathematical functions, with formulas, graphs and mathematical tables. Washington, DC: National Bureau of Standards 1972

Azéma, J., Yor, M.: Une solution simple au problème de Skorokhod. Sem. Probab. XIII. (Lect. Notes Math., Vol. 721, pp. 90–115 and 625–633) Berlin: Springer 1979

Bartlett, M.S.: Approximate confidence intervals. Biometrika **40,** 12–19 (1953)

Bartlett, M.S.: Approximate confidence intervals. II. More than one unknown parameter. Biometrika **40,** 306–317 (1953b)

Brillinger, D.: The calculation of cumulants via conditioning. Ann. Inst. Statist. Math. **21,** 375–390 (1969)

Chacon, R.V., Walsh, J.B.: Une solution simple au problème de Skorokhod. Sem. Probab. X. (Lect. Notes Math., Vol. 511, pp. 19–23) Berlin: Springer 1976

Clark, J.M.C.: The representation of functionals of Brownian motion by stochastic integrals. Ann. Math. Statist. **41,** 1282–1295 (1970)

Dambis, K.E.: On the decomposition of continuous submartingales. Theoret. Probab. Appl. **10,** 401–410 (1965)

Dubins, L.E.: On a theorem of Skorokhod. Ann. Math. Statist. **39,** 2094–2097 (1968)

Dubins, L.E. Schwartz, G.: On continuous martingales. Proc. Nat. Acad. Sci. **53,** 913–916 (1965)

Hall, P., Heyde, C.C.: Martingale limit theory and its application. New York: Academic Press 1980

Heath, D.: Interpolation of martingales. Ann. Probab. **5,** 804–806 (1977)

Jacod, J., Shiryaev, A.N.: Limit theorems for stochastic processes. Berlin: Springer 1987

Karatzas, I., Shreve, S.E.: Brownian motion and stochastic calculus. Berlin: Springer 1988

Khoshnevisan, D.: An embedding of compensated compound Poisson processes with applications to local times. Ann. Probab. **21,** 340–361 (1993)

McCullagh, P.: Tensor methods in statistics. London: Chapman and Hall 1987

Meilijson, I.: On the Azèma–Yor stopping time. Sem. Prob. XVII. (Lect. Notes Math., Vol. 986, pp. 225–226) Berlin: Springer 1983

Monroe, I.: On embedding right continuous martingales in Brownian motion. Ann. Math. Statist. **43,** 1293–1311 (1972)

Mykland, P.A.: Asymptotic expansions and bootstrapping distributions for dependent variables: a martingale approach. Ann. Statist. **20,** 623–654 (1992)

Mykland, P.A.: Asymptotic expansions for martingales. Ann. Probab. **21,** 800–818 (1993)

Mykland, P.A.: Bartlett type identities for martingales. Ann. Statist. **22,** 21–38 (1994)

Mykland, P.A.: Martingale expansions and second order inference. Ann. Statist. (to appear).

Mykland, P.A., Ye, J.: Cumulants and Bartlett identities in Cox regression. Tech. Report no. 322, Department of Statistics, University of Chicago 1992

Root, D.H.: On the existence of certain stopping times on Brownian motion. Ann. Math. Statist. **40,** 715–718 (1969)

Rost, H.: Skorohod stopping times of minimal variance. Sem. Probab. X. (Lect. Notes Math., Vol. 511, pp. 194–208) Berlin: Springer 1976

Sawyer, S.: A uniform rate of convergence for the maximum absolute value of partial sums in probability. Commun. Proc. Appl. Math. 20, 647–658 (1967)

Skorokhod, A.V.: Studies in the theory of random processes. Reading, MA: Addison-Wesley 1965

Speed, T.P.: Cumulants and partition lattices. Austral. J. Statist. **25,** 378–388 (1983)