Strong approximation of stochastic processes using random walks

Balázs Székely

Supervisor: Tamás Szabados

Institute of Mathematics
Budapest University of Technology and Economics
2004

Contents

1 Introduction **1**

2 Approximation of continuous martingales **3**

2.1 Random walks and the Wiener process . . . 3
2.2 Approximation of continuous martingales . . . 7
2.2.1 Approximation of the quadratic variation process . . . 8
2.2.2 Strong approximation of continuous martingales . . . 10
2.3 Symmetrically evolving martingales . . . 14
2.3.1 Distributional characterization of Ocone martingales . . . 14
2.3.2 Properties of Ocone martingales . . . 19
2.3.3 Some examples of Ocone and non-Ocone martingales . . . 21
2.4 Approximation of local time, excursion, meander and bridge . . . 25
2.4.1 Approximation of local time . . . 26
2.4.2 Excursion, bridge and meander – Marchal's construction . . . 34
2.5 Pathwise stochastic integration . . . 35
2.5.1 Karandikar's integral approximation . . . 35
2.5.2 Discretization applied to the integrator . . . 39
2.5.3 The non cadlag-case . . . 43

3 Approximation of the exponential functional of Brownian motion **46**

3.1 Introduction to the exp functional of Brownian motion . . . 46
3.2 The distribution of the discrete exponential functional . . . 47
3.2.1 Review of some fractal notion . . . 48
3.2.2 The non-overlapping case . . . 49
3.2.3 The general case . . . 50
3.3 The moments of the discrete exponential functional . . . 51
3.3.1 Permutations with given descent set . . . 52
3.4 Approximation of the exponential functional of Brownian motion . . . 56
3.4.1 The moments of the exponential functional of BM . . . 58
3.4.2 Properties of the exponential functional process . . . 61

Chapter 1

Introduction

The main concept of this thesis is to draw the attention to the fact that both for theoretical and practical reasons, it is useful to search for strong (i.e. pathwise, almost sure) approximations of stochastic processes by simple random walks (RWs). The prototype of such efforts was the construction of Brownian motion (BM) as an almost sure limit of simple RW paths, given by Frank Knight in 1962 [15]. Later this construction was simplified and somewhat improved by Pál Révész [22] and then by Tamás Szabados [28]. This sort of results states that one can find a sequence of time and space scaled random walks (B_m) such that it converges to a Brownian motion B almost surely uniformly on every compact intervals as m goes to infinity:

$$\sup_{t\in[0,T]} |B_m(t) - B(t)| \to 0.$$

Besides the theoretical value of discrete approximations, in some applications the discrete model could be more natural than the continuous one. It provides a general tool for proving statements for continuous time stochastic processes. First, one can prove the discrete version of the statement and then take limit for obtaining the continuous version. Of course, we cannot predict which part of this procedure will be easier.

In this thesis, we present both theoretical results, e.g. approximation of continuous martingales, and also show examples in which the discrete approximation method can be carried out naturally. During our research some additional statements turned out to be true which are not closely related to our present subject. However, we present some of them because we think that the reader may find them interesting.

This study consists of two main parts, Chapter 2 and Chapter 3. Chapter 2 is mainly based on the paper *Strong approximation of continuous local martingales by simple random walks* and some recent unpublished results. In Chapter 3, we present our results on the so-called exponential functional of Brownian motion which were discussed in the papers *An exponential functional of random walks* and *Moments of an exponential functional of random walks and permutations with given descent sets.*

Chapter 2 discusses a generalization of the result by Knight to continuous local martingales M. We will show that the quadratic variation process $\langle M, M\rangle$ can be almost surely uniformly approximated by a discrete quadratic variation processes N_m which are based on stopping times of a Skorohod-type embedding of nested simple RWs into M. This corresponds to an earlier similar result by Karandikar [12]. In section 2.5 we present some of his related results, for instance discrete approximation of certain stochastic integrals.

Theorems 2 and 3 give an approximation of M by a nested sequence of RWs B_m, time-changed by $\langle M, M\rangle$ and N_m, respectively. The approximations almost surely uniformly converge on bounded intervals.

It is important to note that the DDS Brownian motion W and the quadratic variation $\langle M, M\rangle$ are not independent in general, just like the approximating RW B_m and the discrete quadratic variation N_m. Since this could be a hindrance both in the theory and applications, a necessary and sufficient condition is given for the independence in Theorem 4. Namely, W and $\langle M, M\rangle$ are independent if and only if M has symmetric increments given the past. This is a reformulation of an earlier result by Ocone and Dubins-Émery-Yor. We also present some recent theorems on the properties of Ocone martingales which provide examples for Ocone martingales. This investigations on this special family of martingales can be found in Section 2.3.

Naturally arise the question if we can extend our discrete approximation method to approximate stochastic integrals with respect to continuous martingales. In Section 2.5 it turns out that the construction works for cadlag integrands. To get approximation result for a larger class of integrands, e.g. of the form $f'_-(M_t)$ where f'_- is the derivative of the difference of two convex functions, we prove that we can approximate the Brownian local time in the same manner as we do with martingales. This work is done in Section 2.4.

Chapter 3 focuses on a certain application of the discrete approximation of Brownian motion. Here, we investigate the exponential functional of Brownian motion

$$\mathcal{I}_\nu = \int_0^\infty \exp\left(B(t) - \nu t\right)\, \mathrm{d}t$$

and mainly its discrete version the exponential functional of random walk. The properties of the discrete exponential functional are rather different from the continuous one: typically its distribution is singular w.r.t. Lebesgue measure, all of its positive integer moments are finite and they characterize the distribution. On the other hand, using suitable random walk approximations to Brownian motion, the resulting discrete exponential functionals converge a.s. to the exponential functional of Brownian motion, hence their limit distribution is the same as in the continuous case, namely, the one of the reciprocal of a gamma random variable, so absolutely continuous w.r.t. Lebesgue measure. This way we give a new, elementary proof for an earlier result by Dufresne and Yor as well.

Beyond these results we have found a recursion of certain moments in the expansion of the moments of the discrete approximation.

In this thesis we denote by alphabetical numbering the known results and by arabic numbering the results of the authors.

Chapter 2

Approximation of continuous martingales

2.1 Random walks and the Wiener process

A main tool of this thesis is an elementary construction of the Wiener process (= BM). The specific construction we are going to use in the sequel, taken from [28], is based on a nested sequence of simple random walks that uniformly converges to the Wiener process on bounded intervals with probability 1. This will be called *RW construction* in the sequel. One of our intentions in this chapter is to extend the underlying *"twist and shrink"* algorithm to continuous local martingales.

We summarize the major steps of the RW construction here, see [29] as well. We start with an infinite matrix of i.i.d. random variables $X_m(k)$, $\mathbf{P}\{X_m(k) = \pm 1\} = 1/2$ ($m \geq 0$, $k \geq 1$), defined on the same underlying probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Each row of this matrix is a basis of an approximation of the Wiener process with a dyadic step size $\Delta t = 2^{-2m}$ in time and a corresponding step size $\Delta x = 2^{-m}$ in space, illustrated by the next table.

Table 2.1: The starting setting for the RW construction of BM

Δt	Δx	i.i.d. sequence	RW
1	1	$X_0(1), X_0(2), X_0(3), \ldots$	$S_0(n) = \sum_{k=1}^{n} X_0(k)$
2^{-2}	2^{-1}	$X_1(1), X_1(2), X_1(3), \ldots$	$S_1(n) = \sum_{k=1}^{n} X_1(k)$
2^{-4}	2^{-2}	$X_2(1), X_2(2), X_2(3), \ldots$	$S_2(n) = \sum_{k=1}^{n} X_2(k)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

The second step of the construction is *twisting*. From the independent random walks we want to create dependent ones so that after shrinking temporal and spatial step sizes, each consecutive RW becomes a refinement of the previous one. Since the spatial unit will be halved at each consecutive row, we define stopping times by $T_m(0) = 0$, and for $k \geq 0$,

$$T_m(k+1) = \min\{n : n > T_m(k), |S_m(n) - S_m(T_m(k))| = 2\} \qquad (m \geq 1) \tag{2.1}$$

These are the random time instants when a RW visits even integers, different from the previous one. After shrinking the spatial unit by half, a suitable modification of this RW will visit the same integers in the same order as the previous RW. We operate here on each point $\omega \in \Omega$ of the sample space separately, i.e. we fix a sample path of each RW. We define twisted RWs $\tilde{S}_m$ recursively for $k = 1, 2, \dots$ using $\tilde{S}_{m-1}$, starting with $\tilde{S}_0(n) = S_0(n)$ $(n \geq 0)$. With each fixed m we proceed for $k = 0, 1, 2, \dots$ successively, and for every n in the corresponding bridge, $T_m(k) < n \leq T_m(k+1)$. Any bridge is flipped if its sign differs from the desired:

$$\tilde{X}_m(n) = \begin{cases} X_m(n) & \text{if } S_m(T_m(k+1)) - S_m(T_m(k)) = 2\tilde{X}_{m-1}(k+1), \\ -X_m(n) & \text{otherwise,} \end{cases}$$

and then $\tilde{S}_m(n) = \tilde{S}_m(n-1) + \tilde{X}_m(n)$. Then $\tilde{S}_m(n)$ $(n \geq 0)$ is still a simple symmetric random walk [28, Lemma 1]. The twisted RWs have the desired refinement property:

$$\frac{1}{2}\tilde{S}_m(T_m(k)) = \tilde{S}_{m-1}(k) \qquad (m \geq 1, k \geq 0).$$

The last step of the RW construction is *shrinking*. The sample paths of $\tilde{S}_m(n)$ $(n \geq 0)$ can be extended to continuous functions by linear interpolation, this way one gets $\tilde{S}_m(t)$ $(t \geq 0)$ for real t. Then we define the *mth approximating RW* by

$$\tilde{B}_m(t) = 2^{-m}\tilde{S}_m(t2^{2m}).$$

Using the definition of T_m and $\tilde{B}_m$ we also get the general *refinement property*

$$\tilde{B}_{m+1}\left(T_{m+1}(k)2^{-2(m+1)}\right) = \tilde{B}_m\left(k2^{-2m}\right) \qquad (m \geq 0, k \geq 0). \tag{2.2}$$

Note that a refinement takes the same dyadic values in the same order as the previous shrunken walk, but there is a *time lag* in general:

$$T_{m+1}(k)2^{-2(m+1)} - k2^{-2m} \neq 0. \tag{2.3}$$

Then we quote some important facts from [28] about the above RW construction that will be used in the sequel. These will be stated in somewhat stronger forms but can be read easily from the proofs in the cited reference, cf. Lemmas 2-4 and Theorem 3 there.

Lemma A. *Suppose that $X_1, X_2, \dots, X_N$ is an i.i.d. sequence of random variables, $\mathbf{E}(X_k) = 0$, $\mathbf{Var}(X_k) = 1$, and their moment generating function is finite in a neighborhood of 0. Let $S_j = X_1 + \dots + X_j$, $1 \leq j \leq N$. Then for any $C > 1$ and $N \geq N_0(C)$ one has*

$$\mathbf{P}\left\{\sup_{1 \leq j \leq N} |S_j| \geq (2CN\log N)^{\frac{1}{2}}\right\} \leq 2N^{1-C}.$$

We mention that this basic fact, that appears in the above-mentioned reference [28], essentially depends on a large deviation theorem.

We have a more convenient result in a special case of Hoeffding's inequality, cf. [10]. Let $X_1, X_2, \dots$ be a sequence of bounded i.i.d. random variables, such that $b_i \leq X_i \leq a_i$, and let $S_n = \sum_{i=1}^n X_i$. Then by Hoeffding's inequality, for any $x > 0$ we have

$$\mathbf{P}\left\{|S_n - \mathbf{E}(S_n)| \geq x\left(\frac{1}{4}\sum_{i=1}^n (a_i - b_i)^2\right)^{\frac{1}{2}}\right\} \leq 2\,\mathrm{e}^{-\frac{x^2}{2}}.$$

If $\mathbf{E}(X_i) = 0$ and $b_i = -a_i$ here, then $\frac{1}{4}\sum_{i=1}^n (a_i - b_i)^2 = \sum_{i=1}^n a_i^2 = \mathbf{Var}(S_n)$ if and only if $X_i = a_i X_i'$, where $\mathbf{P}\{X_i' = \pm 1\} = \frac{1}{2}$, $1 \leq i \leq n$.

Thus if $S = \sum_r a_r X'_r$, where not all a_r are zero and $\mathbf{Var}(S) = \sum_r a_r^2 < \infty$, we get

$$\mathbf{P}\left\{|S| \geq x\,(\mathbf{Var}(S))^{\frac{1}{2}}\right\} \leq 2\,\mathrm{e}^{-\frac{x^2}{2}} \qquad (x \geq 0). \tag{2.4}$$

The summation above may extend either to finitely many or to countably many terms. Let $S_1, S_2, \dots, S_N$ be arbitrary sums of the above type: $S_k = \sum_r a_{kr} X'_{kr}$, $\mathbf{P}\{X'_{kr} = \pm 1\} = \frac{1}{2}$, $1 \leq k \leq N$, where X'_{kr} and X'_{ls} can be dependent when $k \neq l$. Then by the inequality (2.4) we obtain the following analog of Lemma A: for any $C > 1$ and $N \geq 1$,

$$\begin{aligned} &\mathbf{P}\left\{\sup_{1\leq k\leq N} |S_k| \geq (2C\log N)^{\frac{1}{2}} \sup_{1\leq k\leq N} (\mathbf{Var}(S_k))^{\frac{1}{2}}\right\} \\ &\leq \sum_{k=1}^{N} \mathbf{P}\left\{|S_k| \geq (2C\log N\,\mathbf{Var}(S_k))^{\frac{1}{2}}\right\} \leq 2N\mathrm{e}^{-C\log N} = 2N^{1-C}. \end{aligned} \tag{2.5}$$

Lemma A easily implies that the time lags (2.3) are uniformly small if m is large enough.

Lemma B. *For any $K > 0$, $C > 1$, and for any $m \geq m_0(C)$, we have*

$$\begin{aligned} \mathbf{P}\left\{\sup_{0\leq k2^{-2m}\leq K} |T_{m+1}(k)2^{-2(m+1)} - k2^{-2m}| \geq \left(\frac{3}{2}CK\log_* K\right)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m}\right\} \\ \leq 2(K2^{2m})^{1-C}, \end{aligned}$$

where $\log_* x = \max\{1, \log x\}$.

This lemma and the refinement property (2.2) implies the uniform closeness of two consecutive approximations if m is large enough.

Lemma C. *For any $K > 0$, $C > 1$, and for any $m \geq m_1(C)$, we have*

$$\begin{aligned} \mathbf{P}\left\{\sup_{0\leq k2^{-2m}\leq K} |\tilde{B}_{m+1}(k2^{-2m}) - \tilde{B}_m(k2^{-2m})| \geq K_*^{\frac{1}{4}} (\log_* K)^{\frac{3}{4}} m 2^{-\frac{m}{2}}\right\} \\ \leq 3(K2^{2m})^{1-C}, \end{aligned}$$

where $K_* = \max\{1, K\}$.

Based on this lemma, it is not difficult to show the following convergence result.

Theorem A. *The shrunken RWs $\tilde{B}_m(t)$ ($t \geq 0, m = 0, 1, 2, \dots$) almost surely uniformly converge to a Wiener process $W(t)$ ($t \geq 0$) on any compact interval $[0, K]$, $K > 0$. For any $K > 0$, $C \geq 3/2$, and for any $m \geq m_2(C)$, we have*

$$\mathbf{P}\left\{\sup_{0\leq t\leq K} |W(t) - \tilde{B}_m(t)| \geq K_*^{\frac{1}{4}} (\log_* K)^{\frac{3}{4}} m 2^{-\frac{m}{2}}\right\} \leq 6(K2^{2m})^{1-C}.$$

Now taking $C = 3$ in Theorem A and using the Borel–Cantelli lemma, we get

$$\sup_{0\leq t\leq K} |W(t) - \tilde{B}_m(t)| < O(1) m 2^{-\frac{m}{2}} \qquad \text{a.s.} \qquad (m \to \infty)$$

and

$$\sup_{0\leq t\leq K} |W(t) - \tilde{B}_m(t)| < K^{\frac{1}{4}} (\log K)^{\frac{3}{4}} \qquad \text{a.s.} \qquad (K \to \infty)$$

for any m large enough, $m \geq m_2(3)$.

Next we are going to study the properties of another nested sequence of random walks, obtained by Skorohod embedding. This sequence is not identical, though asymptotically equivalent to the above RW construction, cf. [28, Theorem 4]. Given a Wiener process W, first we define the stopping times which yield the Skorohod embedded process $B_m(k2^{-2m})$ into W. For every $m \geq 0$ let $s_m(0) = 0$ and

$$s_m(k+1) = \inf\{s : s > s_m(k), |W(s) - W(s_m(k))| = 2^{-m}\} \qquad (k \geq 0). \tag{2.6}$$

With these stopping times the embedded process by definition is

$$B_m(k2^{-2m}) = W(s_m(k)) \qquad (m \geq 0, k \geq 0). \tag{2.7}$$

This definition of B_m can be extended to any real $t \geq 0$ by pathwise linear interpolation. The next lemma describes some useful facts about the relationship between $\tilde{B}_m$ and B_m. These follow from [28, Lemmas 5,7 and Theorem 4], with some minor modifications.

In general, roughly saying, $\tilde{B}_m$ is more useful when someone wants to generate stochastic processes from scratch, while B_m is more advantageous when someone needs a discrete approximation of given processes, like in the case of stochastic integration.

Lemma D. *For any $C \geq 3/2$, $K > 0$, take the following subset of the sample space:*

$$A_m = \left\{ \sup_{n>m} \sup_{0 \leq k2^{-2m} \leq K} |2^{-2n} T_{m,n}(k) - k2^{-2m}| < 6(CK_* \log_* K)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m} \right\}, \tag{2.8}$$

where $T_{m,n}(k) = T_n \circ T_{n-1} \circ \cdots \circ T_m(k)$ for $n > m \geq 0$ and $k \geq 0$. Then for any $m \geq m_3(C)$,

$$\mathbf{P}\{A_m^c\} \leq 4(K2^{2m})^{1-C}.$$

Moreover, $\lim_{n\to\infty} 2^{-2n} T_{m,n}(k) = t_m(k)$ exists almost surely and on the set A_m we have

$$\tilde{B}_m(k2^{-2m}) = W(t_m(k)) \qquad (0 \leq k2^{-2m} \leq K),$$

cf. (2.7). Further, on A_m except for a zero probability subset, $s_m(k) = t_m(k)$ and

$$\sup_{0 \leq k2^{-2m} \leq K} |s_m(k) - k2^{-2m}| \leq 6(CK_* \log_* K)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m} \qquad (m \geq m_3(C)). \tag{2.9}$$

If the Wiener process is built by the RW construction described above using a sequence $\tilde{B}_m$ ($m \geq 0$) of nested RWs and then one constructs the Skorohod embedded RWs B_m ($m \geq 0$), it is natural to ask what the approximating properties of the latter are. The answer described by the next theorem is that they are essentially the same as the ones of $\tilde{B}_m$, cf. Theorem A.

Lemma 1. *For every $K > 0$, $C \geq 3/2$ and $m \geq m_3(C)$ we have*

$$\mathbf{P}\left\{ \sup_{0 \leq t \leq K} |W(t) - B_m(t)| \geq K_*^{\frac{1}{4}} (\log_* K)^{\frac{3}{4}} m 2^{-\frac{m}{2}} \right\} \leq 10(K2^{2m})^{1-C}.$$

Proof. By the triangle inequality,

$$\sup_{0 \leq t \leq K} |W(t) - B_m(t)| \leq \sup_{0 \leq t \leq K} \left| W(t) - \tilde{B}_m(t) \right| + \sup_{0 \leq t \leq K} \left| \tilde{B}_m(t) - B_m(t) \right|.$$

By Lemma D and equation (2.7), on the set A_m defined by (2.8) we have

$$\tilde{B}_m(k2^{-2m}) = W(s_m(k)) = B_m(k2^{-2m}),$$

except for a zero probability subset when $m \geq m_3(C)$. Since both $\tilde{B}_m(t)$ and $B_m(t)$ are obtained by pathwise linear interpolation based on the vertices at $k2^{-2m} \in [0, K]$, they are identical on A_m, except for a zero probability subset of it when $m \geq m_3(C)$. Thus

$$\begin{aligned} &\mathbf{P}\left\{\sup_{0\leq t\leq K} |W(t) - B_m(t)| \geq K_*^{\frac{1}{4}} (\log_* K)^{\frac{3}{4}} m 2^{-\frac{m}{2}}\right\} \\ &\leq \quad \mathbf{P}\{A_m^c\} + \mathbf{P}\left\{\sup_{0\leq t\leq K} \left|W(t) - \tilde{B}_m(t)\right| \geq K_*^{\frac{1}{4}} (\log_* K)^{\frac{3}{4}} m 2^{-\frac{m}{2}}\right\} \end{aligned}$$

Then by Theorem A and Lemma D we get the statement of the theorem. □

2.2 Approximation of continuous martingales

Beside the RW construction of standard Brownian motion, the other main tool applied in this section is a theorem of Dambis (1965) and Dubins–Schwarz (1965) and an extension of it, cf. Theorems B and C below. Briefly saying, these theorems state that any continuous local martingale $(M(t), t \geq 0)$ can be transformed into a standard Brownian motion by time-change. Then somewhat loosely speaking, the resulting Brownian motion takes on the same values in the same order as $M(t)$, only the corresponding time instants may differ. These and other necessary matters about continuous local martingales will be taken from and discussed in the style of [23] in the sequel.

Below it is supposed that an increasing family of sub-σ-algebras $(\mathcal{F}_t, t \geq 0)$ is given in the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and the given continuous local martingale M is adapted to it.

In the case of a continuous local martingale $M(t)$ vanishing at 0 its quadratic variation $\langle M, M\rangle_t$ is a process with almost surely continuous and non-decreasing sample paths vanishing at 0. This will be one of the two time-changes we are going to use in the sequel. The other one is a quasi-inverse of the quadratic variation:

$$T_s = \inf\{t : \langle M, M\rangle_t > s\}, \tag{2.10}$$

where $\inf(\emptyset) = \infty$ by definition. Then the sample paths of the process T_s are almost surely increasing, but only right-continuous, since such a path has a jump at any value where the quadratic variation has a constant level-stretch. Beside this, T_s may be infinite valued. The duality between the two time-changes is expressed by $\langle M, M\rangle_t = \inf\{s : T_s > t\}$. Observe that T_s cannot have constant level-stretches since this would imply jumps for $\langle M, M\rangle_t$. Also the continuity of $\langle M, M\rangle_t$ gives that $\langle M, M\rangle_{T_s} = s$ $(s \geq 0)$, while we have only $T_{\langle M,M\rangle_t} \geq t$ $(t \geq 0)$ in the opposite direction. It is clear that

$$\langle M, M\rangle_t < s \Longrightarrow t < T_s, \text{ but } t < T_s \Longrightarrow \langle M, M\rangle_t \leq s, \tag{2.11}$$

while

$$\langle M, M\rangle_t \leq, \geq, > s \Longleftrightarrow t \leq, \geq, > T_s, \tag{2.12}$$

respectively.

Theorem B. *[23, V (1.6), p.181] If M is a continuous $(\mathcal{F}_t)$-local martingale vanishing at 0 and such that $\langle M, M\rangle_\infty = \infty$ a.s., then $W(s) = M(T_s)$ is an $(\mathcal{F}_{T_s})$-Brownian motion and $M(t) = W(\langle M, M\rangle_t)$.*

Similar statement is true when $\langle M, M\rangle_\infty < \infty$ is possible. Note that on the set $\{\langle M, M\rangle_\infty < \infty\}$ the limit $M(\infty) = \lim_{t\to\infty} M(t)$ exists with probability 1, cf. [23, IV (1.26), p. 131].

Theorem C. *[23, V (1.7), p.182] If M is a continuous $(\mathcal{F}_t)$-local martingale vanishing at 0 and such that $\langle M, M\rangle_\infty < \infty$ with positive probability, then there exists an enlargement $(\widetilde{\Omega}, \widetilde{\mathcal{F}}_t, \widetilde{\mathbf{P}})$ of $(\Omega, \mathcal{F}_{T_t}, \mathbf{P})$ and a Wiener process $\widetilde{\beta}$ on $\widetilde{\Omega}$ independent of M such that the process*

$$W(s) = \begin{cases} M(T_s) & \text{if } s < \langle M, M\rangle_\infty, \\ M(\infty) + \widetilde{\beta}(s - \langle M, M\rangle_\infty) & \text{if } s \geq \langle M, M\rangle_\infty \end{cases}$$

is a standard Brownian motion and $M(t) = W(\langle M, M\rangle_t)$ *for* $t \geq 0$.

From now on, W will always refer to the Wiener process obtained from M by the above time-change, the so-called *DDS Wiener process (or DDS Brownian motion)* of M.

Now Skorohod-type stopping times can be defined for M, similarly as for W in (2.6). For $m \geq 0$, let $\tau_m(0) = 0$ and

$$\tau_m(k+1) = \inf\left\{t : t > \tau_m(k), |M(t) - M(\tau_m(k))| = 2^{-m}\right\} \qquad (k \geq 0). \tag{2.13}$$

The $(m+1)$st stopping time sequence is a refinement of the mth in the sense that $(\tau_m(k))_{k=0}^{\infty}$ is a subsequence of $(\tau_{m+1}(j))_{j=0}^{\infty}$ so that for any $k \geq 0$ there exist j_1 and j_2, $\tau_{m+1}(j_1) = \tau_m(k)$ and $\tau_{m+1}(j_2) = \tau_m(k+1)$, where the difference $j_2 - j_1 \geq 2$, even.

Lemma 2. *With the stopping times defined by (2.13) from a continuous local martingale M one can directly obtain the sequence of shrunken RWs that almost surely converges to the DDS Wiener process W of M, cf. (2.7):*

$$B_m(k2^{-2m}) = W(s_m(k)) = M(\tau_m(k)), \qquad s_m(k) = \langle M, M\rangle_{\tau_m(k)}$$

[but $\tau_m(k) \leq T_{s_m(k)}$], where for $m \geq 0$, the non-negative integer k is taking values (depending on ω) until $s_m(k) \leq \langle M, M\rangle_\infty$.

Proof. By Theorems B and C it follows that $W(\langle M, M\rangle_{\tau_m(k)}) = M(\tau_m(k))$. This implies that $s_m(k) \leq \langle M, M\rangle_{\tau_m(k)}$. Then consider first the case $k = 1$. If $s_m(1) < \langle M, M\rangle_{\tau_m(1)}$ held, then $T_{s_m(1)} < \tau_m(1)$ would follow by (2.12), and this would lead to a contradiction because $M(T_{s_m(1)}) = W(s_m(1)) = \pm 2^{-m}$. For values $k > 1$, induction with a similar argument can show the statement of the lemma. □

In the sequel B_m will always denote the sequence of shrunken RWs defined by Lemma 2.

2.2.1 Approximation of the quadratic variation process

Our next objective is to show that the quadratic variation of M can be obtained as an almost sure limit of a point process related to the above stopping times that we will call *a discrete quadratic variation process*:

$$\begin{aligned} N_m(t) &= 2^{-2m}\#\{r : r > 0, \tau_m(r) \leq t\} \\ &= 2^{-2m}\#\{r : r > 0, s_m(r) \leq \langle M, M\rangle_t\} \qquad (t \geq 0). \end{aligned} \tag{2.14}$$

Clearly, the paths of N_m are non-decreasing pure jump functions, the jumping times being exactly the stopping times $\tau_m(k)$. Moreover, $N_m(\tau_m(k)) = k2^{-2m}$ and the magnitudes of jumps are constant 2^{-2m} when m is fixed.

Lemma 3. *Let M be a continuous local martingale vanishing at 0, let $\langle M, M\rangle$ be the quadratic variation, T be its quasi-inverse (2.10), and N_m be the discrete quadratic variation defined in (2.14). Fix $K > 0$ and take a sequence $a_m = O(m^{-2-\epsilon}2^{2m})K$ with some $\epsilon > 0$, where $a_m \geq K \vee 1$ for any $m \geq 1$ ($x \vee y = \max(x, y)$, $x \wedge y = \min(x, y)$).*

(a) Then for any $C \geq 3/2$ and $m \geq m_4(C)$ we have

$$\mathbf{P}\left\{\sup_{0 \leq t \leq K} |\langle M, M\rangle_t \wedge a_m - N_m(t \wedge T_{a_m})| \geq 12(Ca_m \log_* a_m)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m}\right\} \\ \leq 3(a_m 2^{2m})^{1-C}.$$

(b) Suppose that the quadratic variation satisfies the following tail-condition: a sequence (a_m) fulfilling the above assumptions can be chosen so that

$$\mathbf{P}\left\{\langle M,M\rangle_t > a_m\right\} \leq D(t) m^{-1-\epsilon}, \tag{2.15}$$

where $D(t)$ is some finite valued function of $t \in \mathbb{R}_+$. Then for any $C \geq 3/2$ and $m \geq m_4(C)$ it follows that

$$\mathbf{P}\left\{\sup_{0\leq t\leq K} |\langle M,M\rangle_t - N_m(t)| \geq 12(Ca_m \log_* a_m)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m}\right\} \leq 3(a_m 2^{2m})^{1-C} + D(K)m^{-1-\epsilon}.$$

Proof. The basic idea of the proof is that the Skorohod stopping times of a Wiener process are asymptotically uniformly distributed as shown by (2.9), while the case of a continuous local martingale can be reduced to the former by the DDS representation, cf. Lemma 2.

Introduce the abbreviation $h_{a,m} = 11.1\,(Ca_m \log_* a_m)^{\frac{1}{2}}\, m^{\frac{1}{2}} 2^{-m}$. Then $h_{a,m} = O(m^{-\epsilon}) \to 0$ as $m \to \infty$. We need a truncation here using the sequence a_m, since the quadratic variation $\langle M,M\rangle_t$ is not a bounded random variable in general. By (2.12) and (2.14),

$$\begin{aligned} N_m(t\wedge T_{a_m}) &= 2^{-2m}\#\{r : r>0, \tau_m(r) \leq t\wedge T_{a_m}\} \\ &= 2^{-2m}\#\{r : r>0, s_m(r) \leq \langle M,M\rangle_t \wedge a_m\}. \end{aligned}$$

On the event

$$A_{a,m} = \left\{\sup_{0\leq r2^{-2m}\leq 2a_m} |s_m(r) - r2^{-2m}| \leq h_{a,m}\right\},$$

if $r = \lfloor(\langle M,M\rangle_t \wedge a_m + h_{a,m})2^{2m}\rfloor + 1$, then $s_m(r) > \langle M,M\rangle_t \wedge a_m$, so $s_m(r)$ is not included in $N_m(t\wedge T_{a_m})$. Observe here that $a_m + h_{a,m} + 2^{-2m} \leq 2a_m$ if m is large enough, $m \geq m_4(C)$, where we also suppose that $m_4(C) \geq m_3(C)$ and $m_3(C)$ is defined by Lemma D. This explains why the sup is taken for $r2^{-2m} \leq 2a_m$ in the definition of $A_{a,m}$. Similarly on $A_{a,m}$, if $r = \lfloor(\langle M,M\rangle_t \wedge a_m - h_{a,m})2^{2m}\rfloor$, then $s_m(r) \leq \langle M,M\rangle_t \wedge a_m$, so $s_m(r)$ must be included in $N_m(t\wedge T_{a_m})$. Hence

$$\langle M,M\rangle_t \wedge a_m - h_{a,m} - 2^{-2m} \leq N_m(t\wedge T_{a_m}) \leq \langle M,M\rangle_t \wedge a_m + h_{a,m} + 2^{-2m}, \tag{2.16}$$

for any $t \in [0,K]$ on $A_{a,m}$.

Now $6\,(C2a_m \log_*(2a_m))^{\frac{1}{2}}\, m^{\frac{1}{2}} 2^{-m} \leq 11.1\,(Ca_m \log_* a_m)^{\frac{1}{2}}\, m^{\frac{1}{2}} 2^{-m} = h_{a,m}$, since $\log_*(2a_m) \leq (1 + \log 2)\log_* a_m$. Hence it follows by Lemma D that

$$\mathbf{P}\left\{A^c_{a,m}\right\} \leq 4(2a_m 2^{2m})^{1-C} \leq 3(a_m 2^{2m})^{1-C},$$

when $C \geq 3/2$ and $m \geq m_4(C)$. Noticing that $0.9\,(Ca_m \log_* a_m)^{\frac{1}{2}}\, m^{\frac{1}{2}} 2^{-m} > 2^{-2m}$ for any $m \geq 1$, this and (2.16) prove (a).

Part (b) follows from (a), the inequality

$$\begin{aligned} |\langle M,M\rangle_t - N_m(t)| &\leq |\langle M,M\rangle_t - \langle M,M\rangle_t \wedge a_m| \\ &+ |\langle M,M\rangle_t \wedge a_m - N_m(t\wedge T_{a_m})| \\ &+ |N_m(t\wedge T_{a_m}) - N_m(t)|, \end{aligned} \tag{2.17}$$

and from the following simple relationships between events:

$$\{\langle M,M\rangle_t \wedge a_m \neq \langle M,M\rangle_t\} = \{\langle M,M\rangle_t > a_m\}$$

and

$$\{N_m(t\wedge T_{a_m}) \neq N_m(t)\} = \{t > T_{a_m}\} = \{\langle M,M\rangle_t > a_m\},$$

cf. (2.12). □

We mention that when the quadratic variation $\langle M, M\rangle_t$ is almost surely bounded above by a finite valued function $g(t)$ for $t > 0$, statements (a) and (b) of Lemma 4 simplify as

$$\mathbf{P}\left\{\sup_{0\le t\le K} |\langle M, M\rangle_t - N_m(t)| \ge 12(Cg_*(K)\log_* g(K))^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m}\right\} \le 3(g(K)2^{2m})^{1-C}$$

for any $K > 0$, $C \ge 3/2$ and $m \ge m_4(C)$.

The statement of the next theorem corresponds to the main result in Karandikar [12], though the method applied is different and here we give a rate of convergence as well.

Theorem 1. *Using the same notations as in Lemma 3, we have*

$$\sup_{0\le t\le K} |\langle M, M\rangle_t - N_m(t)| < O(1)m^{\frac{1}{2}}2^{-m} \qquad a.s. \qquad (m \to \infty)$$

and

$$\sup_{0\le t\le K} |\langle M, M\rangle_t - N_m(t)| < K^{\frac{1}{2}}(\log K)^{\frac{1}{2}} \qquad a.s. \qquad (K \to \infty)$$

for any m large enough, $m \ge m_4(3)$.

Proof. To show the first statement take e.g. $C = 3/2$ and $a_m = K\log(m+2)$ in Lemma 3 (a). Consider the inequality (2.17). Since $\langle M, M\rangle_K$ is finite-valued and $a_m \to \infty$, if m is large enough, depending on ω, $\langle M, M\rangle_K < a_m$ holds and then $t < T_{a_m}$ holds as well by (2.11). These remarks show that the first and the third terms on the right hand side of inequality (2.17) are zero if m is large enough. Further, statement of Lemma 3 (a) can be applied to the second term. This, with the Borel–Cantelli lemma, proves the theorem.

The second statement of theorem follows similarly from Lemma 3 (a) by the Borel–Cantelli lemma, taking $C = 3$ and $a_m = K$. □

2.2.2 Strong approximation of continuous martingales

Now we are ready to discuss the strong approximation of continuous local martingales by time-changed random walks.

Lemma 4. *Let M be a continuous local martingale vanishing at 0, let $\langle M, M\rangle$ be the quadratic variation and T be its quasi-inverse (2.10). Denote by B_m the sequence of shrunken RWs embedded into M by Lemma 2. Fix $K > 0$ and take a sequence $a_m = O(m^{-7-\epsilon}2^{2m})K$ with some $\epsilon > 0$, where $a_m \ge K \vee 1$ for any $m \ge 1$.*

(a) Then for any $C \ge 3/2$ and $m \ge m_3(C)$ we have

$$\mathbf{P}\left\{\sup_{0\le t\le K} |M(t\wedge T_{a_m}) - B_m(\langle M, M\rangle_t \wedge a_m)| \ge a_m^{\frac{1}{4}}(\log_* a_m)^{\frac{3}{4}} m 2^{-\frac{m}{2}}\right\} \le 10(a_m 2^{2m})^{1-C}.$$

(b) Under the tail-condition (2.15), for any $C \ge 3/2$ and $m \ge m_3(C)$ it follows that

$$\mathbf{P}\left\{\sup_{0\le t\le K} |M(t) - B_m(\langle M, M\rangle_t)| \ge a_m^{\frac{1}{4}}(\log_* a_m)^{\frac{3}{4}} m 2^{-\frac{m}{2}}\right\} \le 10(a_m 2^{2m})^{1-C} + D(K)m^{-1-\epsilon}.$$

Proof. First, take the DDS Wiener process $W(s)$ obtained from $M(t)$ by the time-change T_s as described by Theorems B and C. Since below we are going to use $W(s)$ and also the time change T_s only for arguments $s \leq \langle M,M\rangle_\infty$, we can always assume that $W(s) = M(T_s)$ and $M(t) = W(\langle M,M\rangle_t)$, irrespective of the fact whether $\langle M,M\rangle_\infty = \infty$ or not. Second, define the nested sequence of shrunken RWs $B_m(s)$ by Lemma 2. Then a quasi-inverse time-change $\langle M,M\rangle_t$ is applied to $B_m(s)$ that gives $B_m(\langle M,M\rangle_t)$ which will be the sequence of time-changed shrunken RWs approximating $M(t)$.

Since T_s may have jumps, we get that

$$\begin{aligned}\sup_{0\leq t\leq K}\left|M(t) - B_m(\langle M,M\rangle_t)\right| &\geq \sup_{0\leq s\leq \langle M,M\rangle_K}\left|M(T_s) - B_m(\langle M,M\rangle_{T_s})\right| \\ &= \sup_{0\leq s\leq \langle M,M\rangle_K}\left|W(s) - B_m(s)\right|. \end{aligned} \tag{2.18}$$

Recalling however that the intervals of constancy are the same for $M(t)$ and for $\langle M,M\rangle_t$ [23, IV (1.13), p.125], there is in fact equality in (2.18). To go on, we need a truncation using the sequence a_m, since the quadratic variation $\langle M,M\rangle_t$ is not a bounded random variable in general. Then (2.18) (with equality as explained above) and (2.12) imply

$$\begin{aligned}&\sup_{0\leq t\leq K}\left|M(t\wedge T_{a_m}) - B_m(\langle M,M\rangle_t\wedge a_m)\right| \\ &= \sup_{0\leq s\leq \langle M,M\rangle_K}\left|M(T_s\wedge T_{a_m}) - B_m(\langle M,M\rangle_{T_s}\wedge a_m)\right| \\ &= \sup_{0\leq s\leq a_m\wedge\langle M,M\rangle_K}\left|W(s) - B_m(s)\right| \\ &\leq \sup_{0\leq s\leq a_m}\left|W(s) - B_m(s)\right|.\end{aligned}$$

Hence by Lemma 1, with $m \geq m_3(C)$,

$$\begin{aligned}&\mathbf{P}\left\{\sup_{0\leq t\leq K}\left|M(t\wedge T_{a_m}) - B_m(\langle M,M\rangle_t\wedge a_m)\right| \geq a_m^{\frac14}(\log_* a_m)^{\frac34} m 2^{-\frac{m}{2}}\right\} \\ &\leq \mathbf{P}\left\{\sup_{0\leq s\leq a_m}\left|W(s) - B_m(s)\right| \geq a_m^{\frac14}(\log_* a_m)^{\frac34} m 2^{-\frac{m}{2}}\right\} \\ &\leq 10(a_m 2^{2m})^{1-C}.\end{aligned}$$

This proves (a).

To show (b) it is enough to consider the inequality

$$\begin{aligned}&\sup_{0\leq t\leq K}\left|M(t) - B_m(\langle M,M\rangle_t)\right| \\ \leq\ &\sup_{0\leq t\leq K}\left|M(t) - M(t\wedge T_{a_m})\right| + \sup_{0\leq t\leq K}\left|M(t\wedge T_{a_m}) - B_m(\langle M,M\rangle_t\wedge a_m)\right| \\ &+ \sup_{0\leq t\leq K}\left|B_m(\langle M,M\rangle_t\wedge a_m) - B_m(\langle M,M\rangle_t)\right|.\end{aligned}$$

From this point the proof is similar to the proof of Lemma 3 (b). □

Kiefer [14] proved in the Brownian case $M = W$ that using Skorohod embedding one cannot embed a standardized RW into W with convergence rate better than $O(1)n^{-\frac14}(\log n)^{\frac12}(\log\log n)^{\frac14}$, where n is the number of points used in the approximation. Since the next theorem gives a rate of convergence $O(1)n^{-\frac14}\log n$ (the number of points used is $n = K2^{2m}$), this rate is close to the best we can have with a Skorohod-type embedding. The same remark is valid for Theorem 3 below.

Theorem 2. *Applying the same notations as in Lemma 4, we have*

$$\sup_{0\le t\le K} |M(t) - B_m(\langle M, M\rangle_t)| < O(1)m2^{-\frac{m}{2}} \qquad a.s. \qquad (m \to \infty)$$

and

$$\sup_{0\le t\le K} |M(t) - B_m(\langle M, M\rangle_t)| < K^{\frac{1}{4}}(\log K)^{\frac{3}{4}} \qquad a.s. \qquad (K \to \infty)$$

for any m large enough, $m \ge m_3(3)$.

Proof. The statements follow from Lemma 4 in a similar way as Theorem 1 followed from Lemma 3. □

We mention that when M is a continuous local martingale vanishing at 0 and there is a deterministic function f on $\mathbb{R}_+$ such that $\langle M, M\rangle_t = f(t)$ a.s., then it follows that M is Gaussian and has independent increments, see [23, V (1.14), p.186]. Conversely, if M is a continuous Gaussian martingale, then $\langle M, M\rangle_t = f(t)$ a.s., see [23, IV (1.35), p.133]. In this case the "twist and shrink" construction of Brownian motion described in Section 2 can be extended to a construction of $M(t)$ (or a simulation algorithm in practice). Namely, we have

$$\left|M(t) - \tilde{B}_m(f(t))\right| \le O(1)m2^{-\frac{m}{2}} \qquad \text{a.s.} \qquad (m \to \infty).$$

Here $\tilde{B}_m(t) = 2^{-m}\tilde{S}_m(t2^{2m})$ $(m \ge 0)$ denotes the nested sequence of the RW construction described in Section 2.

Combining the previous results one can replace $\langle M, M\rangle_t$ by the discrete quadratic variation $N_m(t)$ when approximating $M(t)$ by time-changed shrunken RWs.

Lemma 5. *Let M be a continuous local martingale vanishing at 0, let $\langle M, M\rangle$ be the quadratic variation, T be its quasi-inverse (2.10) and N_m be the discrete quadratic variation defined by (2.14). Denote by B_m the sequence of shrunken RWs embedded into M by Lemma 2. Fix $K > 0$ and take a sequence $a_m = O(m^{-7-\epsilon}2^{2m})K$ with some $\epsilon > 0$, where $a_m \ge K \vee 1$ for any $m \ge 1$.*

(a) Then for any $C \ge 3/2$ and $m \ge m_5(C)$ we have

$$\mathbf{P}\left\{\sup_{0\le t\le K} |M(t\wedge T_{a_m}) - B_m(N_m(t\wedge T_{a_m})| \ge 2a_m^{\frac{1}{4}}(\log_* a_m)^{\frac{3}{4}}m2^{-\frac{m}{2}}\right\} \le 14(a_m2^{2m})^{1-C}.$$

(b) Under the tail-condition (2.15), for any $C \ge 3/2$ and $m \ge m_5(C)$ it follows that

$$\mathbf{P}\left\{\sup_{0\le t\le K} |M(t) - B_m(N_m(t))| \ge 2a_m^{\frac{1}{4}}(\log_* a_m)^{\frac{3}{4}}m2^{-\frac{m}{2}}\right\} \le 14(a_m2^{2m})^{1-C} + D(K)m^{-1-\epsilon}.$$

Proof. For proving (a) we use the triangle inequality

$$\begin{aligned}\sup_{0\le t\le K} &|M(t\wedge T_{a_m}) - B_m(N_m(t\wedge T_{a_m})| \\ &\le \sup_{0\le t\le K} |M(t\wedge T_{a_m}) - B_m(\langle M, M\rangle_t \wedge a_m)| \\ &\quad + \sup_{0\le t\le K} |B_m(\langle M, M\rangle_t \wedge a_m) - B_m(N_m(t\wedge T_{a_m})|. \end{aligned} \tag{2.19}$$

Since the first term on the right hand side can be estimated by Theorem 2, we have to consider the second term. For $m \geq 1$ introduce the abbreviation

$$L_{a,m} = 13(Ca_m \log_* a_m)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m} \geq 12(Ca_m \log_* a_m)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m} + 2^{-2m}.$$

By Theorem 1 (a), with $C \geq 3/2$ and $m \geq m_4(C)$ it follows that

$$\begin{aligned}
&\sup_{0\leq t\leq K} |B_m(\langle M,M\rangle_t \wedge a_m) - B_m(N_m(t\wedge T_{a_m})| \\
&\leq \sup_{0\leq t\leq K} \left|B_m\left(\lfloor(\langle M,M\rangle_t \wedge a_m)2^{2m}\rfloor 2^{-2m}\right) - B_m(N_m(t\wedge T_{a_m})\right| + 2^{-m} \\
&\leq \sup_{0\leq k2^{-2m}\leq \langle M,M\rangle_K\wedge a_m} \ \sup_{|r-k|2^{-2m}\leq L_{a,m}} |B_m(k2^{-2m}) - B_m(r2^{-2m})| + 2^{-m} \\
&\leq \sup_{0\leq k2^{-2m}\leq a_m} \ \sup_{0\leq r2^{-2m}\leq L_{a,m}} |B_m^{(k)}(r2^{-2m})| + 2^{-m},
\end{aligned}$$

except for an event of probability $\leq 3(a_m 2^{2m})^{1-C}$, since the difference of a shrunken RW at two dyadic points equals the value of some shrunken RW $B_m^{(k)}$ at a dyadic point.

Then we can apply estimate (2.5) with some $C' > 1$ for the last expression:

$$\mathbf{P}\left\{\sup_{0\leq k2^{-2m}\leq a_m} \ \sup_{0\leq r2^{-2m}\leq L_{a,m}} |B_m^{(k)}(r2^{-2m})| \geq \left(2C'\log N \sup_{k,r} \mathbf{Var}(B_m^{(k)}(r2^{-2m}))\right)^{\frac{1}{2}}\right\} \leq 2N^{1-C'},$$

where $N = \lfloor a_m 2^{2m}\rfloor \lfloor L_{a,m} 2^{2m}\rfloor$ and $\sup_{k,r} \mathbf{Var}\left(B_m^{(k)}(r2^{-2m})\right) \leq L_{a,m}$. Choose here C' so that $1 - C' = \frac{2}{3}(1-C)$. Then a simple computation shows that $2N^{1-C'} \leq (a_m 2^{2m})^{1-C}$, also $\log N \leq 8m \log_* C \log_* a_m$, and

$$\left(2C'\log N \sup_{k,r} \mathbf{Var}(B_m^{(k)}(r2^{-2m}))\right)^{\frac{1}{2}} + 2^{-m} \leq a_m^{\frac{1}{4}} (\log_* a_m)^{\frac{3}{4}} m 2^{-\frac{m}{2}}$$

if $m \geq m_5(C) \geq m_4(C)$. This argument and Theorem 2(a) applied to (2.19) give (a).

Statement (b) again follow from (a) in a similar way as in Lemma 3. □

Theorem 3. *With the same notations as in Lemma 5, we have*

$$\sup_{0\leq t\leq K} |M(t) - B_m(N_m(t))| < O(1) m 2^{-\frac{m}{2}} \qquad a.s. \qquad (m\to\infty)$$

and

$$\sup_{0\leq t\leq K} |M(t) - B_m(N_m(t))| < K^{\frac{1}{4}} (\log K)^{\frac{3}{4}} \qquad a.s. \qquad (K\to\infty)$$

for any m large enough, $m \geq m_5(3)$.

Proof. The statements follow from Lemma 5 in a similar way again as Theorem 1 followed from Lemma 3. □

2.3 Symmetrically evolving martingales

It is important both from theoretical and practical (e.g. simulation) points of view that the shrunken RW B_m and the corresponding discrete quadratic variation process N_m be independent when approximating M as in Theorem 3. This leads to the question of independence of the DDS Brownian motion W and quadratic variation $\langle M, M\rangle$ in the case of a continuous local martingale M. For, by Lemma 2, B_m depends only on W and, by (2.14), N_m is determined by $\langle M, M\rangle$ alone. Conversely, if the processes B_m and N_m are independent for any m large enough, then so are W and $\langle M, M\rangle$ too by Lemma 1 and 1. It will turn out from the next theorem that the basic notion in this respect is the symmetry of the increments of M given the past. Thus we will say that a stochastic process $M(t)$ $(t \geq 0)$ is *symmetrically evolving* (or has symmetric increments given the past) if for any positive integer n, reals $0 \leq s < t_1 < \cdots < t_n$ and Borel sets of the line $U_1, \ldots U_n$ we have

$$\mathbf{P}\left\{\Gamma \mid \mathcal{F}_s^0\right\} = \mathbf{P}\left\{\Gamma^- \mid \mathcal{F}_s^0\right\}, \tag{2.20}$$

where $\Gamma = \{M(t_1) - M(s) \in U_1, \ldots, M(t_n) - M(s) \in U_n\}$, Γ^- is the same, but each U_j replaced by $-U_j$, and $\mathcal{F}_s^0 = \sigma(M(u), 0 \leq u \leq s)$ is the filtration generated by the past of M. If $M(t)$ has finite expectation for any $t \geq 0$, then this condition expresses a very strong martingale property.

Condition (2.20) is clearly equivalent to the following one: for arbitrary positive integers n, j, reals $0 \leq s_j < \cdots < s_1 \leq s < t_1 < \ldots t_n$ and Borel-sets $V_1, \ldots, V_j, U_1, \ldots, U_n$ one has

$$\mathbf{P}\left\{\Gamma \cap \Lambda\right\} = \mathbf{P}\left\{\Gamma^- \cap \Lambda\right\}, \tag{2.21}$$

where Γ and Γ^- are defined above and $\Lambda = \{M(s_1) \in V_1, \ldots, M(s_j) \in V_j\}$.

Our Theorem 4 below is basically a reformulation of Dubins-Émery-Yor's Theorem of [4]. Their theorem is strongly built on Ocone's Theorem A of [20]. In Ocone's paper it is shown that a continuous local martingale M is conditionally (w.r.t. to the sigma algebra generated by $\langle M, M\rangle$) Gaussian martingale if and only if it is J-invariant. Here J-invariance means that M and $\int_0^t \alpha \, \mathrm{d}M$ have the same law for any predictable process α with range in $\{-1, 1\}$. In fact, it is proved there too that J-invariance is equivalent to H-invariance which means that it is enough to consider deterministic integrands of the form $\alpha^{(r)}(t) = \mathbf{I}_{[0,r]}(t) - \mathbf{I}_{(r,\infty)}(t)$. Moreover, Theorem B there extends the above result to càdlàg local martingales with symmetric jumps. In the sequel local martingales with these properties are called Ocone martingales.

Dubins, Émery and Yor in [4] proved that these conditions are equivalent to the independence of the *DDS* Brownian motion and the quadratic variation. Further, in this paper and the paper of Vostrikova and Yor in [36] shorter proofs with additional equivalent conditions were given in the case when M is a continuous martingale. In these references the equivalent condition of the independence of the DDS BM and $\langle M, M\rangle$ explicitly appears. Besides, in [4], the conjecture that a continuous martingale M has the same law as its Lévy transform $\hat{M} = \int \operatorname{sgn}(M) \, \mathrm{d}M$ if and only if its DDS BM and $\langle M, M\rangle$ are independent is proved to be equivalent to the conjecture that the Lévy transform is ergodic. Below we give a new, long, but elementary proof for any continuous local martingale M that the DDS BM and $\langle M, M\rangle$ are independent if and only if M is symmetrically evolving i.e. Ocone. Then, in Subsection 2.3.2 we present some remarkable properties of Ocone martingales and some of our recent results. Here, we also give a couple of examples for martingales being Ocone or non-Ocone.

2.3.1 Distributional characterization of Ocone martingales

Theorem 4. *(a) If the Wiener process $W(t)$ $(t \geq 0)$ and the non-decreasing, vanishing at 0, continuous stochastic process $C(t)$ $(t \geq 0)$ are independent, then $M(t) = W(C(t))$ is a symmetrically evolving continuous local martingale vanishing at 0, with quadratic variation C.*

(b) Conversely, if M is a symmetrically evolving continuous local martingale, then its DDS Brownian motion W and its quadratic variation $\langle M, M\rangle$ are independent processes.

Proof. To prove (a) suppose that W and C are independent. By [23, V (1.5), p. 181], $M(t) = W(C(t))$ is a continuous local martingale. For simplicity, we will use only three sets in showing that M is symmetrically evolving, i.e. equation (2.21) holds, the generalization being straightforward:

$$
\begin{aligned}
& \mathbf{P}\left\{M(s_1) \in V_1, M(t_1) - M(s) \in U_1, M(t_2) - M(s) \in U_2\right\} \\
= & \int \mathbf{P}\left\{W(x_1) \in V_1\right\} \int_{U_1} \mathbf{P}\left\{W(y_2) - W(y_1) \in U_2 - u\right\} \mathbf{P}\left\{W(y_1) - W(y_0) \in \mathrm{d}u\right\} \\
& \times \mathbf{P}\left\{C(s_1) \in \mathrm{d}x_1, C(s) \in \mathrm{d}y_0, C(t_1) \in \mathrm{d}y_1, C(t_2) \in \mathrm{d}y_2\right\} \\
= & \int \mathbf{P}\left\{W(x_1) \in V_1\right\} \int_{U_1} \mathbf{P}\left\{W(y_1) - W(y_2) \in U_2 - u\right\} \mathbf{P}\left\{W(y_0) - W(y_1) \in \mathrm{d}u\right\} \\
& \times \mathbf{P}\left\{C(s_1) \in \mathrm{d}x_1, C(s) \in \mathrm{d}y_0, C(t_1) \in \mathrm{d}y_1, C(t_2) \in \mathrm{d}y_2\right\} \\
= & \mathbf{P}\left\{M(s_1) \in V_1, M(s) - M(t_1) \in U_1, M(s) - M(t_2) \in U_2\right\},
\end{aligned}
$$

using the independence of B and C on one hand and the symmetry and independence of the increments of Brownian motion on the other hand.

For proving (b) we want to show that the sequences $\tau_m(k)$ $(k = 1, 2, \ldots)$ and $M(\tau_m(j)) - M(\tau_m(j-1))$ $(j = 1, 2, \ldots)$ are independent. Since $N_m(t)$ depends only on the number of stopping times $\tau_m(k) \leq t$, cf. (2.14), while the shrunken random walk B_m is determined by the steps $2^{-m}X_m(j) = M(\tau_m(j)) - M(\tau_m(j-1))$, cf. Lemma 2, this would imply their independence and so the independence of W and $\langle M, M\rangle$ too by Lemma 1 and 1. For this it is enough to show that with arbitrary integers $m \geq 0$, $n \geq 1$, $0 \leq k < n$, reals $t_1, \ldots, t_n$ and $\delta_1 = \pm 2^{-m}, \ldots, \delta_n = \pm 2^{-m}$ (we fix these parameters for the remaining part of the proof) one has

$$\mathbf{P}\left\{A \cap B_{\leq k} \cap B_{>k}\right\} = \mathbf{P}\left\{A \cap B_{\leq k} \cap B_{>k}^{-}\right\}, \tag{2.22}$$

where $A_{\leq k} = \bigcap_{r=1}^{k} \{\tau_m(r) \leq t_r\}$, $A_{>k}$ is similar, but with $r = k+1, \ldots n$, $A = A_{\leq k} \cap A_{>k}$, $B_{\leq k} = \bigcap_{r=1}^{k} \{M(\tau_m(r)) - M(\tau_m(r-1)) = \delta_r\}$, $B_{>k}$ is similar, but with $r = k+1, \ldots n$, $B = B_{\leq k} \cap B_{>k}$, and finally $B_{>k}^{-}$ is the same as $B_{>k}$, but each δ_j is replaced by $-\delta_j$. For, if one can reflect all δ_js for $k < j \leq n$ without changing the probability, then one has the same probability with arbitrary changed signs of δ_js too, since any such change can be reduced to a finite sequence of reflections of the above type. Let B^* be similar to B, but with arbitrarily changed signs of δ_js. Then, as we said, (2.22) implies that $\mathbf{P}\{A \cap B\} = \mathbf{P}\{A \cap B^*\}$. Since $\mathbf{P}\{B\} = \mathbf{P}\{B^*\}$ by Lemma 2, the desired independence follows.

We will prove (2.22) in several steps.

Step 1. In condition (2.20) one can replace s by an arbitrary stopping time σ adapted to the filtration $(\mathcal{F}_s^0)$: for any $u_j \geq 0$ $(1 \leq j \leq N)$,

$$\mathbf{P}\left\{F \mid \mathcal{F}_\sigma^0\right\} = \mathbf{P}\left\{F^- \mid \mathcal{F}_\sigma^0\right\}, \tag{2.23}$$

where

$$F = \bigcap_{j=1}^{N} \left\{M(u_j + \sigma) - M(\sigma) \in U_j\right\}, \tag{2.24}$$

and F^- is the same, but each U_j replaced by $-U_j$. This is somewhat similar to the optional stopping theorem, see [23, II (3.2), p.69]. Indeed, for discrete valued stopping times σ the statement is obvious, since then

$$\mathbf{P}\left\{F \mid \mathcal{F}_\sigma^0\right\} = \sum_{s_r} \mathbf{I}\left\{\sigma = s_r\right\} \mathbf{P}\left\{F \mid \mathcal{F}_{s_r}^0\right\},$$

where $\{s_r\}$ denotes the range of σ, including possibly ∞, and $\mathbf{I}\{S\}$ denotes the indicator of the set S. For every stopping time σ there exists a decreasing sequence of discrete valued stopping times σ_i

almost surely converging to σ. Let us denote the events defined according to (2.24) for σ_i by F_i and F_i^-, respectively. Further, denote the operators projecting $L^2(\Omega)$ onto its subspace of random variables measurable w.r.t. $\mathcal{F}^0_{\sigma_i}$ and $\mathcal{F}^0_\sigma$ by P_i and P, respectively. Then

$$\begin{aligned} & \|\mathbf{P}\left\{F_i \mid \mathcal{F}^0_{\sigma_i}\right\} - \mathbf{P}\left\{F \mid \mathcal{F}^0_\sigma\right\}\|_2 = \|P_i \mathbf{I}\{F_i\} - P\mathbf{I}\{F\}\|_2 \\ \leq\ & \|P_i(\mathbf{I}\{F_i\} - \mathbf{I}\{F\})\|_2 + \|P_i\mathbf{I}\{F\} - P\mathbf{I}\{F\}\|_2 \\ \leq\ & \|\mathbf{I}\{F_i\} - \mathbf{I}\{F\}\|_2 + \|\mathbf{E}(\mathbf{I}\{F\} \mid \mathcal{F}^0_{\sigma_i}) - \mathbf{E}(\mathbf{I}\{F\} \mid \mathcal{F}^0_\sigma)\|_2, \end{aligned}$$

which goes to 0 as $i \to \infty$. Here we used that $\mathbf{E}(\mathbf{I}\{F\} \mid \mathcal{F}^0_{\sigma_i})$ is a bounded, reversed-time martingale converging to $\mathbf{E}(\mathbf{I}\{F\} \mid \mathcal{F}^0_\sigma)$. Hence for any $\epsilon > 0$,

$$\begin{aligned} \|\mathbf{P}\left\{F \mid \mathcal{F}^0_\sigma\right\} - \mathbf{P}\left\{F^- \mid \mathcal{F}^0_\sigma\right\}\|_2 \leq\ & \|\mathbf{P}\left\{F \mid \mathcal{F}^0_\sigma\right\} - \mathbf{P}\left\{F_i \mid \mathcal{F}^0_{\sigma_i}\right\}\|_2 \\ & + \|\mathbf{P}\left\{F^- \mid \mathcal{F}^0_\sigma\right\} - \mathbf{P}\left\{F_i^- \mid \mathcal{F}^0_{\sigma_i}\right\}\|_2 < \epsilon, \end{aligned}$$

if i is large enough. This shows that the left hand side of the inequality is zero, so (2.23) holds.

Step 2. Then for arbitrary reals $0 \leq u_j < v_j$ and Borel sets U_j $(1 \leq j \leq N)$ we have

$$\mathbf{P}\left\{G \mid \mathcal{F}^0_\sigma\right\} = \mathbf{P}\left\{G^- \mid \mathcal{F}^0_\sigma\right\},$$

where

$$G = \bigcap_{j=1}^{N} \{M(v_j + \sigma) - M(u_j + \sigma) \in U_j\}, \tag{2.25}$$

and G^- is the same, but each U_j is replaced by $-U_j$. For simplicity we prove this only for two factors, the general case being similar:

$$\begin{aligned} & \mathbf{P}\left\{M(v_1+\sigma) - M(u_1+\sigma) \in U_1, M(v_2+\sigma) - M(u_2+\sigma) \in U_2 \mid \mathcal{F}^0_\sigma\right\} \\ =\ & \int \mathbf{I}\{x_1 - x_2 \in U_1\}\, \mathbf{I}\{x_3 - x_4 \in U_2\}\, \mathbf{P}\{M(v_1+\sigma) - M(\sigma) \in \mathrm{d}x_1, \\ & M(u_1+\sigma) - M(\sigma) \in \mathrm{d}x_2, M(v_2+\sigma) - M(\sigma) \in \mathrm{d}x_3, M(u_2+\sigma) - M(\sigma) \in \mathrm{d}x_4 \mid \mathcal{F}^0_\sigma\} \\ =\ & \int \mathbf{I}\{x_1 - x_2 \in U_1\}\, \mathbf{I}\{x_3 - x_4 \in U_2\}\, \mathbf{P}\{M(\sigma) - M(v_1+\sigma) \in \mathrm{d}x_1, \\ & M(\sigma) - M(u_1+\sigma) \in \mathrm{d}x_2, M(\sigma) - M(v_2+\sigma) \in \mathrm{d}x_3, M(\sigma) - M(u_2+\sigma) \in \mathrm{d}x_4 \mid \mathcal{F}^0_\sigma\} \\ =\ & \mathbf{P}\left\{M(u_1+\sigma) - M(v_1+\sigma) \in U_1, M(u_2+\sigma) - M(v_2+\sigma) \in U_2 \mid \mathcal{F}^0_\sigma\right\}. \end{aligned}$$

Step 3. Let $\Delta\tau_m(i) = \tau_m(i) - \tau_m(i-1)$ and $a \in [0, \infty)$. Consider the event

$$\begin{aligned} S(a) &= \{\Delta\tau_m(k+1) \geq a\} \\ &= \left\{\inf\left\{u > 0 : |M(u + \tau_m(k)) - M(\tau_m(k))| \geq 2^{-m}\right\} > a\right\} \\ &= \left\{\sup_{0<u\leq a} \{|M(u + \tau_m(k)) - M(\tau_m(k))|\} < 2^{-m}\right\} \\ &= \bigcap_{0<u\leq a} \left\{|M(u + \tau_m(k)) - M(\tau_m(k))| < 2^{-m}\right\}. \end{aligned} \tag{2.26}$$

Introduce the set of dyadic numbers $D_l = \left\{r2^{-l} : r \in \mathbb{Z}\right\}$ $(l \geq 0)$, and the events

$$S_{j,l}(a) = \bigcap_{q \in D_l, 0<q\leq a} \left\{|M(q + \tau_m(k)) - M(\tau_m(k))| \leq 2^{-m} - 2^{-j}\right\}. \tag{2.27}$$

when $j \geq m$. For j fixed, $(S_{j,l})_{l=0}^{\infty}$ is a decreasing sequence of events. Take

$$S_j(a) = \bigcap_{l=0}^{\infty} S_{j,l}(a).$$

Since $S_{j,l}(a)$ is increasing with growing j, so is $S_j(a)$. Put

$$S^*(a) = \bigcup_{j=m}^{\infty} S_j(a) = \bigcup_{j=m}^{\infty} \bigcap_{l=0}^{\infty} S_{j,l}(a).$$

We want to show that $S^*(a) = S(a)$, where the latter is defined by (2.26).

First fix an $\omega \in S(a)$. (We suppress ω in the notations below.) Then with this ω,

$$\sup_{0<u\leq a} \left\{|M(u + \tau_m(k)) - M(\tau_m(k))| < 2^{-m}\right\} =: s < 2^{-m}.$$

If $j > m$ is such that $2^{-j} < 2^{-m} - s$, then $\omega \in S_{j,l}(a)$ for any $l \geq 0$. So $\omega \in S_j(a)$ for each j large enough, consequently, $\omega \in S^*(a)$.

Second, fix an $\omega \notin S(a)$. Then there exists a real $u_0 \leq a$ (depending on ω) so that $|M(u_0+\tau_m(k)) - M(\tau_m(k))| = 2^{-m}$. Since the path of M is a continuous function, for any $j \geq m$ there exists an $l \geq 0$ and $q \in D_l, 0 < q \leq a$ such that $|M(q+\tau_m(k)) - M(\tau_m(k))| > 2^{-m} - 2^{-j}$. That is, $\omega \notin S_j(a)$ if $j \geq m$, thus $\omega \notin S^*(a)$.

In other words, we have proved above that

$$\begin{aligned} \{\Delta\tau_m(k+1) > a\} = S(a) &= \lim_{j\to\infty} \lim_{l\to\infty} S_{j,l}(a) \\ = \lim_{j\to\infty} \lim_{l\to\infty} \bigcap_{q\in D_l, 0<q\leq a} &\left\{|M(q + \tau_m(k)) - M(\tau_m(k))| \leq 2^{-m} - 2^{-j}\right\}. \end{aligned}$$

Consequently, any event $\{\Delta\tau_m(k+1) > a\} = S(a)$ can be written in terms of monotonic sequences of intersections of finitely many events of the form

$$\{|M(q + \tau_m(k)) - M(\tau_m(k))| \leq c\} \qquad (c \geq 0).$$

Moreover, such an approximation can be applied to $\{\Delta\tau_m(k+1) \in (a, b]\} = S(a) \setminus S(b)$ as well, with any $0 \leq a < b$.

Step 4. First, Steps 2 and 3 imply that for any $a \geq 0$,

$$\mathbf{P}\left\{G \cap S_{j,l}(a) \mid \mathcal{F}^0_{\tau_m(k)}\right\} = \mathbf{P}\left\{G^- \cap S_{j,l}(a) \mid \mathcal{F}^0_{\tau_m(k)}\right\},$$

because of the absolute value in definition (2.27) of $S_{j,l}(a)$. Throughout Step 4 G and G^- are defined according to (2.25) with $\sigma = \tau_m(k)$, but otherwise with arbitrary parameters, possibly different from case to case. Then taking limit as $j \to \infty$ and $l \to \infty$ it follows from Steps 2 and 3 that

$$\begin{aligned} &\mathbf{P}\left\{G \cap \{\Delta\tau_m(k+1) > a_{k+1}\} \mid \mathcal{F}^0_{\tau_m(k)}\right\} \\ = \;&\mathbf{P}\left\{G^- \cap \{\Delta\tau_m(k+1) > a_{k+1}\} \mid \mathcal{F}^0_{\tau_m(k)}\right\}. \end{aligned}$$

We want to extend this symmetry property by induction over $i = k+1, \ldots, n+1$. Taking arbitrary reals $a_i \geq 0$ and integers $l \geq 0$, $r_i > 0$, $(k+1 \leq i \leq n+1)$ define the events

$$B_i = \bigcap_{p=k+1}^{i} \{M(\tau_m(p)) - M(\tau_m(p-1)) = \delta_p\},$$

$$H_i = \bigcap_{p=k+1}^{i} \{\Delta\tau_m(p) > a_p\},$$

$$K_{i,l}(r) = \bigcap_{p=k+1}^{i} \left\{\Delta\tau_m(p) \in \left((r_p-1)2^{-l}, r_p 2^{-l}\right]\right\}$$

$$L_{i,l}(r) = \left\{\delta_i \left(M(r_i 2^{-l} + \cdots + r_{k+1}2^{-l} + \tau_m(k)) - M(r_{i-1}2^{-l} + \cdots + r_{k+1}2^{-l} + \tau_m(k))\right) > 0\right\},$$

and B_i^-, $L_{i,l}^-(r)$ similarly, but multiplying each δ_p by (-1). Suppose that we have already proved that

$$\mathbf{P}\left\{G \cap B_{i-1} \cap H_i \mid \mathcal{F}^0_{\tau_m(k)}\right\} = \mathbf{P}\left\{G^- \cap B^-_{i-1} \cap H_i \mid \mathcal{F}^0_{\tau_m(k)}\right\},$$

where $B_k = \Omega$. Define the following event as a generalization of (2.27):

$$S_{i,j,l}(a,r) = \bigcap_{q\in D_l, 0<q\leq a} \left\{\left|M(q + r_i 2^{-l} + \cdots + r_{k+1}2^{-l} + \tau_m(k)) - M(r_{i-1}2^{-l} + r_{k+1}2^{-l} + \tau_m(k))\right| \leq 2^{-m} - 2^{-j}\right\},$$

where $j \geq m$. Then by the induction hypothesis we get that

$$\mathbf{P}\left\{G \cap B_{i-1} \cap H_i \cap \bigcup_{r_{k+1},\ldots,r_i=1}^{2^{2l}+1} K_{i,l}(r) \cap L_{i,l}(r) \cap S_{i,j,l}(a_{i+1}, r) \mid \mathcal{F}^0_{\tau_m(k)}\right\}$$
$$= \mathbf{P}\left\{G^- \cap B^-_{i-1} \cap H_i \cap \bigcup_{r_{k+1},\ldots,r_i=1}^{2^{2l}+1} K_{i,l}(r) \cap L^-_{i,l}(r) \cap S_{i,j,l}(a_{i+1}, r) \mid \mathcal{F}^0_{\tau_m(k)}\right\},$$

where we agree that when $r_p = 2^{2l}+1$, the interval $\left((r_p-1)2^{-l}, r_p 2^{-l}\right]$ in the definition of $K_{i,l}(r)$ is replaced by $\left((r_p-1)2^{-l}, \infty\right] = (2^l, \infty]$. Notice here that the events in $K_{i,l}$ can be written in terms of a difference of events appearing in H_i, while the events in $L_{i,l}$ and $S_{i,j,l}$ are both of the type appearing in G, though $S_{i,j,l}$ is not affected by reflections because of the absolute values in its definition. Then taking limit as $j \to \infty$ and $l \to \infty$ it follows that

$$\mathbf{P}\left\{G \cap B_i \cap H_{i+1} \mid \mathcal{F}^0_{\tau_m(k)}\right\} = \mathbf{P}\left\{G^- \cap B^-_i \cap H_{i+1} \mid \mathcal{F}^0_{\tau_m(k)}\right\}.$$

This completes the induction.

Comparing the notations introduced in this step with the ones introduced above, observe that $B_{>k} = B_n$ and $H_{>k} = H_n = H_{n+1}$, if $a_{n+1} = 0$. Thus one obtains that

$$\mathbf{P}\left\{B_{>k} \cap H_{>k} \mid \mathcal{F}^0_{\tau_m(k)}\right\} = \mathbf{P}\left\{B^-_{>k} \cap H_{>k} \mid \mathcal{F}^0_{\tau_m(k)}\right\}.$$

Step 5. The result of Step 4 implies that

$$\begin{aligned}
&\mathbf{P}\left\{A_{>k} \cap B_{>k} \mid \mathcal{F}^0_{\tau_m(k)}\right\} \\
= \quad & \int \mathbf{I}\left\{\tau_m(k) + x_{k+1} \le t_{k+1}, \dots, \tau_m(k) + x_{k+1} + \dots + x_n \le t_n\right\} \\
& \times \mathbf{P}\left\{B_{>k} \cap \{\Delta\tau_m(k+1) \in \mathrm{d}x_{k+1}, \dots, \Delta\tau_m(n) \in \mathrm{d}x_n\} \mid \mathcal{F}^0_{\tau_m(k)}\right\} \\
= \quad & \int \mathbf{I}\left\{\tau_m(k) + x_{k+1} \le t_{k+1}, \dots, \tau_m(k) + x_{k+1} + \dots + x_n \le t_n\right\} \\
& \times \mathbf{P}\left\{B^-_{>k} \cap \{\Delta\tau_m(k+1) \in \mathrm{d}x_{k+1}, \dots, \Delta\tau_m(n) \in \mathrm{d}x_n\} \mid \mathcal{F}^0_{\tau_m(k)}\right\} \\
= \quad & \mathbf{P}\left\{A_{>k} \cap B^-_{>k} \mid \mathcal{F}^0_{\tau_m(k)}\right\}.
\end{aligned}$$

Step 6. Finally, it follows from Step 5 that

$$\begin{aligned}
\mathbf{P}\{A \cap B\} &= \mathbf{P}\{A_{\le k} \cap A_{>k} \cap B_{\le k} \cap B_{>k}\} \\
&= \mathbf{E}\left(\mathbf{E}\left(\mathbf{I}\{A_{\le k} \cap B_{\le k}\}\,\mathbf{I}\{A_{>k} \cap B_{>k}\} \mid \mathcal{F}^0_{\tau_m(k)}\right)\right) \\
&= \mathbf{E}\left(\mathbf{I}\{A_{\le k} \cap B_{\le k}\}\,\mathbf{P}\left\{A_{>k} \cap B_{>k} \mid \mathcal{F}^0_{\tau_m(k)}\right)\right\} \\
&= \mathbf{E}\left(\mathbf{I}\{A_{\le k} \cap B_{\le k}\}\,\mathbf{P}\left\{A_{>k} \cap B^-_{>k} \mid \mathcal{F}^0_{\tau_m(k)}\right)\right\} \\
&= \mathbf{P}\left\{A_{\le k} \cap A_{>k} \cap B_{\le k} \cap B^-_{>k}\right\}.
\end{aligned}$$

This proves (2.22), and so completes the proof of the theorem. □

2.3.2 Properties of Ocone martingales

In this subsection we give some equivalent conditions for martingales being Ocone martingale. Originally, Ocone proved the equivalence of the condition (ii-iv). The equuivalence of the parts (i-v-vi) are due to Émery-Dubins-Yor. We remark that Ocone's original setting is more general: he deals with local martingales not only in continuous but in the cadlag case.

Theorem D. *Let M be a continuous martingale with natural filtration $\mathcal{F} = (\mathcal{F}_t)$. The following five statement are equivalent:*

(i) *the DDS-Brownian motion β^M of M and $\langle M\rangle$ are independent;*

(ii) *conditionaly on $\langle M\rangle$ M is Gaussian martingale;*

(iii) *for every $\mathcal{F}$-predictable process H taking values in $\{-1, 1\}$, the two pairs of processes have the same law*

$$\left(\int H\,\mathrm{d}M, \langle M\rangle\right) \overset{\mathrm{d}}{=} (M, \langle M\rangle);$$

(iv) *for every deterministic function h of the form $\mathbf{I}_{[0,a]} - \mathbf{I}_{(a,\infty)}$, the martingale $\int h\,\mathrm{d}M$ has the same law as M;*

(v) *for every $\mathcal{F}$-predictable process H measurable for the σ-field $\mathcal{B}(\mathbb{R}_+) \otimes \sigma(\langle M\rangle)$ and such that $\int_0^\infty H_s^2\,\mathrm{d}\langle M\rangle_s < \infty$ a.s.,*

$$\mathbf{E}\left[\exp\left(i\int_0^\infty H_s\,\mathrm{d}M_s\right)\Big|\langle M\rangle\right] = \exp\left(-\frac{1}{2}\int_0^\infty H_s^2\,\mathrm{d}\langle M\rangle_s\right);$$

(vi) for every deterministic function h of the form $\sum_{j=1}^{n} c_j \mathbf{I}_{[0,a_j]}$,

$$\mathbf{E}\left[\exp\left(i\int_0^\infty h(s)\,\mathrm{d}M_s\right)\right] = \mathbf{E}\left[\exp\left(-\frac{1}{2}\int_0^\infty h^2(s)\,\mathrm{d}\langle M\rangle_s\right)\right] .$$

Rather interestingly, in the theory of Ocone martingales there is a ten years old conjecture originally introduced by Yor [4]. It says that a divergent martingale M is Ocone martingale if and only if the Lévy transform $\widehat{M} = \int \operatorname{sign}(M)\,\mathrm{d}M$ of the martingale M has the same law as M. If it were true it would mean that one could formally compress the conditions of Theorem D into only one condition.

Dubins, Émery and Yor proved that this conjecture is equivalent to an other conjecture known since the late 70's. This conjecture is also about the Lévy transformation but it deals with the Brownian case. If B is a standard Brownian motion started at 0 and $\mathcal{L}$ is its local time at 0 then the Lévy characterization says that

$$\int \operatorname{sign}(B)\,\mathrm{d}B = |B| - \mathcal{L}(B)$$

is also a Brownian motion on the measure space (W, μ) where W is the Wiener space and μ is the Wiener measure. The conjecture says that the transformation $\mathcal{T} : B \to \int \operatorname{sign}(B)\,\mathrm{d}B$ is ergodic on the above introduced measure space.

A result by Dubins and Smorodinsky [6, 1992] increases the plausibility of this conjecture. They established that the discrete version of the Lévy transformation taking effect on the standard symmetric random walk is ergodic on the corresponding measure space.

To end this subsection, we want to mention an interesting property of the transformation $\mathcal{T}$. The processes $\mathcal{T}^n B$, $n \geq 0$, are pairwise weakly orthogonal in the following sense [23]:

Two local martingales M and N are said to be weakly orthogonal if $\mathbf{E} M_s N_t = 0$ for every s and $t \geq 0$. This condition is equivalent to that of for all $s \geq 0$ $\mathbf{E}\langle M, N\rangle = 0$. In our present case this means that the martingales $\mathcal{T}^n M_t$ and $\mathcal{T}^k M_t$ are not jointly Gaussian.

Proposition 1. *The martingales $\mathcal{T}^n W$, $(n \geq 0)$ are pairwise weakly orthogonal. More precisely, for all $s, t \geq 0$ and $n \neq k$ non-negative integer we have*

$$\mathbf{E}\mathcal{T}^n W_s \mathcal{T}^k W_t = 0 . \tag{2.28}$$

Proof. It is enough to prove the statement for $\mathbf{E}\mathcal{T}^n W_s \mathcal{T} W_t = 0$. For simplicity, we will prove it for the case $\mathbf{E} W_s \mathcal{T} W_t = 0$.

Let us consider both W_s and $\mathcal{T} W_t$ on the interval $[0, T]$, $0 \leq s, t \leq T$. Let us introduce the notations $I_s^t = \int_s^t \operatorname{sign}(W_u)\,\mathrm{d}W_u$ and $I(F)_s^t = \sum_{t_k \in F \cap [s,t]} \operatorname{sign}(W_{t_k})(W_{t_{k+1}} - W_{t_k})$ for arbitrary partition F of $[0, T]$.

First, there exists a sequence of simple processes converging to $(\operatorname{sign}(W_t), t \geq 0)$ in $L_2(\mathbf{P} \times \lambda)$ on $[0, T]$. We follow Itô's method, see [17].

Let $X = (X_t, 0 \leq t \leq T)$ be an $(\mathcal{F}_t^W)$-adapted process with finite L_2 norm and $\varphi_n(u) = \lfloor \frac{u}{T} 2^n \rfloor \frac{T}{2^n}$. Itô showed that there is a subsequence (n_k) such that $X^{n_k} = (X_{\varphi_{n_k}(t-u)+u}, 0 \leq t \leq T) \to X$ in $L_2(\mathbf{P} \times \lambda)$ as k tends to infinity for Lebesgue almost every u from $[0, 1]$.

For the present case $X_t = \operatorname{sign}(W_t)$, it means that for arbitrary $s < t$ $\mathbf{E}(I(F_k)_s^t - I_s^t)^2 \to 0$ $(k \to \infty)$ if the partition F_k contains the points which can be obtained via Itô's method. Since this and the Hölder inequality we get the following estimate for all s, t:

$$\mathbf{E} W_s (I(F_k)_0^t - I_0^t) \leq \sqrt{\mathbf{E} W_s^2} \sqrt{\mathbf{E}(I(F_k)_0^t - I_0^t)^2} \to 0 \quad (k \to \infty), \tag{2.29}$$

from which one gets $\mathbf{E}W_sI(F_k)_0^t \to \mathbf{E}W_sI_0^t$ as k tends to infinity.

For proving (2.28) we first deal with the case $t \leq s$.
Using that the Brownian motion has independent increment and the fact that $\mathbf{E}\mathrm{sign}W_{t_k} = 0$, we have

$$\begin{aligned}\mathbf{E}W_sI(F_k)_0^t &= \sum_{t_{i+1}\leq t, t_i\in F_k} \mathbf{E}W_s\mathrm{sign}(W_{t_i})(W_{t_{i+1}} - W_{t_i})\\ &= \sum_{t_{i+1}\leq t, t_i\in F_k} \mathbf{E}\left(\left(W_{t_i} + (W_{t_{i+1}} - W_{t_i}) + W_s - W_{t_{i+1}}\right)\mathrm{sign}(W_{t_i})(W_{t_{i+1}} - W_{t_i})\right)\\ &= \sum_{t_{i+1}\leq t, t_i\in F_k} \mathbf{E}\mathrm{sign}(W_{t_i})(W_{t_{i+1}} - W_{t_i})^2 .\end{aligned}$$

Now, take into consideration that for all $0 \leq t_i < t_{i+1}$, $\mathrm{sign}(W_{t_i})$ and $(W_{t_{i+1}} - W_{t_i})^2$ are independent.

Hence, $\mathbf{E}W_sI_0^t = 0$.

In the case $s < t$ we have decomposition $\mathbf{E}W_sI_0^t = \mathbf{E}W_sI_0^s + \mathbf{E}W_sI_s^t$. By the previous paragraph the first term is equal to 0. By (2.29) and by the equation

$$\begin{aligned}\mathbf{E}W_sI(F_k)_s^t &= \sum_{s\leq t_i<t_{i+1}\leq t, t_i\in F_k} \mathbf{E}W_s\mathrm{sign}(W_{t_i})(W_{t_{i+1}} - W_{t_i})\\ &= \sum_{s\leq t_i<t_{i+1}\leq t, t_i\in F_k} \mathbf{E}W_s\mathrm{sign}(W_{t_i})\mathbf{E}(W_{t_{i+1}} - W_{t_i}) = 0\end{aligned}$$

so the second term is also 0. □

2.3.3 Some examples of Ocone and non-Ocone martingales

In this subsection we present some remarkable properties of Ocone martingales and some of our recent results. Here, we also give a couple of examples for martingales being Ocone or non-Ocone.

First, let us recall some definitions and lemmas which will be used in this subsection.

Definition 1. *A continuous local martingale M such that $\langle M\rangle_\infty = \infty$ is said to be pure if, calling B its DDS Brownian motion, $\langle M\rangle_t$ is $\mathcal{F}_\infty^B$ measurable for every t or equivalently*

$$\mathcal{F}_\infty^M = \mathcal{F}_\infty^B.$$

Lemma E. *(Vostrikova-Yor [36, Proposition]) Suppose that M is Ocone martingale. M enjoys the martingale representation property iff $\langle M\rangle$ is a deterministic process.*

Lemma F. *[23, Section 4, Chapter V.] A pure local martingale is extremal so it enjoys the martingale representation property.*

Using these lemmas we can prove the following simple result.

Proposition 2. *Let M be a continuous Ocone martingale. Suppose $\langle M\rangle_\infty = \infty$ and M is of the form*

$$M_t = x + \int_0^t \sigma(M_s)\,\mathrm{d}\beta_s, \tag{2.30}$$

where σ is a nowhere vanishing function and β is a Brownian motion. Then M is Gaussian martingale.

Proof. By [23, Chapter V. (1.11) Proposition] if a continuous local martingale satisfies the differential equation (2.30) then M is measurable with respect to $\mathcal{F}_\infty^B$ where B is the DDS Brownian motion of M. So this martingale is pure which implies that it enjoys the martingale representation property (Lemma F).

By Lemma E, M is a martingale with deterministic quadratic variation process. Martingales with this property are Gaussian. □

Van Zanten proved the following interesting theorem.

Theorem E. *([34]) Let M be a martingale with bounded jumps and let a_n, b_n be sequences of positive numbers, both increasing to infinity. For each n, define the rescaled martingale M^n by*

$$M_t^n = \frac{1}{\sqrt{n}} M_{b_n t}.$$

Then the following statements hold:

1. *If M^n converges weakly to some process N in $D(0,\infty)$ then the N is necessarily a continuous Ocone martingale.*
2. *Let N be a continuous Ocone martingale. Then M^n converges to N in $D(0,\infty)$ if and only if the quadratic variation sequence $\langle M^n\rangle$ converges to $\langle N\rangle$ in $D(\S,\infty)$.*

Example A. Using this theorem he proved that the martingales

$$W_t^+ = \int_0^t \mathbf{I}_{(0,\infty)}(W_s)\,\mathrm{d}W_s, \qquad W_t^- = \int_0^t \mathbf{I}_{(-\infty,0)}(W_s)\,\mathrm{d}W_s$$

are non-Gaussian Ocone martingales. Indeed, the are Ocone martingales because of the self-similarity of the Brownian motion. They are non-Gaussian since the cited theorem of Vostrikova-Yor [36, Proposition] in the proof of Theorem 2

Example B. ([36]). Let (B_t, C_t), $t \geq 0$ a planar Brownian motion. The process

$$\mathcal{A}_t = \frac{1}{2}\int_0^\infty (C_s\,\mathrm{d}B_s - B_s\,\mathrm{d}C_s) \tag{2.31}$$

is an example of Ocone martingale. This assertion is a consequence of the following general theorem which was inspired by Marc Yor.

Theorem 5. *Let $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ be a regular function. Denote the adjoint of its derivative by $\Psi(\mathbf{x}) = (\Phi')^T(\mathbf{x})$ and suppose that the following conditions hold:*

$$\Psi(\mathbf{x})\Phi(\mathbf{x}) = \mathbf{x} \ \textit{ and } \mathbf{x}\cdot\Phi(\mathbf{x}) = 0 \textit{ for any } \mathbf{x}\in\mathbb{R}^d \tag{2.32}$$

where $\cdot$ stands for the usual scalar product in $\mathbb{R}^d$. If B is a standard d-dimensional Brownian motion then the martingale

$$M_t = \int_0^t \Phi(B_s)\cdot\mathrm{d}B_s \tag{2.33}$$

is an Ocone martingale.

Corollary 1. *If $\Phi(\mathbf{x}) = A\mathbf{x}$ where A is a regular matrix then the conditions above are equivalent with the conditions that A is orthogonal and anti-symmetric.*

Proof. The derivative of the function $\Phi : \mathbf{x} \mapsto A\mathbf{x}$ is $\Psi : \mathbf{x} \mapsto A\mathbf{x}$. So the first condition in (2.32) is equivalent to that of $A^T A = Id$. The second condition can be written in the form $\mathbf{x}^T A \mathbf{x} = 0$ for all $\mathbf{x} \in \mathbb{R}^d$, in other words A is anti-symmetric. □

In the light of Corollary 1, the martingale in Example B (2.31) can be written in the following form

$$\mathcal{A}_t = \int_0^t AB_s \cdot \mathrm{d}B_s,$$

where

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

is clearly orthogonal and anti-symmetric.

We remark that matrices with the required properties: regularity, orthogonality, anti-symmetricity are available in even order dimension. If the matrix is in $\mathbb{R}^{d\times d}$ where d is odd and posses the properties above then the matrix is a direct sum of a one dimensional null matrix and a matrix of even dimension with the listed properties.

Proof of Theorem 5. We will use the theorem of Yor and Vostrikova [36, p 426, Theorem 3']:
Let us consider a $(\mathcal{F}_t)$-martingale (M_t) and another martingale (N_t) which is pure, i.e. $N_t = \gamma_{\langle N\rangle_t}$, $t \geq 0$, with $(\langle N\rangle_t, t \geq 0)$ measurable with respect to the σ-field $\mathcal{F}^\gamma = \sigma\{\gamma_u, u \geq 0\}$ of the DDS-Brownian motion γ. Suppose that $\langle M\rangle$ is measurable with respect to $\mathcal{F}^N$.
Then $(M_t, t \geq 0)$ is an Ocone martingale as soon as N and M are orthogonal.

The quadratic variation of the martingale M in (2.33) is $\langle M\rangle = \int_0^\cdot |\Phi(B_s)|^2 \,\mathrm{d}s$. The filtration of $\langle M\rangle$ is that of $|\Phi(B_t)|^2 = \phi_1^2(B_t) + \cdots + \phi_d^2(B_t)$. Using Itô's formula one gets

$$|\Phi(B_t)|^2 = \sum_{i=1}^d \int_0^t \sum_{j=1}^d 2\left(\phi_i \partial_j \phi_i\right)(B_s)\,\mathrm{d}B_{j,s} + \frac{1}{2}\sum_{i=1}^d \int_0^t \sum_{j=1}^d \partial_j^2 \phi_i^2(B_s)\,\mathrm{d}s.$$

The first term is equal to

$$2\int_0^t \Psi(B_s) \circ \Phi(B_s) \cdot \mathrm{d}B_s = 2\sum_i \int_0^t B_{i,s}\,\mathrm{d}B_{i,s} =: N_t$$

by using the first condition in (2.32).
The second term can be written in the following form:

$$\frac{1}{2}\int_0^t \sum_{i=1}^d \sum_{j=1}^d \left(\partial_j \left(2\phi_i\, \partial_j \phi_i\right)\right)(B_s)\,\mathrm{d}s = \int_0^t \left(\sum_{j=1}^d \partial_j \left(\sum_{i=1}^d \phi_i\, \partial_j \phi_i\right)\right)(B_s)\,\mathrm{d}s$$

$$= \int_0^t \left(\sum_{j=1}^d \partial_j \pi_j\right)(B_s)\,\mathrm{d}s = td,$$

where $\pi_j(\mathbf{x}) = x_j$ the jth coordinate of the vector $\mathbf{x}$. The second equality follows from the first part of the condition (2.32). Summarizing these results we get that $|\Phi(B_t)|^2$ and so $\langle M\rangle_t$ is N_t measurable.

We recall that N_t is pure martingale. Consider the following form of $\langle N \rangle$:

$$\langle N \rangle_t = \int_0^t \left(\sum_{i=1}^d B_i^2(s) \right) \mathrm{d}s = \int_0^t \left(N_s^2 + sd \right) \mathrm{d}s.$$

Let T be the quasi-inverse of $\langle N \rangle$. Using the previous equation we have the following

$$t = \langle N \rangle_{T_t} = \int_0^{T_t} \left(\sum_{i=1}^d B_i^2(s) \right) \mathrm{d}s = \int_0^{T_t} \left(N_s^2 + sd \right) \mathrm{d}s.$$

By applying a time-change on the integral one gets

$$t = \int_0^t \left(N_{T_s}^2 + T_s d \right) \mathrm{d}T_s = \int_0^t \left(\gamma_s^2 + T_s d \right) \mathrm{d}T_s$$

where γ is the DDS-Brownian motion of N. Hence,

$$\mathrm{d}T_t = \frac{1}{(\gamma_t^2 + T_t d)} \mathrm{d}t,$$

that is, T and so $\langle M \rangle$ is γ measurable. We have used that $\langle N \rangle$ is strictly increasing and that T is continuous.

We will prove that M and N are orthogonal martingales. This proves our theorem. Using the second condition in (2.32) one finally gets

$$\langle M, N \rangle_t = \int_0^t \Phi(B_s) \cdot (\Psi(B_s) \circ \Phi(B_s)) \, \mathrm{d}s = 0.$$

□

We remark that the cited theorem of Vostrikova and Yor can be easily proved by showing that the DDS-Brownian motions of M and N, β and γ, are independent processes. These two Brownian motions are independent iff every $f, g \in L_2(\mathbb{R}_+)$ the following equation satisfies

$$\mathbf{E}\left(\exp\left\{ \int_0^\infty f(s) \, \mathrm{d}\beta_s - \frac{1}{2} \int_0^\infty f^2(s) \, \mathrm{d}s \right\} \exp\left\{ \int_0^\infty g(s) \, \mathrm{d}\gamma_s - \frac{1}{2} \int_0^\infty g^2(s) \, \mathrm{d}s \right\} \right) = 1$$

or with other words

$$\mathbf{E}\left(\mathcal{E}_\infty^{f,\beta} \cdot \mathcal{E}_\infty^{g,\gamma} \right) = 1,$$

with the notation

$$\mathcal{E}_t^{f,\beta} = \exp\left\{ \int_0^t f(s) \, \mathrm{d}\beta_s - \frac{1}{2} \int_0^t f^2(s) \, \mathrm{d}s \right\},$$

$\mathcal{E}_t^{g,\gamma}$ is defined similarly. Time-changing the exponential martingales $\mathcal{E}_t^{f,\beta}$ and $\mathcal{E}_t^{g,\gamma}$ with the time-change processes $\langle M \rangle$ and $\langle N \rangle$ respectively we get that we have to prove the following

$$\mathbf{E}\left(\mathcal{E}_\infty^{f(\langle M \rangle),M} \cdot \mathcal{E}_\infty^{g(\langle N \rangle),N} \right) = 1$$

since

$$\mathbf{E}\left(\mathcal{E}_\infty^{f(\langle M \rangle),M} \cdot \mathcal{E}_\infty^{g(\langle N \rangle),N} \right) = \mathbf{E}\left(\mathcal{E}_\infty^{f,\beta} \cdot \mathcal{E}_\infty^{g,\gamma} \right).$$

This will be done by showing

$$\mathbf{E}\left(\mathcal{E}_t^{f(\langle M \rangle),M} \cdot \mathcal{E}_t^{g(\langle N \rangle),N} \right) = 1 \tag{2.34}$$

for all $t > 0$.

By using Itô's formula on exponential martingales we get

$$\mathcal{E}_t^{f(\langle M\rangle),M} \cdot \mathcal{E}_t^{g(\langle N\rangle),N} = 1 + \int_0^t \mathcal{E}_s^{g(\langle N\rangle),N} \, \mathrm{d}\mathcal{E}_s^{f(\langle M\rangle),M} + \int_0^t \mathcal{E}_s^{f(\langle M\rangle),M} \, \mathrm{d}\mathcal{E}_s^{g(\langle N\rangle),N} + \frac{1}{2}\int_0^t \mathcal{E}_s^{f(\langle M\rangle),M} \cdot \mathcal{E}_s^{g(\langle N\rangle),N} f(\langle M\rangle_s) g(\langle N\rangle_s) \, \mathrm{d}\langle M, N\rangle_s.$$

The second and the third term are martingales so their expectation is 0 for all t the last term is also zero since M and N are orthogonal so the expectation in (2.34) is equals to 1.

Finally, we show two martingales which are non-Ocone martingales (The examples are taken from [36]).

Example C. Let B be a Brownian motion. The martingale

$$M_t = \int_0^t B_s \, \mathrm{d}B_s, \ t \geq 0$$

is non-Ocone. Its quadratic variation process $\langle M\rangle_t = \int_0^t B_s^2 \, \mathrm{d}s$ is non deterministic which would contradict that M is pure (see Lemma E and F).

Example D. Let (C_s, B_s), $s \geq 0$ be a planar Brownian motion. The martingale

$$\pi_t = B_t C_t = \int_0^t (C_s \, \mathrm{d}B_s + B_s \, \mathrm{d}C_s) \ t \geq 0$$

is non-Ocone.

The quadratic variation process of this martingale is $\langle \pi\rangle_t = \int_0^t R_s^2 \, \mathrm{d}s = \int_0^t (B_s^2 + C_s^2) \, \mathrm{d}s$, $t \geq 0$. The σ-algebra generated by $\langle \pi\rangle$ is that of $(R_t, \ t \geq 0)$. By the definition of π we have $|2B_tC_t| \leq R_t^2$. Since this inequality, conditionally on $(R_t, \ t \geq 0)$ (B_tC_t) is bounded so it can not be Gaussian.

2.4 Approximation of local time, excursion, meander and bridge

The Brownian motion construction in our focus is especially suitable for obtaining results on the approximation of local time of the Brownian motion. Indeed, if two points of an approximating random walk are at the same altitude, they will remain at the same altitude after refinements forever. However, this may be disadvantageous for approximating the other three processes, the excursion, bridge and meander because the first 0 hit changes randomly from one approximation to the next one if we look into the excursion.

We also present a different, rather combinatorial, construction for BM, excursion, bridge, meander which is due to Phillipe Marchal. His algorithms allow to generate directly the excursion, the bridge and the meander. Moreover, many self-similarity properties are easily seen from this construction, and one can also recover various distributions of BM from the urn schemes that are embedded in the construction. In his paper he also gave a local time approximating algorithm but this one is not so natural as in the first three cases

2.4.1 Approximation of local time

Let $\mathcal{L}$ denote the Brownian local time at level 0. Let (B_m) be our usual sequence of scaled random walks that converges to B almost surely uniformly on every compact interval. Let $\ell_m(k) := \#\{0 \le l < k \mid \widetilde{S}_m(l) = 0\}$ and $\mathcal{L}_m(t) := \frac{1}{2^m}\ell_m\left(\lfloor t2^{2m}\rfloor\right)$ the local time of the mth approximation at point 0 up to time t. We will prove that the sequence of the local time of the discrete approximations is almost surely uniformly converges to the local time of the Brownian motion. The main idea of the proof is the following. When the random walk, say the mth, hits 0 at time k the finer random walk, the $(m+1)$th, has $\frac{1}{2}(T_{m+1}(k+1) - T_{m+1}(k))$ 0 hits between $T_{m+1}(k)$ and $T_{m+1}(k+1)$ (for the definition T_{m+1} see (2.1)) which is geometrically distributed random variable with parameter $\frac{1}{2}$ so we can evaluate the time the Brownian motion spends at 0.

Theorem 6. *As* $n \to \infty$

$$\mathbf{P}\left(\sup_{[0,K]} |\mathcal{L}_n(t) - \mathcal{L}(t)| \ge C\, K^{1/4}(\log_* K)^2 n^2 2^{-n/2}\right) \le \lambda(C)\left(K^{1/2}2^n\right)^{2-C}, \tag{2.35}$$

with appropriate C dependent positive constant $\lambda(C)$ after an appropriately large n and fixed C.

Proper choice of the constant C and proper usage of the Borel-Cantelli lemma gives the following corollary:

Corollary 2.

$$\sup_{[0,K]} |\mathcal{L}_n(t) - \mathcal{L}(t)| < O(1)n^2 2^{-\frac{n}{2}} \qquad a.s. \qquad (n \to \infty)$$
$$\sup_{[0,K]} |\mathcal{L}_n(t) - \mathcal{L}(t)| < K^{\frac{1}{4}}(\log K)^2 \qquad a.s. \qquad (K \to \infty)$$

Before the proof of Theorem 6 we recall Lemma B in the following form.

Lemma B. *Let*

$$A_m = \left\{\sup_{0\le k2^{-2m}\le K} |T_{m+1}(k)2^{-2(m+1)} - k2^{-2m}| \ge \left(\frac{3}{2}CK\log_* K\right)^{\frac{1}{2}} m^{\frac{1}{2}} 2^{-m}\right\}.$$

where $\log_* x = \max\{1, \log x\}$. *Then, for any* $K > 0$, $C > 1$, *and for any* $m \ge m_0(C)$, *we have*

$$\mathbf{P}\{A_m\} \le 2(K2^{2m})^{1-C}.$$

We need the following LDP type result on the convergence of local time.

Lemma 6. *Denote by $\ell(k)$ the local time of the symmetric random walk at 0 until the kth step. For j: $\sqrt{k} \le j \le k$ if $k \to \infty$ then*

$$\mathbf{P}(\ell(2k) = j) \le \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{k}}\exp\left(-\frac{1}{2}\left(\frac{j}{\sqrt{k}}\right)^2\right)$$

and

$$\mathbf{P}\left(\frac{\ell(2k)}{\sqrt{2k}} \ge c(\log 2k)^{3/4}\right) < \frac{1}{\sqrt{2\pi}}\frac{1}{c(\log 2k)^{3/4}}e^{-\frac{1}{2}c^2(\log 2k)^{3/2}}.$$

Proof of Lemma 6. For proving these, we apply normal approximation for the distribution of ℓ:

$$\mathbf{P}\left(\ell(2k)=j\right)=2^{-2k+j}\binom{2k-j}{k}\sim\frac{1}{2}\frac{1}{\sqrt{\pi(2k-j)/2}}\exp\left(-\frac{\left(\frac{2k-(2k-j)}{2}\right)^2}{\frac{2k-j}{2}}\right)$$
$$=\frac{1}{2}\frac{1}{\sqrt{\pi k}}\cdot\left(1-\frac{j}{2k}\right)^{-1/2}\exp\left(-\frac{1}{2}\left(\frac{j}{\sqrt{k}}\right)^2\cdot\left(1-\frac{j}{2k}\right)^{-1}\right)\le\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{k}}\exp\left(-\frac{1}{2}\left(\frac{j}{\sqrt{k}}\right)^2\right)$$

if $\sqrt{k}\le j\le k$ and $k\to\infty$. Using this approximation and the asymptotic property of the normal distribution

$$1-\Phi(x)<\frac{1}{x\sqrt{2\pi}}e^{-x^2/2},$$

if x is large enough, one can readily write

$$\mathbf{P}\left(\frac{\ell(2k)}{\sqrt{2k}}\ge c(\log 2k)^{3/4}\right)=\sum_{c(\log 2k)^{3/4}\le\frac{j}{\sqrt{2k}}}\mathbf{P}\left(\ell(2k)=j\right)\approx\left(1-\Phi\left(c(\log 2k)^{3/4}\right)\right)$$
$$<\frac{1}{\sqrt{2\pi}}\frac{1}{c(\log 2k)^{3/4}}e^{-\frac{1}{2}c^2(\log 2k)^{3/2}}.$$

□

Proof of Theorem 6. We will prove that two consecutive approximations are close enough to each other. Namely

$$\mathbf{P}\left(\sup_{[0,K]}|\mathcal{L}_{m+1}(t)-\mathcal{L}_m(t)|\ge C\,K^{1/4}(\log_* K)^2m^22^{-m/2}\right)\le 3\left(K^{1/2}2^m\right)^{2-C}.\tag{2.36}$$

Therefore, $\{\mathcal{L}_m\}$ is a Cauchy sequence in the topology of the uniform convergence on compact sets:

$$\mathbf{P}\left(\sup_{[0,K]}|\mathcal{L}_{n+j}(t)-\mathcal{L}_n(t)|\ge C\,K^{1/4}(\log_* K)^2n^22^{-n/2}\ \text{ for some } j\ge 1\right)\le\lambda(C)\left(K^{1/2}2^n\right)^{2-C},\tag{2.37}$$

with appropriate C dependent positive constant $\lambda(C)$ after an appropriately large n and fixed C. Hence, (2.35) follows.

For proving (2.36) we first apply a triangle inequality

$$\sup_{[0,K]}|\mathcal{L}_{m+1}(t)-\mathcal{L}_m(t)|=\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}|\ell_{m+1}(4k)-2\ell_m(k)|$$
$$\le\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}|\ell_{m+1}(4k)-\ell_{m+1}(T_{m+1}(k))|+\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}|\ell_{m+1}(T_{m+1}(k))-2\ell_m(k)|.$$

We will estimate

$$p_1=\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}|\ell_{m+1}(4k)-\ell_{m+1}(T_{m+1}(k))|\ge C\,K^{1/4}(\log_* K)^2m^22^{-m/2}\right)$$
$$p_2=\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}|\ell_{m+1}(T_{m+1}(k))-2\ell_m(k)|\ge C\,K^{1/4}(\log_* K)^2m^22^{-m/2}\right)$$

separately.

Estimation of $\mathbf{p_1}$

Instead of p_1 we will estimate a strictly larger probability

$$\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\left|\ell_{m+1}(4k)-\ell_{m+1}(T_{m+1}(k))\right|\ge C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right).$$

Since $\{S(k+\min\{4k,T_{m+1}(k)\})-S(\min\{4k,T_{m+1}(k)\})\}_k$ has the same law as $\{S(k)\}_k$ one can write

$$\begin{aligned}&\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\left|\ell_{m+1}(4k)-\ell_{m+1}(T_{m+1}(k))\right|\ge C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right)\\&=\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\ell_{m+1}\left(\min\{4k,T_{m+1}(k)\},\ \max\{4k,T_{m+1}(k)\}\right)\ge C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right)\\&\qquad=\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\ell_{m+1}\left(0,|4k-T_{m+1}(k)|\right)\ge C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right)\end{aligned}$$

Using Lemma B one one has

$$\begin{aligned}&\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\ell_{m+1}\left(0,|4k-T_{m+1}(k)|\right)\ge C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right)\\&\le\mathbf{P}\left(A_m^c;\ \sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\ell_{m+1}\left(0,|4k-T_{m+1}(k)|\right)\ge C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right)+\mathbf{P}(A_m)\\&\le\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\ell_{m+1}\left(\left(\frac{2}{3}CK\log_*K\right)^{\frac12}m^{\frac12}2^m\right)\ge C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right)+2(K2^{2m})^{1-C}\\&\le\sum_{k=1}^{K2^{2m}}\mathbf{P}\left(\ell_{m+1}\left(\left(\frac{2}{3}CK\log_*K\right)^{\frac12}m^{\frac12}2^m\right)\ge C^{1/2}K^{1/4}\log_* Km2^{m/2}\right)+2(K2^{2m})^{1-C}.\qquad(2.38)\end{aligned}$$

We want to apply Lemma 6. For, proceed with the following transformation:

$$\begin{aligned}&\mathbf{P}\left(\ell_{m+1}\left(\left(\frac{2}{3}CK\log_*K\right)^{\frac12}m^{\frac12}2^m\right)\ge C^{1/2}K^{1/4}\log_* Km2^{m/2}\right)\\&=\mathbf{P}\left(\frac{\ell_{m+1}\left(\left(\frac{2}{3}CK\log_*K\right)^{\frac12}m^{\frac12}2^m\right)}{\left(\frac{2}{3}CK\log_*K\right)^{1/4}m^{1/4}2^{m/2}}\ge\frac{C^{1/2}K^{1/4}\log_* Km2^{m/2}}{\left(\frac{2}{3}CK\log_*K\right)^{1/4}m^{1/4}2^{m/2}}\right)\\&=\mathbf{P}\left(\frac{\ell_{m+1}\left(\left(\frac{2}{3}CK\log_*K\right)^{\frac12}m^{\frac12}2^m\right)}{\left(\frac{2}{3}CK\log_*K\right)^{1/4}m^{1/4}2^{m/2}}\ge\left(\frac{3}{2}\right)^{1/4}C^{1/4}(\log_* K)^{3/4}m^{3/4}\right).\end{aligned}$$

Routine calculation shows that

$$c2^{-1}\left(\log C+\log_* K+m\right)^{3/4}<c\left(\log\left(\left(\frac{2}{3}CK\log_* K\right)^{\frac{1}{2}}m^{\frac{1}{2}}2^m\right)\right)^{3/4}$$
$$\leq c\left(2\log_* C\log_* Km\right)^{3/4}=\left(\frac{3}{2}\right)^{1/4}C^{1/4}(\log_* K)^{3/4}m^{3/4},$$

where $c=\frac{3^{1/4}}{2}C^{1/4}(\log_* C)^{-3/4}$. Now, using the last two displayed lines then applying Lemma 6 we get

$$\mathbf{P}\left(\frac{\ell_{m+1}\left(\left(\frac{2}{3}CK\log_* K\right)^{\frac{1}{2}}m^{\frac{1}{2}}2^m\right)}{\left(\frac{2}{3}CK\log_* K\right)^{1/4}m^{1/4}2^{m/2}}\geq\left(\frac{3}{2}\right)^{1/4}C^{1/4}(\log_* K)^{3/4}m^{3/4}\right)$$
$$<\mathbf{P}\left(\frac{\ell_{m+1}\left(\left(\frac{2}{3}CK\log_* K\right)^{\frac{1}{2}}m^{\frac{1}{2}}2^m\right)}{\left(\frac{2}{3}CK\log_* K\right)^{1/4}m^{1/4}2^{m/2}}\geq c\left(\log\left(\left(\frac{2}{3}CK\log_* K\right)^{\frac{1}{2}}m^{\frac{1}{2}}2^m\right)\right)^{3/4}\right)$$
$$<e^{-\frac{1}{4}c^2(\log C+\log_* K+m)^{3/2}}<e^{-\frac{1}{8}\frac{C^{1/2}}{\log^{3/2}C}(\log C+\log_* K+m)^{3/2}}<e^{-\frac{1}{8}\frac{C^{1/2}}{\log^{3/2}C}\left(\log^{3/2}C+\log_*^{3/2}K+m^{3/2}\right)}. \tag{2.39}$$

Henceforth, one can continue the estimation of the first term in (2.38):

$$\sum_{k=1}^{K2^{2m}}\mathbf{P}\left(\ell_{m+1}\left(\left(\frac{2}{3}CK\log_* K\right)^{\frac{1}{2}}m^{\frac{1}{2}}2^m\right)\geq C^{1/2}K^{1/4}\log_* Km2^{m/2}\right)$$
$$<K2^{2m}e^{-\frac{1}{8}\frac{C^{1/2}}{\log^{3/2}C}\left(\log^{3/2}C+\log_*^{3/2}K+m^{3/2}\right)}<e^{\log K\log 2+2m\log 2-\frac{1}{8}\frac{C^{1/2}}{\log^{3/2}C}\left(\log^{3/2}C+\log_*^{3/2}K+m^{3/2}\right)}. \tag{2.40}$$

Hence, by (2.38) and (2.40), we get

$$p_1<\mathbf{P}\left(\sup_{1\leq k\leq K2^{2m}}\frac{1}{2^{m+1}}\left|\ell_{m+1}(4k)-\ell_{m+1}(T_{m+1}(k))\right|\geq C^{1/2}K^{1/4}\log_* Km2^{-m/2}\right)$$
$$\leq e^{\log K\log 2+2m\log 2-\frac{1}{8}\frac{C^{1/2}}{\log^{3/2}C}(\log C+\log_* K+m)^{3/2}}+2(K2^{2m})^{1-C}. \tag{2.41}$$

Estimation of the probability $\mathbf{p_2}$

Taking conditional expectation with respect to $\ell_m(K2^{2m})$ one has the following sum

$$\mathbf{P}\left(\sup_{1\leq k\leq K2^{2m}}\frac{1}{2^{m+1}}\left|\ell_{m+1}(T_{m+1}(k))-2\ell_m(k)\right|\geq C\,K^{1/4}(\log_* K)^2m^22^{-m/2}\right)$$
$$=\sum_{n=1}^{K2^{2m}}\mathbf{P}\left[\sup_{1\leq k\leq K2^{2m}}\frac{1}{2^{m+1}}\left|\ell_{m+1}(T_{m+1}(k))-2\ell_m(k)\right|\geq C\,K^{1/4}(\log_* K)^2m^22^{-m/2}\,\middle|\,\ell_m(K2^{2m})=n\right]$$
$$\cdot\,\mathbf{P}\left(\ell_m(K2^{2m})=n\right). \tag{2.42}$$

Supposing $\ell_m(K2^{2m}) = n$ we have the equality:

$$\sup_{1\le k\le K2^{2m}} \frac{1}{2^{m+1}} \left|\ell_{m+1}(T_{m+1}(k)) - 2\ell_m(k)\right| = \sup_{1\le k\le n} \frac{1}{2^{m+1}} \left|\sum_{i=1}^{k} \gamma_i - 2k\right|,$$

where $\{\gamma_i\}_i$ is an i.i.d sequence of *geometrical* variables with parameter $1/2$. Taking into consideration this remark one can write the sum (2.42) in he following form:

$$\sum_{n=1}^{K2^{2m}} \mathbf{P}\left(\frac{1}{2^{m+1}} \sup_{1\le k\le n}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right) \mathbf{P}\left(\ell_m(K2^{2m}) = n\right).$$

Divide this sum into two parts at level $N = C^{1/2}K^{1/2}(\log_* K)^{3/4}m^{3/4}2^m$ one finds:

$$\begin{aligned}
&\sum_{n=1}^{K2^{2m}} \mathbf{P}\left(\frac{1}{2^{m+1}} \sup_{1\le k\le n}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right) \mathbf{P}\left(\ell_m(K2^{2m}) = n\right)\\
&= \sum_{n=1}^{N} \mathbf{P}\left(\frac{1}{2^{m+1}} \sup_{1\le k\le n}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right) \mathbf{P}\left(\ell_m(K2^{2m}) = n\right) \qquad (2.43)\\
&+ \sum_{n=N}^{K2^{2m}} \mathbf{P}\left(\frac{1}{2^{m+1}} \sup_{1\le k\le n}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right) \mathbf{P}\left(\ell_m(K2^{2m}) = n\right). \qquad (2.44)
\end{aligned}$$

Overestimating the sum (2.43) we get

$$\begin{aligned}
&\sum_{n=1}^{N} \mathbf{P}\left(\frac{1}{2^{m+1}} \sup_{1\le k\le n}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right) \mathbf{P}\left(\ell_m(K2^{2m}) = n\right)\\
&< N\cdot \mathbf{P}\left(\sup_{1\le k\le N}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{m/2}\right)\\
&< N\cdot \mathbf{P}\left(\sup_{1\le k\le N}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge 2^{1/2}(2CN\log N)^{1/2}\right)\\
&< N^{1+1-C} < \left(K^{1/2}(\log_* K)^{3/4}m^{3/4}2^m\right)^{2-C}, \qquad (2.45)
\end{aligned}$$

if C is large enough. We have basically applied Lemma A here.

On one hand, $\mathbf{Var}(\gamma_1 - 2) = 2$. This is the origin of the $2^{1/2}$ multiplier at the right in the probability in the third row. (In Lemma A we have to ensure that the variance of X_i is 1.) On the other hand,

$$\begin{aligned}
&(2CN\log N)^{1/2}\\
&= \left(2CC^{1/2}K^{1/2}(\log_* K)^{3/4}m^{3/4}2^m\left(\frac{1}{2}\log C + \frac{1}{2}\log K + \frac{3}{4}\log\log_* K + \frac{3}{4}\log m + m\log 2\right)\right)^{1/2}\\
&< 2C^{3/4}(\log C)^{1/2}K^{1/4}(\log_* K)^2 m^2 2^{m/2} < CK^{1/4}(\log_* K)^2 m^2 2^{m/2}, \qquad (2.46)
\end{aligned}$$

if C is large enough. This yields the second estimate. The third is a pure application of Lemma A. The last can be obtained by omitting the term $C^{1/2}$ from N.

Overestimating (2.44) one can write

$$\sum_{n=N}^{K2^{2m}} \mathbf{P}\left(\frac{1}{2^{m+1}} \sup_{1\le k\le n}\left|\sum_{i=1}^{k}\gamma_i - 2k\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right)\mathbf{P}\left(\ell_m(K2^{2m}) = n\right)$$

$$< \sum_{n=N}^{K2^{2m}} \mathbf{P}\left(\ell_m(K2^{2m}) = n\right) = \mathbf{P}\left(\ell_m(K2^{2m}) \ge N\right) = \mathbf{P}\left(\frac{\ell_m(K2^{2m})}{K^{1/2}2^m} \ge C^{1/2}\left(\log_* Km\right)^{3/4}\right)$$

$$< e^{-\frac{1}{2}C(\log_* Km)^{3/2}}, \quad (2.47)$$

by a proper application of Lemma 6.
Summarizing (2.45) and (2.47) one has

$$\mathbf{P}\left(\sup_{1\le k\le K2^{2m}} \frac{1}{2^{m+1}}\left|\ell_{m+1}(T_{m+1}(k)) - 2\ell_m(k)\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right)$$

$$\le 2(K2^{2m})^{1-C} + \left(K^{1/2}(\log_* K)^{3/4}m^{3/4}2^m\right)^{2-C} + e^{-\frac{1}{2}C(\log_* Km)^{3/2}}. \quad (2.48)$$

The lines (2.41) and (2.48) together yield (2.36):

$$\mathbf{P}\left(\sup_{[0,K]}\left|\mathcal{L}_{m+1}(t) - \mathcal{L}_m(t)\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right)$$

$$\le e^{\log K \log 2 + 2m\log 2 - \frac{1}{8}\frac{C^{1/2}}{\log^{3/2} C}\left(\log^{3/2} C + \log_*^{3/2} K + m^{3/2}\right)}$$

$$+ \left(K^{1/2}(\log_* K)^{3/4}m^{3/4}2^m\right)^{2-C} < 3\left(K^{1/2}2^m\right)^{2-C}$$

if C is large enough. Now, for proving (2.37) we use the following estimation

$$\sup_{[0,K]}\left|\mathcal{L}_{n+j}(t) - \mathcal{L}_n(t)\right| = \sup_{[0,K]}\left|\sum_{m=n}^{n+j}\mathcal{L}_{m+1}(t) - \mathcal{L}_m(t)\right|$$

$$\le \sum_{m=n}^{n+j}\sup_{[0,K]}\left|\mathcal{L}_{m+1}(t) - \mathcal{L}_m(t)\right| < \sum_{m=n}^{\infty} C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2} < \mathrm{c}\,C\,K^{1/4}(\log_* K)^2 n^2 2^{-n/2} \quad (2.49)$$

with proper constant, say $\mathrm{c} = 20$ for any $j > 1$ after an appropriately large n. We omit the constant c from the latter expressions because it can be built in e.g. the estimate (2.46) if C is large enough. Therefore, we can conclude that

$$\mathbf{P}\left(\sup_{[0,K]}\left|\mathcal{L}_{n+j}(t) - \mathcal{L}_n(t)\right| \ge C\,K^{1/4}(\log_* K)^2 n^2 2^{-n/2} \text{ for some } j \ge 1\right)$$

$$\le \sum_{m=n}^{\infty}\mathbf{P}\left(\sup_{[0,K]}\left|\mathcal{L}_{m+1}(t) - \mathcal{L}_m(t)\right| \ge C\,K^{1/4}(\log_* K)^2 m^2 2^{-m/2}\right)$$

$$\le \sum_{m=n}^{\infty} 3\left(K^{1/2}2^m\right)^{2-C} \le \lambda(C)\left(K^{1/2}2^n\right)^{2-C},$$

with appropriate C dependent positive constant $\lambda(C)$. This last line proves (2.37) and (2.35). □

The following corollary is necessary for proving an stochastic integral approximation result Theorem 9 in Section 2.5.

Before this global convergence theorem we introduce some notation. Let $\mathcal{L}^a(t)$ and $\mathcal{L}_n^a(t)$ respectively be the local time of the standard Brownian motion and its nth approximation respectively at level a until time t. For arbitrary real number a let us define $\mathcal{L}_n^a(t) = \mathcal{L}_n^{\lceil a \rceil_n}(t)$ where $\lceil a \rceil_n = \frac{\lceil a2^n \rceil}{2^n}$. More, $\mathcal{L}_n^{\lceil a \rceil_n}(t) = \frac{1}{2^m}\ell_m^{\lceil a2^n \rceil}(\lfloor t2^{2m} \rfloor) = \frac{1}{2^m}\#\{0 \le l < \lfloor t2^{2m} \rfloor \,|\, \widetilde{S}_m(l) = \lceil a2^n \rceil\}$

Corollary 3.

$$\sup_{a \in \mathbb{R}} \sup_{[0,K]} |\mathcal{L}_n^a(t) - \mathcal{L}^a(t)| < O(1)2^{-\left(\frac{1}{2}-\varepsilon\right)n} \qquad a.s. \qquad (n \to \infty) \tag{2.50}$$

for an arbitrary small positive ε.

The rate of the convergence is worse than in Corollary 2 and we does not say anything about the case when m is fixed and K tends to infinity. The reason is that we use the following statement on the uniformly Hölder continuity of the local time for proving the corollary.

For all fixed K there exist a positive random variable D_K such that

$$\sup_{t \in [0,K]} |\mathcal{L}^x(t) - \mathcal{L}^y(t)| \le D_K |x-y|^\alpha \tag{2.51}$$

for every $\alpha < 1/2$ (see [23, (1.32) Exercise, Chapter VI]).

Proof of Corollary 3. First, we prove

$$\mathbf{P}\left(\sup_{\substack{a = j2^{-2n} \\ j \in \mathbb{Z}}} \sup_{[0,K]} |\mathcal{L}_{n+i}^a(t) - \mathcal{L}_n^a(t)| \ge C\,K^{1/4}(\log_* K)^2 n^2 2^{-n/2} \text{ for some } i \ge 1 \right) \le \lambda(C)\left(K^{1/2}2^n\right)^{4-C}, \tag{2.52}$$

for all fixed j with appropriate C dependent positive constant. This yields

$$\sup_{\substack{a = j2^{-2n} \\ j \in \mathbb{Z}}} \sup_{[0,K]} |\mathcal{L}_n^a(t) - \mathcal{L}^a(t)| < C\,K^{1/4}(\log_* K)^2 n^2 2^{-n/2} \qquad \text{a.s.} \qquad (n \to \infty) \tag{2.53}$$

To prove (2.52) we first apply the following transformation:

$$\sup_{\substack{a = j2^{-2m} \\ j \in \mathbb{Z}}} \sup_{[0,K]} |\mathcal{L}_{m+1}^a(t) - \mathcal{L}_m^a(t)| = \sup_{j \in \mathbb{Z}} \sup_{1 \le k \le K2^{2m}} \frac{1}{2^{m+1}} \left|\ell_{m+1}^{2j}(4k) - 2\ell_m^j(k)\right|.$$

Dividing $\mathbb{Z}$ into two parts one gets

$$\begin{aligned} &\sup_{j \in \mathbb{Z}} \sup_{1 \le k \le K2^{2m}} \frac{1}{2^{m+1}} \left|\ell_{m+1}^{2j}(4k) - 2\ell_m^j(k)\right| \\ &\le \sup_{|j| \le K2^{2m}} \sup_{1 \le k \le K2^{2m}} \frac{1}{2^{m+1}} \left|\ell_{m+1}^{2j}(4k) - 2\ell_m^j(k)\right| + \sup_{|j| > K2^{2m}} \sup_{1 \le k \le K2^{2m}} \frac{1}{2^{m+1}} \left|\ell_{m+1}^{2j}(4k) - 2\ell_m^j(k)\right|. \end{aligned}$$

Here, the second term is 0 because under $K2^{2m}$ steps the random walk can not reach the level further than $K2^{2m}$ from the origin. Therefore, we only have to deal with first part.

$$\begin{aligned}
&\mathbf{P}\left(\sup_{j\in\mathbb{Z}}\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\left|\ell^{2j}_{m+1}(4k)-2\ell^{j}_{m}(k)\right|\ge C\,K^{1/4}(\log_* K)^2m^2 2^{-m/2}\right)\\
&\quad=\mathbf{P}\left(\sup_{|j|\le K2^{2m}}\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\left|\ell^{2j}_{m+1}(4k)-2\ell^{j}_{m}(k)\right|\ge C\,K^{1/4}(\log_* K)^2m^2 2^{-m/2}\right)\\
&\quad\le\sum_{|j|\le K2^{2m}}\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\left|\ell^{2j}_{m+1}(4k)-2\ell^{j}_{m}(k)\right|\ge C\,K^{1/4}(\log_* K)^2m^2 2^{-m/2}\right)\\
&\quad=\sum_{|j|\le K2^{2m}}\mathbf{P}\left(\sup_{\tau_j\le k\le K2^{2m}}\frac{1}{2^{m+1}}\left|\ell^{2j}_{m+1}(4k)-2\ell^{j}_{m}(k)\right|\ge C\,K^{1/4}(\log_* K)^2m^2 2^{-m/2}\right)\\
&\quad\le\sum_{|j|\le K2^{2m}}\mathbf{P}\left(\sup_{1\le k\le K2^{2m}}\frac{1}{2^{m+1}}\left|\ell_{m+1}(4k)-2\ell_{m}(k)\right|\ge C\,K^{1/4}(\log_* K)^2m^2 2^{-m/2}\right)\\
&\qquad\qquad< 2K2^{2m}3\left(K^{1/2}2^m\right)^{2-C}=6\left(K^{1/2}2^m\right)^{4-C},
\end{aligned}$$

where τ_j denotes the first hitting time of the level j. We used the strong Markov property of the standard random walk and Theorem (6).

At this point we apply a (2.49) like argument:

$$\begin{aligned}
&\sup_{\substack{a=j2^{-2n}\\ j\in\mathbb{Z}}}\sup_{[0,K]}\left|\mathcal{L}^a_{n+j}(t)-\mathcal{L}^a_n(t)\right| = \sup_{\substack{a=j2^{-2n}\\ j\in\mathbb{Z}}}\sup_{[0,K]}\left|\sum_{m=n}^{n+j}\mathcal{L}^a_{m+1}(t)-\mathcal{L}^a_m(t)\right|\\
&\quad\le\sum_{m=n}^{n+j}\sup_{\substack{a=j2^{-2n}\\ j\in\mathbb{Z}}}\sup_{[0,K]}\left|\mathcal{L}^a_{m+1}(t)-\mathcal{L}^a_m(t)\right|\le\sum_{m=n}^{n+j}\sup_{\substack{a=j2^{-2m}\\ j\in\mathbb{Z}}}\sup_{[0,K]}\left|\mathcal{L}^a_{m+1}(t)-\mathcal{L}^a_m(t)\right|\\
&\qquad\qquad<\sum_{m=n}^{\infty}C\,K^{1/4}(\log_* K)^2m^2 2^{-m/2}<\mathrm{c}\,C\,K^{1/4}(\log_* K)^2n^2 2^{-n/2}
\end{aligned}$$

with proper constant $\mathrm{c}>0$, which can be omitted if C is large enough, for any $j>1$ after an appropriately large n. So we can conclude that

$$\begin{aligned}
&\mathbf{P}\left(\sup_{\substack{a=j2^{-2m}\\ j\in\mathbb{Z}}}\sup_{[0,K]}\left|\mathcal{L}^a_{n+j}(t)-\mathcal{L}^a_n(t)\right|\ge C\,K^{1/4}(\log_* K)^2n^2 2^{-n/2}\ \text{ for some } j\ge 1\right)\\
&\quad\le\sum_{m=n}^{\infty}\mathbf{P}\left(\sup_{\substack{a=j2^{-2m}\\ j\in\mathbb{Z}}}\sup_{[0,K]}\left|\mathcal{L}^a_{m+1}(t)-\mathcal{L}^a_m(t)\right|\ge C\,K^{1/4}(\log_* K)^2m^2 2^{-m/2}\right)\\
&\qquad\qquad\le\sum_{m=n}^{\infty}6\left(K^{1/2}2^m\right)^{4-C}\le\lambda(C)\left(K^{1/2}2^n\right)^{4-C},
\end{aligned}$$

with appropriate C dependent positive constant which proves (2.52).

Now, we are ready to prove (2.50). By using (2.51) and (2.53) we get

$$\sup_{a\in\mathbb{R}}\sup_{[0,K]}|\mathcal{L}_n^a(t)-\mathcal{L}^a(t)| = \sup_{a\in\mathbb{R}}\sup_{[0,K]}\left|\mathcal{L}_n^{\lceil a\rceil_n}(t)-\mathcal{L}^a(t)\right|$$
$$\leq \sup_{\substack{a=j2^{-2n}\\ j\in\mathbb{Z}}}\sup_{[0,K]}\left|\mathcal{L}_n^a(t)-\mathcal{L}^{\lceil a\rceil_n}(t)\right| + \sup_{a\in\mathbb{R}}\sup_{[0,K]}\left|\mathcal{L}^{\lceil a\rceil_n}(t)-\mathcal{L}^a(t)\right|$$
$$\leq \mathrm{c}\,C\,K^{1/4}(\log_* K)^2 n^2 2^{-n/2} + D_K 2^{-\left(\frac{1}{2}-\varepsilon\right)n} \leq O(1)2^{-\left(\frac{1}{2}-\varepsilon\right)n} \quad a.s.$$

as n tends to infinity.

□

2.4.2 Excursion, bridge and meander – Marchal's construction

We present Phillipe Marchal's algorithm for approximating the Brownian excursion, bridge and meander. His algorithms allow to generate directly the excursion, the bridge and the meander. Moreover, many self-similarity properties are easily seen from this construction, and one can also recover various distributions of BM from the urn schemes that are embedded in the construction. In his paper he also gave a local time approximating algorithm but this one is not so natural as in the first three cases. In spite of the originality of that construction we will not present it here.

Theorem F. *There exists a family $(S^n, n\geq 1)$ of random walks on $\mathbb{Z}$ starting at 0 where for every n, S^n respectively:*

(1) has length $2n$ and is conditioned to return to 0 at time $2n$,

(2) has length $2n$ and conditioned to return to 0 at time $2n$ and to stay positive from time 1 to $2n-1$,

(3) has length n and conditioned to stay positive from time 1 to $n-1$,

and such that almost surely, for every $t\in[0,1]$, $S^n_{\lfloor 2nt\rfloor}/\sqrt{n}$ (or $S^n_{\lfloor nt\rfloor}/\sqrt{n}$ in the third case) converges to $\widetilde{B}_t$ where $(\widetilde{B}_t, 0\leq t\leq 1)$ is respectively:

(1) a Brownian bridge,

(2) a Brownian excursion,

(3) a Brownian meander.

Before the construction we have to introduce some notation and definition.

If the meander of a path $\mathcal{P}$ is positive, a point t in the meander is visible from the right if

$$\mathcal{P}(t) = \min\{\mathcal{P}(n)|\, n\geq t\,\}.$$

A positive step followed by a negative step is called positive hat. A negative step is defined likewise.

Let us describe a procedure to extend a path $\mathcal{P}$. Suppose we are given a time T such that $\mathcal{P}(T)>0$. Then *lifting the Dyck path before time T* means the following. Let

$$T' = 1+\sup\{n\leq T|\,\mathcal{P}(n)<\mathcal{P}(T)\}.$$

Then form the new path $\mathcal{P}'$ by inserting a positive step at time T' and then a negative step at time $T+1$. We remark the if $\mathcal{P}(T-1)=\mathcal{P}(T)-1$, then $T=T'$ and lifting the Dyck path before time T amounts to inserting a hat at time T.

Now, we describe that how to generate S^{n+1} from S^n:

I. Choose a random time t uniformly on $\{0,\dots,2n-1\}$.

II. If $S^n(t)=0$, insert a positive or negative hat at time t with respective probabilities 1/2–1/2.

III. If t is in a positive excursion,

1 with probability 1/2 insert a positive hat at time t
2 with probability 1/2 lift the Dyck path before time t.
IV. If t is in the meander of S^n and if this meander is positive,
1 If t is visible from the right, or if t is visible from the right and $S^n(t)$ is even, proceed as in III.
2 If t is visible from the right and $S^n(t)$ is odd,
a with probability 1/2 insert at time t a positive hat,
b with probability 1/2 insert at time t two positive steps.

Now, let us present the algorithms. In case (1) of F, begin with S^0 the empty path. To obtain S^{n+1} from S^n, choose a random time t uniformly in $\{0,\dots,2n\}$ and apply procedure **II** or **III**.

In case (2), begin with S^1 a positive hat. To obtain S^{n+1} from S^n, choose a random time t uniformly in $\{1,\dots,2n\}$ and apply procedure **III**.

In case (3), begin with S^1 a positive step. To obtain S^{2n+1} from S^{2n-1}, choose a random time t uniformly in $\{1,\dots,2n\}$ and apply procedure **IV**. To obtain S^{2n} from S^{2n-1}, just add a last random step.

The idea of the proof of these statements is a bijection made between Dyck paths and random binary trees. The convergence rate in all case is $O(1/n^{1/4})$

2.5 Pathwise stochastic integration

Our objective in this section is to define a sequence of stochastic sums converging to $\int_0^t Y_s \,\mathrm{d}M(s)$ with probability 1 on every compact interval, where M is a continuous local martingale and Y is an integrable stochastic process with respect to M. It will turn out that one can carry out this procedure for two types of processes. The first is of the form $f'_-(M)$ where f'_- denotes the left-hand side derivative of the difference of two convex functions. The other family of processes is that of with right continuous paths and left limit.

The main idea is that one can take first a dyadic partition on the "spatial" axis that gives the random stopping times of the Skorohod embedding on the time axis. At this point the procedure somewhat reminds a possible definition of Lebesgue integral. So we get an approach of stochastic integrals which is basically different from the usual definition of stochastic integration.

Similar approach can be found in Karandikar's papers. Because of this similarity we think that showing some of his results could make our presentation more complete.

The theorems in Subsection 2.5.1 are taken from the above mentioned paper [13] from 1996. Here, we could find statements on stochastic integration w.r.t Brownian motion and semimartingales and even pathwise construction of solution of SDE.

The main difference of the two approaches is for which process the discretisation is applied. Karandikar applied the discretisation to the integrand while we apply it to the integrator martingale.

2.5.1 Karandikar's integral approximation

Throughout this subsection, we fix a complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and filtration $(\mathcal{F}_t)$ satisfying the usual conditions.

His first theorem is about pathwise stochastic integration w.r.t. Brownian motion is

Theorem G. *Let (W_t) be Brownian motion adapted to the filtration $(\mathcal{F}_t)$ such that $W_t - W_s$ is independent of $\mathcal{F}_s$ for all $0 \leq s \leq t < \infty$. Let f be an r.c.l.l. adapted process and for $n \geq 1$, let $\{\tau_i^n : i \geq 0\}$ be defined by $\tau_0^n = 0$ and for $i \geq i$*

$$\tau_{i+1}^n = \inf\{t \geq \tau_i^n : |f_t(\cdot) - f_{\tau_i^n}(\cdot)| \geq 2^{-n}\}.$$

Let (Y_t^n) be defined as follows. For $\tau_k^n \leq t < \tau_{k+1}^n, k \geq 0$,

$$Y_t^n = \sum_{i=0}^{k-1} f_{\tau_k^n}(W_{\tau_{i+1}^n} - W_{\tau_i^n}) + f_{\tau_k^n}(W_t - W_{\tau_k^n}).$$

Then, for all $T < \infty$ we have

$$\sup_{0 \leq t \leq T} \left| Y_t^n - \int_0^t f \, \mathrm{d}W \right| \to 0 \qquad a.s. \qquad (n \to \infty).$$

We present his proof of this theorem because we will use some notations and key steps of that in the proof of Theorem 7.

Proof. Note that $Y_t^n = \int_0^t f^n \, \mathrm{d}W$, where

$$f_t^n = f_{\tau_k^n} \text{ for } \tau_k^n \leq t < \tau_{k+1}^n$$

and hence by the choice of $\{\tau_i^n\}$ we have

$$|f_t^n - f_t| \leq 2^{-n}\,.$$

Using the standard estimate

$$\mathbf{E} \sup_{0 \leq t \leq T} \left| \int_0^t g \, \mathrm{d}W \right|^2 \leq 4\mathbf{E} \left| \int_0^T g^2 \, \mathrm{d}t \right| \tag{2.54}$$

one gets

$$\mathbf{E} \sup_{0 \leq t \leq T} \left| Y_t^n - \int_0^t f \, \mathrm{d}W \right|^2 \leq 4T2^{-2n}. \tag{2.55}$$

Let

$$U_n = \sup_{0 \leq t \leq T} \left| Y_t^n - \int_0^t f \, \mathrm{d}W \right|.$$

Then (2.55) implies that $\mathbf{E}U_n \leq 2\sqrt{T}2^{-n}$ and hence it follows that

$$\mathbf{E} \sum_n U_n = \sum_n \mathbf{E}U_n \leq \sum_n 2\sqrt{T}2^{-n} < \infty.$$

As a consequence, one gets

$$\sum_n U_n < \infty \text{ a.s.}$$

which gives the required conclusion. □

Theorem H. *Let X be a semimartingale and let Z be an r.c.l.l. adapted process. For $n \geq 1$, let $\{\tau_i^n : i \geq 0\}$ be defined by $\tau_o^n = 0$ and for $i \geq i$*

$$\tau_{i+1}^n = \inf\{t \geq \tau_i^n : |Z_t(\cdot) - Z_{\tau_i^n}(\cdot)| \geq 2^{-n}\}.$$

Let (Y_t^n) be defined as follows. For $\tau_k^n \geq t < \tau_{k+1}^n, k \geq 0$,

$$Y_t^n = \sum_{i=0}^{k-1} Z_{\tau_k^n}(X_{\tau_{i+1}^n} - X_{\tau_i^n}) + Z_{\tau_k^n}(X_t - X_{\tau_k^n}).$$

Then, for all $T < \infty$ we have

$$\sup_{0 \leq t \leq T} \left| Y_t^n - \int_0^t Z \, \mathrm{d}X \right| \to 0 \qquad a.s. \qquad (n \to \infty).$$

Proof. Note that $Y_t^n = \int_0^t Z_-^n \, \mathrm{d}X$, where

$$Z_t^n = f Z_{\tau_k^n} \text{ for } \tau_k^n \leq t < \tau_{k+1}^n$$

for $k \geq 1$ and $Z_0^n = Z_0$. Hence by the choice of $\{\tau_i^n\}$ we have

$$\sup_t |Z_t^n - Z_{t-}| \leq 2^{-n}.$$

The core of the proof is the inequality (2.56). For this we have to introduce some notations. Let $X = M + A$ be a decomposition of a semimartingale X where M is a locally square integrable martingale and A is a process whose paths are of bounded variation an bounded intervals. Define stopping times σ_k increasing to ∞ such that $C_k = \mathbf{E}\langle M, M\rangle_{\sigma_k} < \infty$. We have the following statement:

For a predictable process f and stopping time σ we have

$$\mathbf{E} \sup_{0 \leq t \leq \sigma} \left| \int_0^t f \, \mathrm{d}M \right|^2 \leq 4\mathbf{E} \left| \int_0^\sigma f^2 \, \mathrm{d}\langle M, M\rangle \right|. \tag{2.56}$$

Plugging in our variables we obtain

$$\mathbf{E} \sup_{0 \leq t \leq \sigma_k} \left| \int_0^t Z^n \, \mathrm{d}M - \int_0^t Z_- \, \mathrm{d}M \right|^2 \leq 2^{-2n} C_k$$

and, as in Theorem G we can conclude that

$$\sup_{0 \leq t \leq \sigma_k} \left| \int_0^t Z^n \, \mathrm{d}M - \int_0^t Z_- \, \mathrm{d}M \right| \to 0 \text{ a.s.}$$

for all $k \geq 1$. Since $\sigma_k \to \infty$, we get

$$\sup_{0 \leq t \leq T} \left| \int_0^t Z^n \, \mathrm{d}M - \int_0^t Z_- \, \mathrm{d}M \right| \to 0 \text{ a.s.} \tag{2.57}$$

for all $T < \infty$. As for the $\mathrm{d}A$ integral, uniform convergence of Z_n to Z_- yields

$$\sup_{0 \leq t \leq \sigma_k} \left| \int_0^t Z^n \, \mathrm{d}A - \int_0^t Z_- \, \mathrm{d}A \right| \to 0. \tag{2.58}$$

Together, (2.5.1) and (2.58) yield the result. □

Finally we present Karandikar's result on pathwise approximation of the solution of a SDE. The SDE considered is

$$Z_t = H_t + \int_0^t a(Z)s_- \, \mathrm{d}X_s,$$

where X is an $\mathbb{R}^d$ valued semimartingale, H is a given adapted r.c.l.l. $\mathbb{R}^d$ valued process and

$$a : \mathbb{D}([0,\infty), \mathbb{R}^d) \to \mathbb{D}([0,\infty), L(d,d)),$$

where $L(d,d)$ is the space of $d \times d$ matrices.

We assume that the functional a satisfies the following Lipschitz condition. For each $T < \infty$ there exists a constant C_T such that

$$\|a(\rho_1)(t) - a(\rho_2)(t)\| \leq C_T \sup_{0 \leq s \leq t} \|\rho_1(s) - \rho_2(s)\|$$

for all $\rho_1, \rho_2 \in \mathbb{D}([0,\infty), \mathbb{R}^d)$ and for all $0 \leq t \leq T$. Here $\| \cdot \|$ denotes the Euclidean norm on $\mathbb{R}^d$ and on $L(d,d)$.

For arbitrary fixed $\rho, \eta \in \mathbb{D}([0,\infty), \mathbb{R}^d)$ and $n \geq 1$ we will define a function $\mathcal{S}_n(\eta, \rho) \in \mathbb{D}$. For, we first define some sequence of some variables for these fixed functions.

Let $\{u_i : i \geq 1\}$ and $\xi^i \in \mathbb{D}([0,\infty), \mathbb{R}^d)$ be defined inductively by

$$u_0 = 0 \text{ and } \xi^0_t \equiv \eta_0$$

and having defined u_j, ξ^j for $j \leq i$, let

$$\begin{aligned} u_{i+1} = \inf\{t > u_i : \|\eta(t) - \eta(u_i) + a(\xi^i)(u_i)(\rho(t) - \rho(u_i))\| &\geq 2^{-n} \\ \text{or } \|a(\xi^i)(t) - a(\xi^i)(u_i)\| &\geq 2^{-n}\} \end{aligned}$$

and

$$\xi^{i+1}(t) = \begin{cases} \xi^i(t) & \text{for } t < u_{i+1} \\ \xi^i(u_i) + \eta(u_{i+1}) - \eta(u_i) + a(\xi^i)(u_i)(\rho(u_{i+1}) - \rho(u_i)) & \text{for } t \geq u_{i+1} \end{cases}.$$

By definition ξ^i is a step function with jumps at $u_1, u_2, \ldots, u_{i+1}$.

Now, we are able to define the function $\mathcal{S}_n(\eta, \rho)$. Let $\mathcal{S}_n(\eta, \rho)(0) = \eta(0)$ and for $u_i < t \leq u_{i+1}$ let

$$\mathcal{S}_n(\eta, \rho)(t) = \xi^i(u_i) + \eta_t - \eta_{u_i} + a(\xi^i)(\rho(t) - \rho(u_i)).$$

So have defined a function on $\mathbb{D}([0,\infty), \mathbb{R}^d) \times \mathbb{D}([0,\infty), \mathbb{R}^d)$. Taking limit one gets the function

$$\mathcal{S}(\eta, \rho) = \lim_{n \to \infty} \mathcal{S}_n(\eta, \rho)$$

whenever the limit exists in the topology of uniform convergence of compact subsets.

Theorem I. *Let (X_t) be a semimartingale and let (H_t) be an r.c.l.l. adapted process. Then*

$$Z_t(\omega) = \mathcal{S}(H.(\omega), X.(\omega))(t)$$

is the unique solution to the equation

$$Z_t = H_t + \int_0^t a(Z)_{s-} \, \mathrm{d}X_s,$$

2.5.2 Discretization applied to the integrator

Theorem 7. *Let $(W_t, (\mathcal{F}_t))$ be a Brownian motion with the standard filtration. Let f be an r.c.l.l. adapted process and for $n \geq 1$ let $\{s_n(i) : i \geq 0\}$ be the stopping times defined in (2.6).*

Let (I_t^n) be defined as follows. For $s_n(k) \leq t < s_n(k+1), k \geq 0$,

$$I_t^n = \sum_{i=0}^{k-1} f_{s_n(i)}(W_{s_n(i+1)} - W_{s_n(i)}) + f_{s_n(k)}(W_t - W_{s_n(k)}).$$

Then, for all $T < \infty$ we have

$$\sup_{0 \leq t \leq T} \left| I_t^n - \int_0^t f \, \mathrm{d}W \right| \to 0 \qquad a.s. \qquad (n \to \infty). \tag{2.59}$$

Proof. Essentially, the proof is based on the fact that the sequence $\{s_m(k)\}_k$ is getting dense on every compact interval as m goes to infinity.

Fix $T, M > 0$ and let $\vartheta^M = \inf\{t \,|\, 0 \leq t \leq T, \ |f_t| > M\}$. Since f is r.c.l.l. process, $\max_{[0,T]} |f_t|$ is almost surely finite so $\mathbf{P}(\vartheta^M = T) \to 1$ as $M \to \infty$. Define

$$f_t^M = f_t \, \mathbf{I}\{t < \vartheta^M\}$$

$$f_t^{n,M} = f_{s_n(k)}^M \text{ for } s_n(k) \leq t < s_n(k+1) \tag{2.60}$$

and

$$I_t^{n,M} := \int_0^t f^{n,M} \, \mathrm{d}W. \tag{2.61}$$

We will prove that

$$\sup_{0 \leq t \leq T} \left| I_t^{n,M} - \int_0^t f^M \, \mathrm{d}W \right| \to 0 \ a.s. \tag{2.62}$$

or equivalently

$$\sup_{0 \leq t \leq \vartheta^M} \left| I_t^n - \int_0^t f \, \mathrm{d}W \right| \to 0 \ a.s.$$

Hence, the property $\mathbf{P}(\vartheta^M = T) \to 1$ as $M \to \infty$ yields (2.59). Indeed, the following events are a.s. equal to each other because of the property $\mathbf{P}(\vartheta^M = T) \to 1$ as $M \to \infty$:

$$\left\{ \sup_{0 \leq t \leq T} \left| I_t^n - \int_0^t f \, \mathrm{d}W \right| \underset{n\to\infty}{\longrightarrow} 0 \ a.s. \right\} = \left\{ \text{for all } m \in \mathbb{N} : \sup_{0 \leq t \leq \vartheta^m} \left| I_t^n - \int_0^t f \, \mathrm{d}W \right| \underset{n\to\infty}{\longrightarrow} 0 \ a.s. \right\}$$
$$= \bigcap_{m \in \mathbb{N}} \left\{ \sup_{0 \leq t \leq \vartheta^m} \left| I_t^n - \int_0^t f \, \mathrm{d}W \right| \underset{n\to\infty}{\longrightarrow} 0 \ a.s. \right\}.$$

Since each terms in the last section have probability 1, we get (2.59).

For proving (2.62), we prove a somewhat more general statement. Namely, we show (2.59) for arbitrary f r.c.l.l. adapted process such that $\max_{[0,T]} |f_t| \leq M$ a.s. In the sequel, we suppose f possesses this property. For simplicity we use the notations f_t^{n*} and I_t^n with the following meaning:

$$f_t^{n*} = f_{s_n(k)} \text{ for } s_n(k) \leq t < s_n(k+1) \text{ and } I_t^n := \int_0^t f^{n*} \, \mathrm{d}W \tag{2.63}$$

similar to (2.60) and (2.61).

Using the notations of Theorem G we first define the set $A_{n,k}$ for non-negative integers n and k:

$$A_{n,k} = \left\{\forall i(\tau_i^n < T)\,\exists j\,:\,\tau_i^n \le s_k(j) < \tau_{i+1}^n \wedge T\right\}.$$

Since the sequence $\{s_{n+j}(k)\}_k$ is a refinement of the sequence $\{s_n(k)\}_k$ and $\{s_m(k)\}_k$ is getting dense on every compact interval as n goes to infinity, there exists a positive integer $l(n,k)$ such that for all $l \ge l(n,k)$ the property $A_{n,k} \subset A_{n+1,l}$ holds. Moreover, for fixed n the probability $\mathbf{P}(A_{n,k})$ tends to 1 as $k \to \infty$.

We introduce the following sequences of positive integers $\{k_n\}$ and embedded sets $\{\mathcal{A}_n\}$:

$$k_0 = 0,\ \mathcal{A}_0 = \emptyset,$$

$$k_{n+1} = \inf\left\{k > k_n \,\middle|\, \mathbf{P}(A_{n+1,k}) \ge 1 - \frac{2^{-2n}}{C(n)},\ \mathcal{A}_n \subset A_{n+1,k}\right\},\ \mathcal{A}_{n+1} := A_{n+1,k_{n+1}}, \tag{2.64}$$

where $C(n) = 12\left(\frac{4}{3}\right)^4 T^2 M^4$. By the remarks in the previous paragraph $\{k_n\}$ and $\{\mathcal{A}_n\}$ are well defined and has the following two properties $\mathcal{A}_{n-1} \subset \mathcal{A}_n$ and $\mathbf{P}(\mathcal{A}_n) \to \infty$ as $n \to \infty$.

Let

$$U_n = \sup_{0\le t\le T}\left|I_t^n - \int_0^t f\,\mathrm{d}W\right|.$$

We will prove that U_{k_n} tends to 0 almost surely as n goes to infinity.

Using the triangle-inequality one gets

$$\mathbf{E}\left(\sup_{[0,T]}|U_{k_n}|^2\right) \le \mathbf{E}\left(\sup_{[0,T]}\left|Y_t^n - \int_0^t f\,\mathrm{d}W\right|^2\right) + \mathbf{E}\left(\sup_{[0,T]}\left|I_t^{k_n} - Y_t^n\right|^2\right).$$

(For the definition of Y_t^n see Theorem G.) By the definition of $\mathcal{A}_n$ one finds that for all $t \le T$

$$|f_t^n - f_t^{n*}| \le 2^{-n} \text{ on the set } \mathcal{A}_n$$

where f_t^n is defined in the proof of Theorem G. Using this inequality and (2.54) one gets the following estimate:

$$\begin{aligned}\mathbf{E}\left(\sup_{[0,T]}\left|I_t^{k_n} - Y_t^n\right|^2\right) &\le \mathbf{E}\left(\sup_{[0,T]}\left|I_t^{k_n} - Y_t^n\right|^2 \mathbf{I}\{\mathcal{A}_n\}\right) + \mathbf{E}\left(\sup_{[0,T]}\left|I_t^{k_n} - Y_t^n\right|^2 \mathbf{I}\{\mathcal{A}_n^c\}\right)\\ &\le 4\mathbf{E}\left|\int_0^T (f_t^n - f_t^{n*})^2\,\mathrm{d}t\right| + \mathbf{P}(\mathcal{A}_n^c)\cdot\mathbf{E}\left(\sup_{[0,T]}\left|I_t^{k_n} - Y_t^n\right|^4\right)\\ &\le 4T2^{-2n} + 2^{-2n}.\end{aligned} \tag{2.65}$$

The last inequality is valid because of the definition of $\mathcal{A}_n$ and the following consideration on the last expectation.

$$\mathbf{E}\left(\sup_{[0,T]}\left|I_t^{k_n} - Y_t^n\right|^4\right) \le \left(\frac{4}{3}\right)^4 \mathbf{E}\left(I_T^{k_n} - Y_T^n\right)^4$$

because $\left(I_t^{k_n} - Y_t^n\right)^4$ is submartingale. Simple calculation on stochastic integrals shows that

$$\mathbf{E}\left(I_T^{k_n} - Y_T^n\right)^4 \le 3\mathbf{E}\left(\int_0^T (f_s^{n*} - f_s^n)^2\,\mathrm{d}s\right)^2.$$

Using the condition $\max_{[0,T]}|f_t| \le M$ a.s. we get that the last expression is smaller than $4T^4M^4$ so we get

$$\mathbf{E}\left(\sup_{[0,T]}\left|I_t^{k_n}-Y_t^n\right|^4\right) \le 12\left(\frac{4}{3}\right)^4 T^2M^4.$$

Summarizing these estimation we have

$$\mathbf{E}\left(\sup_{[0,T]}|U_{k_n}|^2\right) \le (8T+1)2^{-2n}.$$

This implies that $\mathbf{E}U_{k_n} \le \sqrt{8T+1}\,2^{-n}$ and hence it follows that

$$\mathbf{E}\sum_n U_{k_n} = \sum_n \mathbf{E}U_{k_n} \le \sum_n \sqrt{8T+1}\,2^{-n} < \infty.$$

As a consequence, one gets

$$\sum_n U_{k_n} < \infty \text{ a.s.}$$

which gives the required conclusion.

Now, we will prove that $U_m \to 0$ almost surely as m goes to infinity. Let n be such that $k_n \le m < k_{n+1}$. Then

$$U_m \le U_{k_n} + \sup_{[0,T]}\left|I_t^{k_n}-I_t^m\right|.$$

Since the sequence $\{s_m(j)\}_j$ is a refinement of the sequence $\{s_{k_n}(j)\}_j$ one has $\mathcal{A}_{n-1} \subset A_{n,m}$. Therefore,

$$\left|f_t^{k_n}-f_t^m\right| \le 2\cdot 2^{-n+1} \text{ on the set } \mathcal{A}_{n-1}$$

so

$$\sup_{[0,T]}\left|I_t^{k_n}-I_t^m\right| \le T2^{-n+2} \text{ on } \mathcal{A}_{n-1}$$

from which one obtains

$$U_m \le U_{k_n} + T2^{-n+2} \text{ on } \mathcal{A}_{n-1}.$$

By the definition of $\{\mathcal{A}_n\}_n$, $\mathbf{P}(\mathcal{A}_n) \to 1$ and $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ so we get U_m tends to 0 a.e. as it was required.

□

The following general form of Theorem 7 can be obtained by using the *DDS*-construction.

Theorem 8. *Let M be continuous $(\mathcal{F}_t)$-martingale with quadratic variation such that $\langle M\rangle_\infty = \infty$. Let Y be an r.c.l.l. $(\mathcal{F}_t)$-adapted process and for $n \ge 1$ let $\{\tau_n(i):\ i \ge 0\}$ be the stopping times defined in (2.13).*

Let (Y_t^n) and (I_t^n) be defined as follows. For arbitrary t let k be such that $\tau_n(k) \le t < \tau_n(k+1)$, $k \ge 0$ let us define

$$Y_t^n := Y_{\tau_n(k)} \tag{2.66}$$

$$I_t^n = \int_0^t Y_s^n\,\mathrm{d}M_s = \sum_{i=0}^{k-1} Y_{\tau_n(i)}(M_{\tau_n(i+1)}-M_{\tau_n(i)}) + Y_{\tau_n(k)}(M_t - M_{\tau_n(k)}).$$

Then, for all $K < \infty$ we have

$$\sup_{0\le t\le K}\left|I_t^n - \int_0^t Y\,\mathrm{d}M\right| \to 0 \qquad a.s. \qquad (n\to\infty). \tag{2.67}$$

The next important corollary is a simple consequence of this theorem.

Corollary 4. *Under the assumption of Theorem 8 one can obtain the following convergence result:*

$$\sup_{0\le t\le K}\left|\sum_{\tau_n(i+1)\le t} Y_{\tau_n(i)}(M_{\tau_n(i+1)}-M_{\tau_n(i)})-\int_0^t Y\,\mathrm{d}M\right|\to 0 \qquad a.s. \qquad (n\to\infty).$$

This corollary implies that I_t^n can be thought of as a stochastic sum of stochastically weighted, independent, identically distributed, symmetrical and $\pm\frac{1}{2^m}$ valued random variables.

Proof. We want to apply Theorem 7. For, consider the following assertion for an $(\mathcal{F}_t)$-progressive process Y:

$$\int_0^t Y_s\,\mathrm{d}M_s=\int_0^t Y_{T_{\langle M\rangle_s}}\,\mathrm{d}M_s=\int_0^t Y_{T_{\langle M\rangle_s}}\,\mathrm{d}B_{\langle M\rangle_s}=\int_0^{\langle M\rangle_t} Y_{T_s}\,\mathrm{d}B_s \tag{2.68}$$

where T is the quasi-inverse of $\langle M\rangle$ and B is the *DDS*-Brownian motion of M. The first equality holds because the intervals of constancy are the same for M and for $\langle M\rangle$ so on the intervals $[s,T_{\langle M\rangle_s}]$ M is constant. The third equality follows from Proposition 1.4 in [23, Chapter V.] for C-continuous time-changed stochastic integrals:

Lemma G. *If $C=(C_u)$ is a time-change and H is an $(\mathcal{F}_t)$-progressive process, then H_{C_u} is $(\mathcal{F}_{C_u})$-progressive and if X is a C-continuous process of finite variation, then*

$$\int_{C_0}^{C_t} H_s\,\mathrm{d}X_s=\int_0^t \mathbf{I}\{C_u<\infty\}\,H_{C_u}\,\mathrm{d}X_{C_u}.$$

*Here a process X is said to be C-*continuous *if X is constant on each interval $[C_{u-},C_u]$.*

In our present case B is $\langle M\rangle$-continuous and $\langle M\rangle_\infty<\infty$ so (2.68) follows.

Let $\sigma_k:=\inf\{t>0\,|\,\langle M\rangle_t>k\}$ and denote by M^{σ_k} the stopped martingale. Using (2.68) one obtains

$$\begin{aligned}\sup_{0\le t\le K\wedge\sigma_k}\left|\int_0^t Y_s^n\,\mathrm{d}M_s-\int_0^t Y_s\,\mathrm{d}M_s\right|&=\sup_{0\le t\le K}\left|\int_0^t Y_s^n\,\mathrm{d}M_s^{\sigma_k}-\int_0^t Y_s\,\mathrm{d}M_s^{\sigma_k}\right|\\&=\sup_{0\le t\le\langle M^{\sigma_k}\rangle_K}\left|\int_0^t Y_{T_s}^n\,\mathrm{d}B_s-\int_0^t Y_{T_s}\,\mathrm{d}B_s\right|\le\sup_{0\le t\le k}\left|\int_0^t Y_{T_s}^n\,\mathrm{d}B_s-\int_0^t Y_{T_s}\,\mathrm{d}B_s\right|.\end{aligned}\tag{2.69}$$

Here, $\int_0^{\langle M\rangle_t} Y_{T_s}^n\,\mathrm{d}B_s$ is equal to the following sum:

$$\begin{aligned}\int_0^{\langle M\rangle_t} Y_{T_s}^n\,\mathrm{d}B_s&=\int_0^t Y_s^n\,\mathrm{d}M_s=\sum_{i=0}^{k-1} Y_{\tau_n(i)}(M_{\tau_n(i+1)}-M_{\tau_n(i)})+Y_{\tau_n(k)}(M_t-M_{\tau_n(k)})\\&=\sum_{\tau_n(i+1)<\langle M\rangle_t} Y_{T_{\langle M\rangle_{\tau_n(i)}}}(B_{\langle M\rangle_{\tau_n(i+1)}}-B_{\langle M\rangle_{\tau_n(i)}})+Y_{T_{\langle M\rangle_{\tau_n(k)}}}(B_{\langle M\rangle_t}-B_{\langle M\rangle_{\tau_n(k)}})\end{aligned}$$

by (2.68). By Lemma 2, this last expression can be written in the form of

$$\begin{aligned}&\sum_{\langle M\rangle_{s_n(i+1)}<\langle M\rangle_t} Y_{T_{s_n(i)}}(B_{s_n(i+1)}-B_{s_n(i)})+Y_{T_{s_n(i)}}(B_{\langle M\rangle_t}-B_{s_n(k)})\\&\qquad=\sum_{s_n(i+1)<t} Y_{T_{s_n(i)}}(B_{s_n(i+1)}-B_{s_n(i)})+Y_{T_{s_n(i)}}(B_{\langle M\rangle_t}-B_{s_n(k)})=\int_0^t (Y_T)_s^{n*}\,\mathrm{d}B_s\end{aligned}$$

with the notation of (2.63). So the last expression of (2.68) is of the form of

$$\sup_{0\le t\le k}\left|\int_0^t (Y_T)_s^{n*}\,\mathrm{d}B_s - \int_0^t Y_{T_s}\,\mathrm{d}B_s\right|$$

which fulfils the condition of Theorem 7. Therefore, it converges to zero as n goes to infinity.

Since $\langle M\rangle_\infty = \infty$, σ_k also tends to infinity as k tends to infinity. Applying this and the previous conclusion of Theorem 7 we get

$$\sup_{0\le t\le K}\left|\int_0^t Y_s^n\,\mathrm{d}M_s - \int_0^t Y_s\,\mathrm{d}M_s\right| \to 0, \qquad \text{a.s.}$$

which was required.

□

2.5.3 The non cadlag-case

Karandikar's approach provides a method for approximating stochastic integrals where the integrand is cadlag process. Therefore, we cannot apply his method for integrals in which the integrand is the sign function of the Brownian motion $\operatorname{sign}(B_t)$ which is often used as a basis of examples in stochastic analysis. In this case we can carry out our approximating method which is applicable on a much smaller class of integrands, namely for $f'_-(B)$ where f'_- is the derivative of the difference of two convex functions.

In [28] Tamás Szabados gave an approximation theorem using discrete Itô formula. This result was valid for integrals like $\int f(B)\,\mathrm{d}B$ for $f\in C^2$, i.e. for two times continuously differentiable real functions.

Using Theorem 3 and Itô-Tanaka formula [23] one could prove the following more general statement of this kind.

Theorem 9. *Let f be a difference of two convex functions and let M be a continuous local martingale such that $\langle M\rangle_\infty = \infty$ almost surely. Then for arbitrary $K>0$*

$$\sup_{t\in[0,K]}\left|\int_0^t f'_-(M_m(s))\,\mathrm{d}M_m(s) - \int_0^t f'_-(M(s))\,\mathrm{d}M(s)\right| \to 0 \tag{2.70}$$

almost surely as m tends to infinity.

Proof. Basically, we follow the method of the proof of Itô-Tanaka theorem [23, Theorem 1.5, Ch VI].

It is enough to prove the formula for a convex f. On every compact interval I f is equal to a convex function g such that g' has compact support. Thus by stopping M and M_m when they first leave a compact set, it suffices to prove the statement when f'' has compact support in which case there are two constants α_I and β_I such that

$$f(x) = \alpha_I x + \beta_I + \frac{1}{2}\int |x-a| f''(\,\mathrm{d}a). \tag{2.71}$$

First, we prove the statement (2.70) for the Brownian case, that is, $M=B$. Here, we have to remark that the stopping does not change the convergence rate and the corresponding probabilities in Theorem A so we have the same convergence result for the stopped processes B and B_m.

Proper usage of the above equation leads us to the Itô-Tanaka formula for Brownian motion B

$$f(B(t)) = f(B(0)) + \int_0^t f'_-(B(s))\,\mathrm{d}B(s) + \frac{1}{2}\int_{\mathbb{R}} \mathcal{L}_t^a f''(\,\mathrm{d}a). \tag{2.72}$$

For having the discrete version of this equation we need the following equation

$$\int_0^t \operatorname{sign}(B_m(s)-a)\,\mathrm{d}B_m(s)$$
$$= |B_m(t)-\lceil a\rceil_m| - \lceil a\rceil_m - \mathcal{L}_m^{\lceil a\rceil_m}(t) + O(2^{-m}) = |B_m(t)-a| - a - \mathcal{L}_m^a(t) + O(2^{-m}). \quad (2.73)$$

where $\lceil a\rceil_m = \frac{\lceil a2^m\rceil}{2^m}$. Here, we used the definition of $\mathcal{L}_m^a(t) = \mathcal{L}_m^{\lceil a\rceil_m}(t)$ (see Corollary 3)and that $|B_m(t)-\lceil a\rceil_m| - |B_m(t)-a| = O(2^{-m})$ where $|O(2^{-m})| \leq 2^{-m}$. Finally, applying the last three and the following statement

$$f'_-(x) = \frac{1}{2}\int_I \operatorname{sign}(x-a) f''(\,\mathrm{d}a) + \alpha_I$$

we obtain the equation

$$f(B_m(t)) = \alpha_I B_m(t) + \beta_I + \frac{1}{2}\int_{\mathbb{R}} |B_m(t)-a|\, f''(\,\mathrm{d}a)$$
$$= \alpha_I(B_m(t)-B_m(0)) + f(B_m(0)) + \int_{\mathbb{R}} \frac{1}{2}\left(\int_0^t \operatorname{sign}(B_m(s)-a)\,\mathrm{d}B_m(s) + \mathcal{L}_m^a(t) + O(2^{-m})\right) f''(\,\mathrm{d}a)$$
$$= \alpha_I B_m(t) + \int_0^t f'_-(B_m(s))\,\mathrm{d}B_m(s) - \alpha_I(B_m(t)-B_m(0)) + \int_{\mathbb{R}} \mathcal{L}_m^a(t) f''(\,\mathrm{d}a) + \int_{\mathbb{R}} O(2^{-m}) f''(\,\mathrm{d}a)$$
$$= \int_0^t f'_-(B_m(s))\,\mathrm{d}B_m(s) + \int_{\mathbb{R}} \mathcal{L}_m^a(t) f''(\,\mathrm{d}a) + \int_{\mathbb{R}} O(2^{-m}) f''(\,\mathrm{d}a).$$

Therefore, we can write

$$\sup_{t\in[0,K]} \left| \int_0^t f'_-(B_m(s))\,\mathrm{d}B_m(s) - \int_0^t f'_-(B(s))\,\mathrm{d}B(s) \right|$$
$$\leq \sup_{t\in[0,K]} |f(B_m(t)) - f(B(t))| + \int_{\mathbb{R}} \sup_{t\in[0,K]} |\mathcal{L}_m^{\lceil a\rceil_m}(t) - \mathcal{L}^a(t)| f''(\,\mathrm{d}a)$$
$$+ \int_{\mathbb{R}} \sup_{t\in[0,K]} O(2^{-m}) f''(\,\mathrm{d}a).$$

On the right hand side each terms converge almost surely to 0 as m tends to infinity. For the firs term this is trivial, the convergence of the second and the third term can be proved using (2.50) and the Lebesgue theorem. So we get

$$\sup_{t\in[0,K]} \left| \int_0^t f'_-(B_m(s))\,\mathrm{d}B_m(s) - \int_0^t f'_-(B(s))\,\mathrm{d}B(s) \right| \to 0 \quad (2.74)$$

almost surely as m tends to infinity.

Now, let us investigate the general case. We reduce it to the Brownian case by using DDS construction and the scaling property of the local time.

Let us introduce the notation $\mathcal{L}_X^a(t)$. It denotes the local time of the process X at level a till time t. By Exercise 1.27 in [23, Ch VI] we have $\mathcal{L}_M^a(t) = \mathcal{L}_B^a(\langle M\rangle_t) = \mathcal{L}^a(\langle M\rangle_t)$ where B denotes the DDS-Brownian motion of M. Simple consideration shows that the discrete version of this identity is also valid: $\mathcal{L}_{M_m,\,m}^a(t) = \mathcal{L}_{B_m,\,m}^a(N_m(t)) = \mathcal{L}_m^{\lceil a\rceil_m}(N_m(t))$

Let σ_k be the first time when $\langle M\rangle$ exceeds the level k. We can write the following estimate for the stopped martingale $M^k(t) = M(t\wedge\sigma_k)$:

$$\sup_{t\in[0,K]}\left|\int_0^t f'_-(M^k_m(s))\,\mathrm{d}M^k_m(s) - \int_0^t f'_-(M^k(s))\,\mathrm{d}M^k(s)\right|$$
$$\leq \sup_{t\in[0,(\langle M,M\rangle_K\wedge k)+1]} \left|\int_0^t f'_-(B_m(s))\,\mathrm{d}B_m(s) - \int_0^t f'_-(B(s))\,\mathrm{d}B(s)\right|$$

on the set $A_{a,m}$ (for the precise definition of $A_{a,m}$ see Lemma 3) after an appropriately large m. Since (2.74), M^k fulfils (2.70) for all fixed k. Taking into consideration that $\langle M\rangle_\infty = \infty$ (so $\sigma_k\to\infty$), (2.70) is valid for M as well. □

Chapter 3

Approximation of the exponential functional of Brownian motion

This chapter is devoted to studying a special application of our discrete approximation method.

Here, we investigate the exponential functional of Brownian motion

$$\mathcal{I}_\nu = \int_0^\infty \exp\left(B(t) - \nu t\right) \mathrm{d}t$$

and mainly its discrete version the exponential functional of random walk.

3.1 Introduction to the exp functional of Brownian motion

The geometric Brownian motion (originally introduced by the economist P. Samuelson in 1965) plays a fundamental role in the Black–Scholes theory of option pricing, modeling the price process of a stock. It can be explicitly given in terms of Brownian motion (BM) B as

$$S(t) = S_0 \exp\left(\sigma B(t) + \left(\mu - \sigma^2/2\right) t\right), \qquad t \geq 0.$$

In the case of Asian options one is interested in the average price process

$$A(t) = \frac{1}{t} \int_0^t S(u)\, \mathrm{d}u, \qquad t \geq 0.$$

The following interesting result is true for the distribution of a closely related, widely investigated exponential functional of BM:

$$\mathcal{I} = \int_0^\infty \exp(B(t) - \nu t)\, \mathrm{d}t \stackrel{\mathrm{d}}{=} \frac{2}{Z_{2\nu}} \qquad (\nu > 0). \tag{3.1}$$

Here $Z_{2\nu}$ is a gamma distributed random variable with index 2ν and parameter 1, while $\stackrel{\mathrm{d}}{=}$ denotes equality in distribution. This result was proved by [7] using discrete approximations with gamma distributed random variables and also by [38], using rather ingenious stochastic analysis tools. For more background information see [39] and [2].

As a consequence, the pth integer moment of $\mathcal{I}$ is finite iff $p < 2\nu$ and

$$\mathbf{E}(\mathcal{I}^p) = 2^p \frac{\Gamma(2\nu - p)}{\Gamma(2\nu)}. \tag{3.2}$$

On the other hand, all negative integer moments, also given by (3.2), are finite and they characterize the distribution of $\mathcal{I}$.

The situation is much nicer when BM with negative drift is replaced in the model by the negative of a subordinator $(\alpha_t, t \geq 0)$, that is, by the negative of a non-decreasing process with stationary and independent increments, starting from the origin. Then, as was shown by [1], all positive integer moments of $\mathcal{J} = \int_0^\infty \exp(-\alpha_t)\,\mathrm{d}t$ are finite:

$$\mathbf{E}(\mathcal{J}^p) = \frac{p!}{\Phi(1)\cdots\Phi(p)}, \qquad \Phi(\lambda) = -\frac{1}{t}\log \mathbf{E}(\exp(-\lambda\alpha_t)), \tag{3.3}$$

and in this case the positive integer moments characterize the distribution of $\mathcal{J}$.

To achieve a similar favorable situation in the BM case, at least in an approximate sense, it is a natural idea to use a simple, symmetric random walk (RW) as an approximation, with a large enough negative drift. Besides, in some applications a discrete model could be more natural than a continuous one. It seems important that, as we shall see below, the discrete case is rather different from the continuous case in many respects.

So let $(X_j)_{j=1}^\infty$ be an i.i.d. sequence with $\mathbf{P}(X_1 = \pm 1) = \frac{1}{2}$ and $S_0 = 0$, $S_k = \sum_{j=1}^k X_j$ $(k \geq 1)$. Introduce the following approximation of $\mathcal{I}$:

$$Y = \sum_{k=0}^\infty \exp(S_k - k\nu) = 1 + \xi_1 + \xi_1\xi_2 + \cdots, \qquad \xi_j = \exp(X_j - \nu), \tag{3.4}$$

where $\nu > 0$. Later, we apply proper scaling to get a real approximation of $\mathcal{I}$. In this paper we investigate the properties of Y type random variables which, in this simple case, will be called the discrete exponential functional of the given RW, or shortly, the discrete exponential functional.

In Section 3.2, below it turns out that the distribution of Y is singular w.r.t. Lebesgue measure if $\nu > 1$. Here, we prove a more general result according to the singularity of the distribution of Y if ξ_j above has more general distribution.

In Section 3.3 we determine the moments of the discrete exponential functional in order to work out a (3.2) and a (3.3) type equations for the discrete case. Dealing with the moments, beyond these results we have found a recursion of certain moments in the expansion of the moments of the discrete approximation. We will describe this recursion in Subsection 3.3.1.

Finally, in Section 3.4 section we use a nested sequence of RWs to obtain a.s. converging approximations of $\mathcal{I}$, and this way an elementary proof of result (3.1) of Dufresne and Yor as well.

3.2 The distribution of the discrete exponential functional

Let us start with a natural generalization: $(\xi_j)_{j=1}^\infty$ be i.i.d., $\xi_j > 0$. Consider first the finite polynomial

$$\begin{aligned} Y_n &= 1 + \xi_1 + \xi_1\xi_2 + \cdots + \xi_1\cdots\xi_n \\ &= 1 + \xi_1(1 + \xi_2 + \xi_2\xi_3 + \cdots + \xi_2\cdots\xi_n) \qquad (n \geq 1), \end{aligned} \tag{3.5}$$

$Y_0 = 1$. This implies the following equality in distribution:

$$Y_n \stackrel{\mathsf{d}}{=} 1 + \xi Y_{n-1}, \tag{3.6}$$

where $\xi \stackrel{\mathsf{d}}{=} \xi_1$, and ξ is independent of Y_{n-1}. Since $Y_n \nearrow Y = 1 + \xi_1 + \xi_1\xi_2 + \cdots$ a.s., we get the basic *self-similarity* of Y in distribution:

$$Y \stackrel{\mathsf{d}}{=} 1 + \xi Y, \tag{3.7}$$

where ξ is independent of Y. We remark that infinite polynomials similar to Y were studied by [35] and many others. There some of the ideas discussed below have already appeared.

A standard application of the strong law of large numbers gives a condition for having an a.s. finite limit Y here, see Theorem 1 in [33]. Namely, when $\mathbf{E}(|\log \xi_j|) < \infty$, one has $Y_n \nearrow Y < \infty$ a.s. if and only if $\mathbf{E}(\log \xi_j) < 0$.

In the special case when Y is defined as in (3.4), but S_n is the partial sum of an *arbitrary* i.i.d. sequence $(X_j)_{j=1}^{\infty}$ with zero expectation, $Y < \infty$ a.s. iff the drift added is negative: $\nu > 0$. Hence this condition is always assumed in our basic example (simple, symmetric RW).

Next we want to show that self-similarity (3.7) implies a simple functional equation for the distribution function $F(y) = \mathbf{P}(Y \leq y)$, $y \in \mathbb{R}$. For a modest generalization of our basic case, let us introduce some notations.

3.2.1 Review of some fractal notion

In (3.5) let ξ_j take the positive values $\gamma_1 < \cdots < \gamma_N$, and let $p_i = \mathbf{P}(\xi = \gamma_i)$. (In our basic case $N = 2$, $\gamma_1 = e^{-1-\nu}$, $\gamma_2 = e^{1-\nu}$, $p_1 = p_2 = \frac{1}{2}$.) Consider the following similarity transformations: $T_i(x) = \gamma_i x + 1$ $(1 \leq i \leq N)$. When

$$\mathbf{E}(\log \xi) = \sum_{i=1}^{N} p_i \, \log \gamma_i < 0$$

holds, by (3.7) we have $\mathbf{P}(Y \leq y) = \mathbf{P}(1 + \xi Y \leq y) = \sum_{i=1}^{N} p_i \, \mathbf{P}(1 + \gamma_i Y \leq y | \xi = \gamma_i) = \sum_{i=1}^{N} p_i \, \mathbf{P}(1 + \gamma_i Y \leq y)$. Thus one obtains the following functional equation for the distribution function of Y:

$$F(y) = \sum_{i=1}^{N} p_i \, F(T_i^{-1}(y)). \tag{3.8}$$

An important special case is when $\gamma_N < 1$ (in the basic case: $\nu > 1$). Then by (3.5), Y is a bounded random variable. Moreover, each mapping T_i is a contraction, having a unique fixpoint $y_i = (1 - \gamma_i)^{-1}$, $0 < y_1 < \cdots < y_N < \infty$. Since each T_i is an increasing function, $T_i(y_j) < T_i(y_k)$ if $j < k$. Also, $T_j(y_i) < T_k(y_i)$ if $j < k$. Then it follows that each T_i maps the fundamental interval $I = [y_1, y_N]$ into itself. Clearly, I contains the range of Y too.

In the case $\gamma_N < 1$ it is useful to rephrase the given problem in the language of fractal theory, see e.g. [8]. Let us introduce the symbolic space $\Sigma = \{\underline{i} = (i_1, i_2, \dots) : i_j = 1, \dots, N\}$, endowed with the countable power of the discrete measure $(p_1, \dots, p_N)$, denoted by $\mathbf{P}$. By (3.5),

$$Y_n = 1 + \gamma_{i_1}(1 + \gamma_{i_2}(\cdots(1 + \gamma_{i_n}))) = (T_{i_1} \circ T_{i_2} \circ \cdots \circ T_{i_n})(1), \tag{3.9}$$

with probability $p_{i_1} p_{i_2} \cdots p_{i_n}$. Thus the canonical projection $\Pi : \Sigma \to I$, $\Pi(\underline{i}) = \lim_{k\to\infty}(T_{i_1} \circ \cdots \circ T_{i_k})(1) = \lim_{k\to\infty}(1 + \gamma_{i_1} + \cdots + \gamma_{i_1} \dots \gamma_{i_k})$ maps Σ onto the range of Y. The *attractor* Λ of the iterated function scheme of similarity transformations $(T_1, \dots, T_N)$ is defined as

$$\Lambda = \bigcap_{k \geq 0} \bigcup_{1 \leq i_1, \dots, i_k \leq N} (T_{i_1} \circ \cdots \circ T_{i_k})(I), \qquad \Lambda = \bigcup_{i=1}^{N} T_i(\Lambda). \tag{3.10}$$

Then Λ is a non-empty, compact, self-similar set. In (3.10) the fundamental interval $I = [y_1, y_N]$ can be replaced by any interval J which is mapped into itself by each T_i, e.g. by $J = [0, y_N] \ni 1$. Thus $\text{range}(Y) \subset \Lambda$, cf. (3.9). The converse is also true, since for any $y \in \Lambda$ and for any $\epsilon > 0$ there is an $\underline{i} \in \Sigma$ and a large enough k such that $y \in (T_{i_1} \circ \cdots \circ T_{i_k})(J)$ and the length $|(T_{i_1} \circ \cdots \circ T_{i_k})(J)| =$

$\gamma_{i_1} \cdots \gamma_{i_k} |J| < \epsilon$. Hence, by (3.9), $y \in$ range(Y), that is, Λ =range(Y). Also, the distribution of Y on the real line, which will be denoted by $\mathbf{P}_Y$, is simply $\mathbf{P} \circ \Pi^{-1}$.

We are going to use the notations

$$I_{i_1 \ldots i_k} = [y_{i_1 \ldots i_k 1}, y_{i_1 \ldots i_k N}] = (T_{i_1} \circ \cdots \circ T_{i_k})(I), \tag{3.11}$$

$y_{i_1 \ldots i_k l} = (T_{i_1} \circ \cdots \circ T_{i_k})(y_l)$ $(l = 1, \ldots, N)$ as well, where $i_j = 1, \ldots N$ and $y_l = (1 - \gamma_l)^{-1}$. The length of such an interval is $|I_{i_1 \ldots i_k}| = \gamma_{i_1} \cdots \gamma_{i_k} |I|$, where $|I| = y_N - y_1$.

3.2.2 The non-overlapping case

Returning to the distribution of Y in the basic case, consider first when the intervals $I_1 = T_1(I) = [y_{11}, y_{12}] = [y_1, y_{12}]$ and $I_2 = T_2(I) = [y_{21}, y_{22}] = [y_{21}, y_2]$ do not overlap, where $y_{12} = 1 + \gamma_1 (1 - \gamma_2)^{-1}$, $y_{21} = 1 + \gamma_2 (1 - \gamma_1)^{-1}$. Thus there is no overlap iff $y_{12} < y_{21}$, i.e., $\nu > \log(e + e^{-1}) \approx 1.127$. Since $F(y) = 0$ if $y < y_1$ and $F(y) = 1$ if $y \geq y_2$, in this non-overlapping case (3.8) simplifies as

$$F(y) = \begin{cases} \frac{1}{2} F(T_1^{-1}(y)) & \text{if } y \in [y_1, y_{12}), \\ \frac{1}{2} & \text{if } y \in [y_{12}, y_{21}), \\ \frac{1}{2} + \frac{1}{2} F(T_2^{-1}(y)) & \text{if } y \in [y_{21}, y_2). \end{cases} \tag{3.12}$$

By the similarities given by T_1 and T_2, applied to (3.12), one obtains that F has constant value $\frac{1}{4}$ over the interval $[y_{112}, y_{121})$ and constant value $\frac{3}{4}$ on $[y_{212}, y_{221})$. Continuing this way by induction one gets that F has constant dyadic values over such *plateau* intervals:

$$F(y) = 2^{-k-1} + \sum_{j=1}^{k} (i_j - 1) 2^{-j}, \qquad y \in [y_{i_1 \ldots i_k 12}, y_{i_1 \ldots i_k 21}), \qquad i_j = 1, 2. \tag{3.13}$$

The sum of the lengths of these plateaus is $|I| \left(1 - (\gamma_1 + \gamma_2)\right) \left(1 + (\gamma_1 + \gamma_2) \right. \left. + (\gamma_1 + \gamma_2)^2 \right. + \cdots)$, so add up to $|I|$. Hence the attractor Λ (the range of Y), i.e., the set of points of increase of F, has zero Lebesgue measure. So it is a Cantor-type set: an uncountable, perfect set of Lebesgue measure zero.

The distribution function F is clearly a continuous singular function. For, if $y_0 \in \Lambda$ and $\epsilon > 0$ is given, take k so that $2^{-k} < \epsilon$. By the construction of Λ, there exists an interval $I_{i_1 \ldots i_k} \ni y_0$. Let the left endpoint of the left neighbor plateau of $I_{i_1 \ldots i_k}$ be η_1 (or $-\infty$), and the right endpoint of the right neighbor plateau be η_2 (or ∞). If $\delta = \min(y_0 - \eta_1, \eta_2 - y_0) > 0$, then for any y such that $|y - y_0| < \delta$ one has $|F(y) - F(y_0)| \leq 2^{-k} < \epsilon$ by (3.13).

It is not difficult to see, cf. [9], that in general, any solution of (3.7) has either absolutely continuous or continuous singular distribution.

We mention that standard results of fractal theory, see Theorem 9.3 in [8], imply that the Hausdorff dimension s of Λ equals the (unique) solution of the equation $\gamma_1^s + \gamma_2^s = 1$. Solving this equation for ν, we get $\nu = s^{-1} \log(e^s + e^{-s})$. Hence the fractal dimension s is a strictly decreasing function of $\nu > \log(e + e^{-1})$, tending to 1 as $\nu \to \log(e + e^{-1})$ and converging to 0 as $\nu \to \infty$. Also, the Hausdorff measure of Λ is $\mathcal{H}^s(\Lambda) = |I|^s$, where the Hausdorff dimension s is the one defined above. It means that

$$\mathcal{H}^s(\Lambda) = \left(\left(1 - e(2\cosh s)^{-1/s}\right)^{-1} - \left(1 - e^{-1}(2\cosh s)^{-1/s}\right)^{-1} \right)^s .$$

Thus $\mathcal{H}^s(\Lambda) \to e^2 - e^{-2}$ as $\nu \to \log(e + e^{-1})$ and $\mathcal{H}^s(\Lambda) \to 0$ as $\nu \to \infty$.

3.2.3 The general case

Next we are going to show that the distribution of Y is singular w.r.t. Lebesgue measure even in the overlapping case if $\nu > 1$. Again, we consider the slight generalization introduced above. The proof below is based on [26] and on personal communication with K. Simon.

Theorem 10. *Let ξ take the values γ_i $(i = 1, \dots, N)$, $0 < \gamma_1 < \dots < \gamma_N < 1$, and let $p_i = \mathbf{P}(\xi = \gamma_i)$. Take an i.i.d. sequence $(\xi_j)_{j=1}^{\infty}$, $\xi_j \overset{\mathsf{d}}{=} \xi$. Then the distribution of $Y = 1 + \xi_1 + \xi_1\xi_2 + \cdots$ is singular w.r.t. Lebesgue measure, if*

$$-\chi_{\mathbf{P}} = \mathbf{E}(\log \xi) = \sum_{i=1}^{N} p_i \log \gamma_i < \sum_{i=1}^{N} p_i \log p_i = -H_{\mathbf{P}}.$$

This will be called the entropy condition. Here $\chi_{\mathbf{P}}$ is the Lyapunov exponent of the iterated function scheme $(T_1, \dots, T_N)$ corresponding to the Bernoulli measure $\mathbf{P}$.

Proof. We are going to use the fractal theoretical approach and notations introduced above.

We want to show that

$$(\bar{D}\mathbf{P}_Y)(x) = \limsup_{r \searrow 0} \frac{\mathbf{P}_Y(B(x,r))}{\lambda(B(x,r))} = \infty \qquad \mathbf{P}_Y \text{a.s.}, \tag{3.14}$$

where $B(x,r)$ denotes the open ball (in the real line) with center at x and radius r and λ is Lebesgue measure. The statement of the theorem easily follows from this. For, take the set $E = \{x \in I : (\bar{D}\mathbf{P}_Y)(x) = \infty\}$. Then (3.14) implies that $\mathbf{P}_Y(E) = 1$, while e.g. Theorem 8.6 in [24] shows that the symmetric derivative $D\mathbf{P}_Y$ exists and is finite λ a.e., so $\lambda(E) = 0$.

Introduce the notation $a_k^{(j)}(\underline{i}) = \#\{l : i_l = j, 1 \le l \le k\}$. Thus

$$\Pi(\underline{i}) = 1 + \sum_{k=1}^{\infty} \prod_{j=1}^{N} \gamma_j^{a_k^{(j)}(\underline{i})}.$$

By the SLLN, the set $A_j = \left\{ \underline{i} \in \Sigma : k^{-1} a_k^{(j)}(\underline{i}) \to p_j \right\}$ has probability 1 for every $j = 1, \dots, N$ and so has $A = \bigcap_{j=1}^{N} A_j$. Let $C = \{x \in I : \Pi^{-1}(x) \cap A \neq \emptyset\}$. Then $\mathbf{P}_Y(C) = 1$.

If $x \in C$, there exists $\underline{i} \in A$ such that $\Pi(\underline{i}) = x$ and $k^{-1} a_k^{(j)}(\underline{i}) \to p_j$ as $k \to \infty$ for all $j = 1, \dots, N$. Fix such an $\underline{i}$ and x. Let r_k be the smallest radius such that $B(x, r_k) \supset I_{i_1 \dots i_k}$, where $\underline{i} = (i_1, \dots, i_k, \dots)$ and $I_{i_1 \dots i_k}$ is defined by (3.11).

The following facts are clear: (a) $x \in \Lambda$, moreover, $x \in I_{i_1 \dots i_k}$, see (3.10) and (3.11); (b) $\frac{1}{2}|I_{i_1 \dots i_k}| < r_k \le c|I_{i_1 \dots i_k}|$, where $c > 1$ is arbitrary; (c) $|I_{i_1 \dots i_k}| = |I| \prod_{j=1}^{N} \gamma_j^{a_k^{(j)}(\underline{i})}$; (d) $\mathbf{P}_Y(B(x, r_k)) \ge \mathbf{P}_Y(I_{i_1 \dots i_k}) = \mathbf{P}(i_1, \dots, i_k) = \prod_{j=1}^{N} p_j^{a_k^{(j)}(\underline{i})}$. Using these facts it follows for any $k \ge 1$ that

$$\frac{\mathbf{P}_Y(B(x, r_k))}{\lambda(B(x, r_k))} \ge (2c|I|)^{-1} \left(\frac{\prod_{j=1}^{N} p_j^{k^{-1} a_k^{(j)}(\underline{i})}}{\prod_{j=1}^{N} \gamma_j^{k^{-1} a_k^{(j)}(\underline{i})}} \right)^k.$$

By our assumptions concerning x and $\underline{i}$, the ratio on the right hand side converges to

$$(p_1^{p_1} \cdots p_N^{p_N}) \,/\, (\gamma_1^{p_1} \cdots \gamma_N^{p_N})$$

as $k \to \infty$. The entropy condition of the theorem implies that this latter ratio is larger than 1. Hence (3.14) holds, and this completes the proof. $\square$

Returning to our basic case, consider the entropy condition when $\gamma_1 = e^{-1-\nu}$, $\gamma_2 = e^{1-\nu}$ and $p_1 = p_2 = \frac{1}{2}$. The condition holds iff $\nu > \log 2 \approx 0.693$, since this is equivalent to $\frac{1}{2}(-1-\nu)+\frac{1}{2}(1-\nu) < \frac{1}{2}\log\frac{1}{2} + \frac{1}{2}\log\frac{1}{2}$. Combining this with the condition $\gamma_2 < 1$, this means that the distribution of Y is singular w.r.t. Lebesgue measure for any $\nu > 1$.

Characterization of the distribution of Y when $0 < \nu \leq 1$ remains open. In that case one of the two similarity mappings, T_2, is not a contraction anymore, and that situation requires more sophisticated tools than the ones above.

3.3 The moments of the discrete exponential functional

Let us consider first the general case: $(\xi_j)_{j=1}^{\infty}$ i.i.d., $\xi_j > 0$, as at the beginning of the previous section. Now we turn our attention to the moments of $Y = 1 + \xi_1 + \xi_1\xi_2 + \cdots$. If Y_n is defined by (3.5), the equality in law (3.6) implies

$$\mathbf{E}(Y_n^p) = \mathbf{E}\left((1 + \xi Y_{n-1})^p\right) = \sum_{k=0}^{p} \binom{p}{k} \mu_k \mathbf{E}(Y_{n-1}^k), \tag{3.15}$$

where $p \geq 0$ integer and $\mu_k = \mathbf{E}(\xi^k)$. As $n \to \infty$, by monotone convergence we obtain

$$\mathbf{E}(Y^p) = \sum_{k=0}^{p} \binom{p}{k} \mu_k \mathbf{E}(Y^k). \tag{3.16}$$

In (3.15) and (3.16) both sides are either finite positive, or $+\infty$.

Proposition 3. *Let $(\xi_j)_{j=1}^{\infty}$ be an i.i.d. sequence, $\xi_j > 0$, and $Y = 1 + \xi_1 + \xi_1\xi_2 + \cdots$. For $p \geq 1$ real, $\mathbf{E}(Y^p) < \infty$ if and only if $\mu_p = \mathbf{E}(\xi^p) < 1$. Then $\mu_q < 1$ for any $1 \leq q \leq p$ and $\mathbf{E}(Y^p) \leq (1 - \mu_p^{1/p})^{-p}$ as well. In this case if $p \geq 1$ is an integer, we also have the recursion formula*

$$\mathbf{E}(Y^p) = \frac{1}{1 - \mu_p} \sum_{k=0}^{p-1} \binom{p}{k} \mu_k \mathbf{E}(Y^k). \tag{3.17}$$

Proof. First, (3.6) and the simple inequality $(1 + x)^p \geq 1 + x^p$ ($x \geq 0$, $p \geq 1$, real) imply that $\mathbf{E}(Y_n^p) \geq 1 + \mu_p \mathbf{E}(Y_{n-1}^p)$. Suppose that $\mu_p \geq 1$. Since $\mathbf{E}(Y_0^p) = 1$, taking limit as $n \to \infty$, one gets that $\mathbf{E}(Y^p) = \infty$.

Conversely, suppose that $\mu_p < 1$. Then by Hölder's (or by Jensen's) inequality, $\mu_q \leq \mu_p^{q/p} < 1$ for any $1 \leq q \leq p$ as well. We want to show that $\mathbf{E}(Y^p)$ is finite. Let us begin by observing that $\mathbf{E}\left((Y_j - Y_{j-1})^p\right) = \mathbf{E}\left((\xi_1 \cdots \xi_j)^p\right) = \mu_p^j$ ($j \geq 1$). Hence by the triangle inequality and $Y_0 = 1$ we get that

$$\mathbf{E}(Y_n^p)^{1/p} \leq 1 + \sum_{j=1}^{n} \mathbf{E}\left((Y_j - Y_{j-1})^p\right)^{1/p} = \sum_{j=0}^{n} \mu_p^{j/p} < \frac{1}{1 - \mu_p^{1/p}},$$

for any $n \geq 1$ when $\mu_p < 1$. Taking limit as $n \to \infty$, it follows that $\mathbf{E}(Y^p) \leq (1 - \mu_p^{1/p})^{-p} < \infty$.

Finally, the recursion formula (3.17) directly follows from (3.16) when $\mu_p < 1$, $p \geq 1$ integer. □

There is a nice analogy between the moments of the exponential functional of a subordinator and the moments of Y, compare (3.3) and (3.18). First, the sum of the coefficients in the numerator of (3.18) is $p!$, as can be seen by induction. For, if one explicitly writes down $\mathbf{E}(Y^p)$, based on the recursion (3.17), taking a common denominator, the numerator of each earlier term except the last one is multiplied by factors $1 - \mu_k$. In the sum of the coefficients of the numerator it means a multiplication by zero. On the other hand, in the last term one multiplies the numerator of $\mathbf{E}(Y^{p-1})$ by $p\mu_{p-1}$, which results the sum $p!$ of the coefficients by the induction.

Second, there is a relationship between the denominators of (3.3) and (3.18) as well. In the special case when Y is defined as in (3.4), but S_n is the partial sum of an arbitrary i.i.d. sequence $(X_j)_{j=1}^{\infty}$ with zero expectation, $\Phi(\lambda) = -n^{-1} \log \mathbf{E}\left(\exp(\lambda(S_n - \nu n))\right) = -\log \mathbf{E}(\xi^{\lambda})$, so $\Phi(k) = -\log \mu_k$, corresponding to the factors in the denominator of (3.3). The factors $1 - \mu_k$ in the denominator of (3.18) are tangents to these.

Finally, let us consider the moments of Y in our *basic case*. Then $\mu_k = \mathbf{E}(\xi^k) = \exp(-k\nu)\cosh(k)$. Since $\cosh(k) < e^k$ for any $k > 0$, it follows that $\mu_k < 1$ for any $k \geq 1$ when $\nu \geq 1$, therefore all positive integer moments of Y are finite in this case by Proposition 3. In particular, in Section 2 we saw that Y is a bounded random variable if $\nu > 1$, hence the positive integer moments characterize its distribution. On the other hand, when $0 < \nu < 1$, only finitely many moments of Y are finite. For, $\mu_k \geq 1$ if $0 < \nu \leq k^{-1} \log\cosh(k) \nearrow 1$ as $k \to \infty$. For example, even $\mu_1 \geq 1$ (and consequently all $\mathbf{E}(Y^p) = \infty$) if $0 < \nu < \log\cosh(1) \approx 0.43378$.

3.3.1 Permutations with given descent set

For integer $p \geq 1$ it follows from (3.17) by induction that $\mathbf{E}(Y^p)$ is a rational function of the moments $\mu_1, \ldots, \mu_p$:

$$\mathbf{E}(Y^p) = \frac{1}{(1-\mu_1)\cdots(1-\mu_p)} \sum_{(j_1,\ldots,j_{p-1})\in\{0,1\}^{p-1}} a^{(p)}_{j_1,\ldots,j_{p-1}} \mu_1^{j_1} \cdots \mu_{p-1}^{j_{p-1}}, \tag{3.18}$$

where the coefficients of the numerator are ***universal*** constants, independent of the distribution of ξ_j.

These universal coefficients $a^{(p)}_{j_1 \cdots j_{p-1}}$ make a symmetrical, Pascal's triangle-like table if each row is listed in the increasing order of the binary numbers $j_{p-1}2^{p-2} + \cdots + j_1 2^0$, defined by the multiindices $(j_1, \ldots, j_{p-1})$, see the rows $p = 1, \ldots, 5$:

Table 3.1: The Pascal's triangle-like table of the coefficients

								1								
							0		1							
							1		1							
						00	01		10	11						
						1	2		2	1						
				000	001	010	011		100	101	110	111				
				1	3	5	3		3	5	3	1				
0000	0001	0010	0011	0100	0101	0110	0111		1000	1001	1010	1011	1100	1101	1110	1111
1	4	9	6	9	16	11	4		4	11	16	9	6	9	4	1

Two natural questions may arise at this point, independently of any probability theory background mentioned above. First, suppose that one defines a recursive sequence $(e_p)_{p=1}^{\infty}$ by (3.17) with coefficients $a^{(p)}_{j_1 \cdots j_{p-1}}$ given by (3.18). Can one attach any direct mathematical meaning to this coefficients $a^{(p)}_{j_1 \cdots j_{p-1}}$ then? The answer is yes, and rather surprisingly (as was conjectured in [30]), the coefficient $a^{(p)}_{j_1 \cdots j_{p-1}}$ is equal to the number of permutations $\pi \in S_p$ having descent $\pi(i) > \pi(i+1)$ exactly where $j_i = 1$, $1 \leq i \leq p-1$, cf. Theorem 11 below.

Second, can one give a direct way to evaluate the coefficients $a^{(p)}_{j_1 \cdots j_{p-1}}$? The affirmative answer to this question is partly included in the previous answer, since several formulae are known for the number of permutations with given descent sets. However, an apparently new recursion was conjectured in

[30], which is analogous to the recursion of binomial coefficients in the ordinary Pascal's triangle. The proof of this is the content of Lemma 8 below.

Lemma 7. *Fix a multiindex $(j_1, \ldots, j_{p-1}) \in \{0,1\}^{p-1}$. Let S be the set of indices k where $j_k = 1$:*

$$S = \{s_1, \ldots, s_m\} = \{k : j_k = 1, 1 \le k \le p-1\}, \quad m = \sum_{k=1}^{p-1} j_k. \tag{3.19}$$

Then the coefficient $a^{(p)}_{j_1 \cdots j_{p-1}}$ defined by (3.18) can be obtained by the recursion

$$\begin{aligned} a^{(p)}_{j_1 \cdots j_{p-1}} &= \sum_{k=0}^{p-1} \binom{p}{k} j_k (-1)^{j_{k+1}+\cdots+j_{p-1}} a^{(k)}_{j_1 \cdots j_{k-1}} \\ &= \sum_{l=0}^{m} \binom{p}{s_l} (-1)^{m-l} a^{(s_l)}_{j_1 \cdots j_{s_l - 1}}, \end{aligned} \tag{3.20}$$

where, by definition, $a^{(0)} = 1$, $j_0 = 1$, $s_0 = 0$ and -1 powered to an empty sum is 1.

Proof. The second equality in (3.20) is a direct consequence of the definitions above. To show the first equality, substitute (3.18) into (3.17):

$$e_p = \frac{1}{1-\mu_p} \sum_{k=0}^{p-1} \binom{p}{k} \frac{\mu_k}{(1-\mu_1)\cdots(1-\mu_k)} \sum_{(j_1,\ldots,j_{k-1}) \in \{0,1\}^{k-1}} a^{(k)}_{j_1 \cdots j_{k-1}} \mu_1^{j_1} \cdots \mu_{k-1}^{j_{k-1}}.$$

Here, multiplying by the common denominator and then collecting the coefficients of $\mu_1^{j_1} \cdots \mu_{p-1}^{j_{p-1}}$ for each $(j_1, \ldots, j_{p-1})$ we obtain

$$\begin{aligned} & e_p (1-\mu_1) \cdots (1-\mu_p) \\ &= \sum_{k=0}^{p-1} \sum_{(j_1,\ldots,j_{k-1}) \in \{0,1\}^{k-1}} \binom{p}{k} a^{(k)}_{j_1 \cdots j_{k-1}} \mu_1^{j_1} \cdots \mu_{k-1}^{j_{k-1}} \mu_k (1-\mu_{k+1}) \cdots (1-\mu_{p-1}) \\ &= \sum_{(j_1,\ldots,j_{p-1}) \in \{0,1\}^{p-1}} \mu_1^{j_1} \cdots \mu_{p-1}^{j_{p-1}} \sum_{k=0}^{p-1} \binom{p}{k} a^{(k)}_{j_1 \cdots j_{k-1}} j_k (-1)^{j_{k+1}} \cdots (-1)^{j_{p-1}}. \end{aligned}$$

This and (3.18) imply the first equality in (3.20). □

Now we turn to the proof of the equality of the coefficient $a^{(p)}_{j_1 \cdots j_{p-1}}$ given by (3.18) and the number of permutations $b^{(p)}(S)$ with descent set S given by (3.19). The descent set of a permutation $\pi \in S_p$ is defined as $D(\pi) = \{i : \pi(i) > \pi(i+1), 1 \le i \le p-1\}$. It is known, cf. [27, p. 69], that the number of permutations $\pi \in S_p$ with a given descent set $S = (s_1, \ldots, s_m)$, $1 \le s_1 < \cdots < s_m \le p-1$, can be obtained by the following inclusion-exclusion formula:

$$b^{(p)}(S) = b^{(p)}(s_1, \ldots, s_m) = \sum_{1 \le i_1 < \cdots < i_j \le k} (-1)^{m-j} \binom{p}{s_{i_1}, s_{i_2}, \ldots, p - s_{i_j}}. \tag{3.21}$$

Theorem 11. *The coefficient $a^{(p)}_{j_1 \cdots j_{p-1}}$ given by (3.18) is equal to the number of permutations $b^{(p)}(S)$ with descent set S given by (3.19).*

Proof. It is enough to show that the numbers $b^{(p)}(S)$ satisfy the same recursion (3.20) as the numbers $a^{(p)}_{j_1\cdots j_{p-1}}$ do, that is,

$$b^{(p)}(s_1,\ldots,s_m)=\sum_{l=0}^{m}\binom{p}{s_l}(-1)^{m-l}b^{(s_l)}(s_1,\ldots,s_{l-1}), \tag{3.22}$$

where, by definition, $s_0=0$ and $b^{(s_0)}=b^{(0)}=1$.

To show this, let us substitute the sieve formula (3.21) into the right hand side of (3.22):

$$\begin{aligned}
&\sum_{l=0}^{m}\binom{p}{s_l}(-1)^{m-l}b^{(s_l)}(s_1,\ldots,s_{l-1})\\
=\;&(-1)^m+\sum_{l=1}^{m}\binom{p}{s_l}(-1)^{m-l}\sum_{1\le i_1<\cdots<i_j\le l-1}(-1)^{l-1-j}\binom{s_l}{s_{i_1},s_{i_2},\ldots,s_l-s_{i_j}}\\
=\;&(-1)^m+\sum_{l=1}^{m}\sum_{1\le i_1<\cdots<i_j<l}(-1)^{m-j-1}\binom{p}{s_{i_1},s_{i_2},\ldots,s_l-s_{i_j},p-s_l}\\
=\;&(-1)^m+\sum_{1\le i_1<\cdots<i_j<l\le m}(-1)^{m-(j+1)}\binom{p}{s_{i_1},s_{i_2},\ldots,s_l-s_{i_j},p-s_l}\\
=\;&\sum_{1\le i_1<\cdots<i_r\le m}(-1)^{m-r}\binom{p}{s_{i_1},s_{i_2},\ldots,p-s_{i_r}}=b^{(p)}(s_1,\ldots,s_m).
\end{aligned}$$

This proves (3.22), and so completes the proof. □

Lemma 7 described a recursion that uses all previous rows of Table 1 to compose the elements of a new row. In the next lemma we show a recursion that uses only the previous row and which is an analog of the recursion formula in the ordinary Pascal's triangle: $\binom{p}{k}=\binom{p-1}{k-1}+\binom{p-1}{k}$. Interestingly, an application of this identity is a key step in the following algebraic proof as well. We also give a simple combinatorial proof which basically translates the well-known method by which permutations in S_p can be obtained from permutations in S_{p-1} by adjoining the number p.

Lemma 8. *The following recursion holds for any $p\ge 2$ and multiindex $(j_1,\ldots,j_{p-1})\in\{0,1\}^{p-1}$:*

$$a^{(p)}_{j_1\cdots j_{p-1}}=\sum_{i=1}^{p-1}\delta_i a^{(p-1)}_{j_1^{(i)}\cdots j_{p-2}^{(i)}}=\sum_{(i_1,\ldots,i_{p-2})\in L(j_1,\ldots,j_{p-1})}a^{(p-1)}_{i_1\ldots i_{p-2}}, \tag{3.23}$$

where $a^{(1)}=1$, $\delta_i=|j_i-j_{i-1}|$ for $i\ge 2$, $\delta_1=1$, $j_k^{(i)}=j_k$ for $1\le k\le i-1$, $j_k^{(i)}=j_{k+1}$ for $i\le k\le p-2$, and $L(j_1,\ldots,j_{p-1})$ is the set of all distinct binary sequences obtained from $(j_1,\ldots,j_{p-1})$ by deleting exactly one digit. For example, $a^{(5)}_{0110}=11=a^{(4)}_{110}+a^{(4)}_{010}+a^{(4)}_{011}$.

Proof. First we prove the second equality in (3.23). For this it is enough to show that if the same binary sequence is obtained from $(j_1,\ldots,j_{p-1})$ when eliminating either the kth or the lth digit ($k<l$), then all digits between the kth and lth (including these two) are uniformly either 0's or 1's (a run of 0's or 1's). Therefore, the two recursions given in (3.23) are the same.

Consider a multiindex $(j_1,\ldots,j_{k-1},j_k,\ldots,j_l,j_{l+1},\ldots,j_{p-1})\in\{0,1\}^{p-1}$. Suppose that we get the same binary sequence by deleting j_k and j_l, respectively: $(j_1,\ldots,j_{k-1},j_{k+1},\ldots,j_l,j_{l+1},\ldots,j_{p-1})=(j_1,\ldots,j_{k-1},j_k,\ldots,j_{l-1},j_{l+1},\ldots,j_{p-1})$. Then $j_k=j_{k+1}=\cdots=j_{l-1}=j_l$, so the second equality in (3.20) really holds.

Now it remains to show the first equality in (3.23), that is, the recursion itself.

A combinatorial proof of the recursion. Given a binary sequence $(j_1, \ldots, j_{p-1})$, let us remove a single 1 from a run of 1's or a single 0 from a run of 0's. Count the number of permutations in S_{p-1} determined by the resulting multiindex $(i_1, \ldots, i_{p-2})$. This number is $a^{(p-1)}_{i_1 \ldots i_{p-2}}$ by Theorem 11. We want to show that there is a uniquely determined adjoining of the number p to any such permutation from S_{p-1} to obtain a permutation from S_p corresponding to the original multiindex $(j_1, \ldots, j_{p-1})$.

If a 0 was deleted from a run of 0's, the number p should be inserted right after the number at the position of the first 1 following the affected run of 0's. (If the given run happens to be the last, p is inserted as the last number.) When a 1 was deleted from a run of 1's, the number p should be inserted right after the number at the position of the last 0 preceding the affected run of 1's. (If the given run happens to be the first, p is inserted as the first number.) Since these insertions are the only ones that reconstruct the original descent set, the recursion is proved.

An algebraic proof of the recursion. We are going to proceed by induction over p. Thus suppose that the recursion holds for all multiindices of lengths smaller than $p-1$. First we use the recursion of Lemma 7 for the terms in the right side of (3.23), then we change the order of summation to obtain

$$
\begin{aligned}
\sum_{i=1}^{p-1} \delta_i a^{(p-1)}_{j_1^{(i)} \ldots j_{p-2}^{(i)}} &= \sum_{i=1}^{p-1} \delta_i \sum_{k=0}^{p-2} \binom{p-1}{k} j_k^{(i)} (-1)^{j_{k+1}^{(i)} + \cdots + j_{p-2}^{(i)}} a^{(k)}_{j_1^{(i)} \ldots j_{k-1}^{(i)}} \\
&= \sum_{k=0}^{p-2} \binom{p-1}{k} \Bigg\{ j_{k+1} (-1)^{j_{k+2} + \cdots + j_{p-1}} \sum_{i=1}^{k} \delta_i a^{(k)}_{j_1^{(i)} \ldots j_{k-1}^{(i)}} \\
&\quad + j_k (-1)^{j_{k+1} + \cdots + j_{p-1}} a^{(k)}_{j_1 \ldots j_{k-1}} \sum_{i=k+1}^{p-1} \delta_i (-1)^{j_i} \Bigg\}.
\end{aligned}
$$

Here in the last expression, one may use the induction hypothesis for the first sum. In the second sum observe that $\delta_i (-1)^{j_i}$ is 1 if $j_{i-1} = 1$ and $j_i = 0$, it is -1 if $j_{i-1} = 0$ and $j_i = 1$, and it equals 0 otherwise. Hence we get the identity $j_k \sum_{i=k+1}^{p-1} \delta_i (-1)^{j_i} = j_k (1 - j_{p-1})$. Thus one obtains that

$$
\begin{aligned}
\sum_{i=1}^{p-1} \delta_i a^{(p-1)}_{j_1^{(i)} \ldots j_{p-2}^{(i)}} &= \sum_{k=0}^{p-2} \Bigg\{ \binom{p-1}{k} j_{k+1} (-1)^{j_{k+2} + \cdots + j_{p-1}} a^{(k+1)}_{j_1 \ldots j_k} \\
&\quad + \binom{p-1}{k} (1 - j_{p-1}) j_k (-1)^{j_{k+1} + \cdots + j_{p-1}} a^{(k)}_{j_1 \ldots j_{k-1}} \Bigg\} \\
= \sum_{k=1}^{p-2} &\Bigg\{ \binom{p-1}{k-1} + \binom{p-1}{k} (1 - j_{p-1}) \Bigg\} j_k (-1)^{j_{k+1} + \cdots + j_{p-1}} a^{(k)}_{j_1 \ldots j_{k-1}} \\
&\quad + (1 - j_{p-1}) (-1)^{j_1 + \cdots + j_{p-1}} + \binom{p-1}{p-2} j_{p-1} a^{(p)}_{j_1 \ldots j_{p-2}}.
\end{aligned}
$$

To rewrite the terms above we use recursion (3.20) in the following case:

$$
a^{(p-1)}_{j_1 \ldots j_{p-2}} = \sum_{k=0}^{p-2} \binom{p-1}{k} j_k (-1)^{j_{k+1}} \ldots (-1)^{j_{p-2}} a^{(k)}_{j_1 \ldots j_{k-1}},
$$

plus the identity $-j_{p-1}(-1)^{j_{p-1}} = j_{p-1}$, and the recursion for binomial coefficients:

$$\begin{aligned}\sum_{i=1}^{p-1} \delta_i a^{(p-1)}_{j_1^{(i)}\cdots j_{p-2}^{(i)}} &= \sum_{k=1}^{p-2} \left\{ \binom{p-1}{k-1} + \binom{p-1}{k} \right\} j_k (-1)^{j_{k+1}+\cdots+j_{p-1}} a^{(k)}_{j_1\cdots j_{k-1}} \\ &\quad - j_{p-1}(-1)^{j_{p-1}} \left(a^{(p-1)}_{j_1\cdots j_{p-2}} - (-1)^{j_1+\cdots+j_{p-2}} \right) \\ &\quad + (1-j_{p-1})(-1)^{j_1+\cdots+j_{p-1}} + \binom{p-1}{p-2} j_{p-1} a^{(p)}_{j_1\cdots j_{p-2}} \\ &= \sum_{k=0}^{p-1} \binom{p}{k} j_k (-1)^{j_{k+1}+\cdots+j_{p-1}} a^{(k)}_{j_1\cdots j_{k-1}} = a^{(p)}_{j_1\cdots j_{p-1}}.\end{aligned}$$

This completes the proof. □

The results above imply that Table 1 has properties analogous to the ones of Pascal's triangle: each entry $a^{(p)}_{j_1\cdots j_{p-1}}$ is a positive integer, the first and the last entries, $a^{(p)}_{0\ldots0}$ and $a^{(p)}_{1\ldots1}$ are 1, the table has symmetries $a^{(p)}_{j_1\cdots j_{p-1}} = a^{(p)}_{j_{p-1}\cdots j_1}$ and $a^{(p)}_{j_1\cdots j_{p-1}} = a^{(p)}_{1-j_1\ldots 1-j_{p-1}}$, and the sum of the 2^{p-1} entries in the pth row is $p!$.

3.4 Approximation of the exponential functional of Brownian motion

In this section we are going to show that taking our usual sequence of RWs the resulting sequence of discrete exponential functionals (3.4) converges almost surely to the corresponding exponential functional $\mathcal{I}$ of BM. Based on this, using convergence of moments, we will give an elementary proof of theorem (3.1) of Dufresne and Yor.

We need a more general result about this approximation. This is a generalization of Lemma C originally introduced in [28, Lemma 4]. The proof can be read easily from the proof there. Namely, for almost every ω there exists an $m_0(\omega)$ such that for any $m \geq m_0(\omega)$ and for any $K \geq e$, one has

$$\sup_{j\geq 1} \sup_{0\leq t\leq K} |B_{m+j}(t) - B_m(t)| \leq K^{\frac{1}{4}} (\log K)^{\frac{3}{4}} m 2^{-\frac{m}{2}}. \tag{3.24}$$

Lemma 9. *Let $B_m(t) = 2^{-m}\widetilde{S}_m(t2^{2m})$, $t \geq 0$, $m \geq 0$, be a sequence of shrunken simple symmetric RWs that a.s. converges to BM $(B(t), t \geq 0)$, uniformly on bounded intervals. Then for any $\nu > 0$, as $m \to \infty$,*

$$\begin{aligned} Y_m &= 2^{-2m} \sum_{k=0}^{\infty} \exp\left(2^{-m}\widetilde{S}_m(k) - \nu k 2^{-2m}\right) \\ \to\quad \mathcal{I} &= \int_0^\infty \exp\left(B(t) - \nu t\right) \mathrm{d}t < \infty \qquad a.s. \end{aligned}$$

Proof. The basic idea of the proof is that the sequence of functions $f_m(t,\omega) = \exp(B_m(t) - \nu t)$, converges to $f(t,\omega) = \exp(B(t) - \nu t)$ for $t \in [0,\infty)$ as $m \to \infty$, for almost every ω. If one can find a function $g(t,\omega) \in L^1[0,\infty)$, that dominates each f_m for $m \geq m_0(\omega)$, then their integrals on $[0,\infty)$ also converge to the integral of f, and then we are practically done.

First, by (3.24), for a.e. ω there exists an $m_0 = m_0(\omega)$ so that for any $K \geq e$,

$$\sup_{m\geq m_0} \sup_{0\leq t\leq K} |B_m(t) - B_{m_0}(t)| \leq K^{\frac{1}{4}} (\log K)^{\frac{3}{4}} \leq K^{\frac{1}{2}} \log K, \tag{3.25}$$

where we supposed that m_0 was chosen large enough so that $m_0 2^{-m_0/2} \le 1$.

Second, by the law of iterated logarithms,

$$\limsup_{t\to\infty} \frac{B_{m_0}(t)}{(2t\log\log t)^{\frac{1}{2}}} = \limsup_{u\to\infty} \frac{\widetilde{S}_{m_0}(u)}{(2u\log\log u)^{\frac{1}{2}}} = 1 \qquad \text{a.s.},$$

where $u = t2^{2m_0}$. Hence for a.e. ω, there is a $K_0 = K_0(\nu,\omega)$, such that for any $t \ge K_0$,

$$B_{m_0}(t) \le 2(t\log\log t)^{\frac{1}{2}} \le 2t^{\frac{1}{2}}\log t, \tag{3.26}$$

where K_0 is chosen so large that $3t^{\frac{1}{2}}\log t \le \nu t/2$ for any $t \ge K_0$.

Since a.s. any path of B_{m_0} is continuous, it is bounded on the interval $[0, K_0]$. Then by (3.25), we have an upper bound uniform in m: for any $m \ge m_0$ and $t \in [0, K_0]$, $B_m(t) \le M(\omega)$. On the other hand, when $t > K_0$, by (3.26), $B_{m_0}(t) \le 2t^{\frac{1}{2}}\log t$ and so by (3.25), $B_m(t) \le 3t^{\frac{1}{2}}\log t$, for any $m \ge m_0$.

Summarizing, the function

$$g(t,\omega) = \begin{cases} e^{M(\omega)} & \text{if } 0 \le t \le K_0(\nu,\omega), \\ e^{-\nu t/2} & \text{if } t > K_0(\nu,\omega), \end{cases}$$

is an integrable function on $[0,\infty)$, dominating $\exp(B_m(t) - \nu t)$ for each $m \ge m_0(\omega)$. This implies that

$$\lim_{m\to\infty} \int_0^\infty \exp\left(B_m(t) - \nu t\right)\, dt = \int_0^\infty \exp\left(B(t) - \nu t\right)\, dt < \infty \qquad \text{a.s.}$$

Finally, compare $\int_0^\infty \exp\left(B_m(t) - \nu t\right) dt$ to $Y_m = 2^{-2m}\sum_{k=0}^\infty \exp\left(B_m(k2^{-2m}) \; - \nu k 2^{-2m}\right)$ that appears in the statement of the lemma. Applying the uniform domination of $\exp\left(B_m(t) - \nu t\right)$ by the function g shown above, both the tail of the integral on the interval $[K_0, \infty)$ and the tail of the sum for $k \ge \lceil K_0 2^{2m}\rceil$ is smaller than $\int_{K_0}^\infty \exp(-\nu t/2)\, dt$, thus their difference is uniformly arbitrarily small for any $m \ge m_0$ if K_0 is large enough. On the interval $[0, K_0]$ the difference of the integral and the sum (which is a Riemann sum of a continuous function) tends to zero uniformly as $m \to \infty$, since on each subinterval of length 2^{-2m}, the difference of $B_m(t)$ and $B_m(k2^{-2m})$ is at most 2^{-m}. This completes the proof of the lemma. □

Next we want to apply the results of the previous sections to Y_m. To do this we introduce the following notations. For $m \ge 0$ and $n \ge 1$ let

$$\begin{aligned} Y_{m,n} &= 2^{-2m}\sum_{k=0}^{n} \exp\left(2^{-m}\widetilde{S}_m(k) - \nu k 2^{-2m}\right) \\ &= 2^{-2m}\left(1 + \xi_{m1} + \xi_{m1}\xi_{m2} + \xi_{m1}\cdots\xi_{mn}\right) \end{aligned} \tag{3.27}$$

and $Y_{m,0} = 2^{-2m}$, where $\xi_{mj} = \exp(2^{-m}\widetilde{X}_m(j) - \nu 2^{-2m})$. Here $\widetilde{X}_m(j) = \widetilde{S}_m(j) - \widetilde{S}_m(j-1)$ ($j = 1, 2, \dots$) is an i.i.d. sequence, $\mathbf{P}(\widetilde{X}_m(j) = \pm 1) = \frac{1}{2}$.

Then $Y_{m,n} \nearrow Y_m$ as $n \to \infty$, $Y_m < \infty$ a.s. iff $\nu > 0$, and Y_m satisfies the following self-similarity in distribution:

$$Y_m \overset{d}{=} 2^{-2m} + \xi_m Y_m \qquad \text{or} \qquad 2^{2m}Y_m \overset{d}{=} 1 + \xi_m 2^{2m} Y_m, \tag{3.28}$$

where ξ_m and Y_m are independent, $\xi_m \overset{d}{=} \xi_{mj}$. Using the notations of Section 2, now $\gamma_1 = \exp(-2^{-m} - \nu 2^{-2m})$, $\gamma_2 = \exp(2^{-m} - \nu 2^{-2m})$, $p_1 = p_2 = \frac{1}{2}$. If $\nu > 2^m$, $\gamma_2 < 1$ holds, so the similarity transformations T_1 and T_2 are contractions, mapping the interval $I = [(1-\gamma_1)^{-1}, (1-\gamma_2)^{-1}]$ into itself. By Theorem 10, the distribution of Y_m is singular w.r.t. Lebesgue measure if $\nu > 2^{2m}\log 2$ ($m \ge 1$). Moreover, there is no overlap in the ranges of T_1 and T_2 iff $\nu > 2^{2m}\log(2\cosh(2^{-m}))$. As $m \to \infty$ this means asymptotically that $\nu > \frac{1}{2} + 2^{2m}\log 2 + o(1)$.

For $m \geq 0$ and k integer let

$$\mu_{mk} = \mathbf{E}(\xi_m^k) = \exp(-\nu k 2^{-2m}) \cosh(k 2^{-m}).$$

Since Proposition 3 is applicable to $2^{2m} Y_m$, one obtains that $\mathbf{E}(Y_m^p) < \infty$ if and only if $\mu_{mp} < 1$ and then the following recursion is valid for $p \geq 1$ integer:

$$\mathbf{E}(Y_m^p) = \frac{1}{1-\mu_{mp}} \sum_{k=0}^{p-1} \binom{p}{k} 2^{-2m(p-k)} \mu_{mk} \mathbf{E}(Y_m^k). \tag{3.29}$$

Now using $\cosh(x) < e^x$ $(x > 0)$, it follows that $\mu_{mk} < \exp\left(k 2^{-m}(1 - \nu 2^{-m})\right)$. So $\mu_{mk} < 1$ for any $k \geq 1$ if $\nu \geq 2^m$. If $0 < \nu < 2^m$, only finitely many positive moments are finite, since $\mu_{mk} \geq 1$ when $0 < \nu < 2^{2m} k^{-1} \log\cosh(k 2^{-m}) \to 2^m$ as $k \to \infty$.

More importantly,

$$\mu_{mk} < \exp\left(k 2^{-2m} \left(\frac{k}{2} - \nu\right)\right) \qquad (k \geq 1), \tag{3.30}$$

since $\cosh(x) < \exp(x^2/2)$ when $x > 0$ (compare the Taylor series). Thus $\mu_{mk} < 1$ for any $m \geq 0$ if $\nu \geq \frac{k}{2}$. This condition is sharp as $m \to \infty$. For, apply $e^x = 1 + x + o(x)$ and $\cosh(x) = 1 + x^2/2 + o(x^2)$ (as $x \to 0$) to the definition of μ_{mk}. Then

$$\mu_{mk} = 1 + k 2^{-2m} \left(\frac{k}{2} - \nu\right) + o(2^{-2m}), \tag{3.31}$$

for any fixed k as $m \to \infty$.

Using these results one can evaluate the moments of the exponential functional.

3.4.1 The moments of the exponential functional of BM

Lemma 10. *If p is a positive integer such that $\frac{p}{2} < \nu$, then*

$$\lim_{m\to\infty} \mathbf{E}(Y_m^p) = \frac{1}{\prod_{k=1}^{p} \left(\nu - \frac{k}{2}\right)} < \infty. \tag{3.32}$$

Proof. By (3.30), for any positive integer p such that $\frac{p}{2} < \nu$ we have $\mu_{mp} < 1$. Since Proposition 3 is valid for $2^{2m} Y_m$, the recursion formula (3.29) holds, and by induction one gets $\mathbf{E}(Y_m^p)$ as a rational function of the moments $\mu_{m1}, \ldots, \mu_{mp}$, similarly to formula (3.18). The argument below formula (3.18) also applies here too, showing that the sum of the coefficients in the numerator of this rational function is $p! 2^{-2mp}$. The extra factor comes from the difference that Y_m is multiplied by 2^{2m} here, compare equations (3.17) and (3.29). Since each $\mu_{mk} \to 1$ as $m \to \infty$, it follows that 2^{2mp} times the numerator tends to $p!$.

By (3.31), we get that $1 - \mu_{mk} = k 2^{-2m}(\nu - \frac{k}{2}) + o(2^{-2m})$ if k is fixed and $m \to \infty$. So 2^{2mp} times the denominator of the rational function tends to $p! \prod_{k=1}^{p}(\nu - \frac{k}{2})$ as $m \to \infty$. This and the limit of the numerator together imply the statement of the lemma. □

Our next objective is to give an asymptotic formula, similar to (3.32), for the negative moments of Y_m as $m \to \infty$.

Lemma 11. *For all integer $p \geq 1$, we have*

$$\lim_{m\to\infty} \mathbf{E}(Y_m^{-p}) = \lim_{m\to\infty} \mathbf{E}(Y_m^{-1}) \prod_{k=1}^{p-1} \left(\nu + \frac{k}{2}\right), \tag{3.33}$$

where $\lim_{m\to\infty} \mathbf{E}(Y_m^{-1}) < \infty$.

Proof. We want to show (3.33) by establishing a recursion. Introduce the notations $z_{m,k} = \mathbf{E}(Y_m^{-k})$ and $\mu_{m,-k} = \mathbf{E}(\xi_m^{-k})$ for $k \geq 1$ integer. By (3.27), $0 < Y_m^{-1} < 2^{2m}$, hence all negative moments $z_{m,k}$ of Y_m are finite.

The self-similarity equation (3.28) implies that $\xi_m Y_m \stackrel{\mathsf{d}}{=} Y_m - 2^{-2m}$ and so

$$\xi_m^{-1} Y_m^{-1} \stackrel{\mathsf{d}}{=} \frac{Y_m^{-1}}{1 - 2^{-2m} Y_m^{-1}}, \tag{3.34}$$

where ξ_m and Y_m are independent. Taking kth moment ($k \geq 1$ integer) on both sides and applying the Taylor series

$$\frac{x^k}{(1-x)^k} = \sum_{n=k}^{\infty} \binom{n-1}{k-1} x^n, \tag{3.35}$$

valid for any $|x| < 1$, one obtains

$$\mu_{m,-k}\, z_{m,k} = \sum_{n=k}^{\infty} \binom{n-1}{k-1} 2^{-2m(n-k)} z_{m,n},$$

with the notations introduced above. This implies that

$$(\mu_{m,-k} - 1) z_{m,k} - k 2^{-2m} z_{m,k+1} = a(m,k), \tag{3.36}$$

where

$$a(m,k) = \sum_{n=k+2}^{\infty} \binom{n-1}{k-1} 2^{-2m(n-k)} z_{m,n} \geq 0.$$

Next we want to give an upper bound for $a(m,k)$, which goes to zero fast enough as $m \to \infty$. Since $\xi_m \geq \gamma_1 = \exp(-2^{-m} - \nu 2^{-2m})$, by (3.27) it follows that

$$Y_m^{-1} \leq 2^{2m} \left(\sum_{j=0}^{\infty} \gamma_1^j \right)^{-1} = 2^{2m}(1 - \gamma_1) \leq 2^{2m}(2^{-m} + \nu 2^{-2m}) \leq 2^{m+1}, \tag{3.37}$$

if $m \geq \log(\nu)/\log(2)$, where we used that $1 - e^{-x} \leq x$, for any real x. This implies that $z_{m,r+j} = \mathbf{E}(Y_m^{-r-j}) \leq \mathbf{E}(Y_m^{-r}) \left(\sup(Y_m^{-1})\right)^j \leq z_{m,r} 2^{(m+1)j}$ for $r, j \geq 0$. Substituting this into the definition of $a(m,k)$, one gets that

$$\begin{aligned}
a(m,k) &\leq z_{m,k+1} \sum_{n=k+2}^{\infty} \binom{n-1}{k-1} 2^{-2m(n-k)} 2^{(m+1)(n-k-1)} \\
&= z_{m,k+1} 2^{k(m-1)-m-1} \sum_{n=k+2}^{\infty} \binom{n-1}{k-1} 2^{(1-m)n} \\
&= z_{m,k+1} 2^{-m-1} \left((1 - 2^{1-m})^{-k} - 1 - k 2^{1-m} \right) \\
&\leq z_{m,k+1} 2^{-m-1} 4k(k+1) 2^{-2m} = z_{m,k+1} 2k(k+1) 2^{-3m},
\end{aligned}$$

if m is large enough, depending on k.

Let us substitute this estimate of $a(m,k)$ into (3.36) and express the following ratio:

$$\frac{z_{m,k+1}}{z_{m,k}} = \frac{\mu_{m,-k} - 1}{k 2^{-2m}(1 + O(2^{-m}))}.$$

Apply the asymptotics (3.31) here with $-k$:

$$\frac{z_{m,k+1}}{z_{m,k}} = \frac{\nu + \frac{k}{2} + o(1)}{1 + O(2^{-m})},$$

as $m \to \infty$. This implies the equality $\lim_{m\to\infty} z_{m,k+1}/z_{m,k} = \nu + \frac{k}{2}$, and thus for any positive integer p,

$$\lim_{m\to\infty} \frac{\mathbf{E}(Y_m^{-p})}{\mathbf{E}(Y_m^{-1})} = \prod_{k=1}^{p-1} \left(\nu + \frac{k}{2}\right). \tag{3.38}$$

It remains to show that $\mathbf{E}(Y_m^{-1})$ has a finite limit as $m \to \infty$. Writing $Y_m^{-2} = Y_m Y_m^{-3}$ and applying the Cauchy-Schwarz inequality, one obtains $\left(\mathbf{E}(Y_m^{-2})\right)^2 \le \mathbf{E}(Y_m^2)\mathbf{E}(Y_m^{-6})$, or $\mathbf{E}(Y_m^{-2}) \le \mathbf{E}(Y_m^2)\mathbf{E}(Y_m^{-6})/\mathbf{E}(Y_m^{-2})$. Suppose first that $\nu > 1$ and take limit here on the right hand side as $m \to \infty$, applying (3.32) and (3.38). It follows that $\sup_{m\ge1} \mathbf{E}(Y_m^{-2}) \le \infty$. As $\mathbf{E}(Y_m^{-2})$ is an increasing function of ν by its definition, hence the same is true for any $\nu \in (0, 1]$ as well. Since by Lemma 9, $Y_m \to \mathcal{I}$ a.s., where each Y_m and also $\mathcal{I}$ take values in $(0,\infty)$ a.s., it follows that $Y_m^{-1} \to I^{-1}$ a.s. Then by the L^2 uniform boundedness of Y_m^{-1} $(m \ge 0)$ shown above, $\mathbf{E}(Y_m^{-1}) \to \mathbf{E}(I^{-1}) < \infty$ follows as well. This ends the proof of the lemma. □

Finally, it turns out that Y_m^{-1} converges to $\mathcal{I}^{-1}$ in any L^p. This makes it possible to recover the result (3.1) of [7] and [38].

Theorem 12. *Let $B_m(t) = 2^{-m}\widetilde{S}_m(t2^{2m})$, $t \ge 0$, $m \ge 0$, be a sequence of shrunken simple symmetric RWs that a.s. converges to BM $(B(t), t \ge 0)$, uniformly on bounded intervals. Take*

$$Y_m = 2^{-2m} \sum_{k=0}^{\infty} \exp\left(B_m(k2^{-2m}) - \nu k 2^{-2m}\right) = 2^{-2m} \left(1 + \xi_{m1} + \xi_{m1}\xi_{m2} + \cdots\right)$$

and

$$\mathcal{I} = \int_0^\infty \exp\left(B(t) - \nu t\right)\, \mathrm{d}t$$

when $\nu > 0$. Then the following statements hold true:

(a) Y_m^{-1} converges to $\mathcal{I}^{-1}$ in L^p for any $p \ge 1$ real and $\lim_{m\to\infty} \mathbf{E}(Y_m^{-p}) = \mathbf{E}(\mathcal{I}^{-p}) < \infty$;

(b) $\mathcal{I} \stackrel{\mathrm{d}}{=} 2/Z_{2\nu}$, where $Z_{2\nu}$ is a gamma distributed random variable with index 2ν and parameter 1;

(c) Y_m converges to $\mathcal{I}$ in L^p for any integer p such that $1 \le p < 2\nu$ (supposing $\nu > \frac{1}{2}$) and then $\lim_{m\to\infty} \mathbf{E}(Y_m^p) = \mathbf{E}(\mathcal{I}^p) < \infty$. The same is true for any real q, $1 \le q < p$.

Proof. By Lemma 9, $Y_m \to \mathcal{I}$ a.s., where each Y_m and also $\mathcal{I}$ take values in $(0,\infty)$ a.s. Hence $Y_m^{-1} \to \mathcal{I}^{-1}$ a.s. By Lemma 11, $\lim_{m\to\infty} \mathbf{E}(Y_m^{-k}) < \infty$ for any $k \ge 1$ integer, so (a) follows.

Thus by (a) and Lemma 11, for any integer $p \ge 1$,

$$a_p = \mathbf{E}(\mathcal{I}^{-p}) = c \prod_{k=1}^{p-1} \left(\nu + \frac{k}{2}\right) = c2^{1-p}(2\nu + 1) \cdots (2\nu + p - 1) < \infty,$$

where $c = \mathbf{E}(\mathcal{I}^{-1})$. By a classical result, see [25], a Stieltjes moment problem is determinate, that is the moments uniquely determine a probability distribution on $[0, \infty)$, if there exist constants $C > 0$ and $R > 0$ such that $a_p \le CR^p(2p)!$ for any $p \ge 1$ integer. In the present case $a_p \le c2^{-p}(p + 1)!$ when $\nu \le 1$ and $a_p \le c(\nu/2)^{p-1}(p + 1)!$ when $\nu > 1$, so the moment problem for $\mathcal{I}^{-1}$ is determinate and it also follows that $\mathcal{I}^{-1}$ has a finite moment generating function in a neighborhood of the origin.

Also, using the moments of the gamma distribution we get

$$b_p = \mathbf{E}(2^{-p} Z_{2\nu}^p) = 2^{-p} \frac{\Gamma(2\nu + p)}{\Gamma(2\nu)} = 2^{-p}(2\nu)(2\nu + 1) \cdots (2\nu + p - 1),$$

for any $p \geq 1$, and $Z_{2\nu}/2$ has a finite moment generating function in a neighborhood of the origin as well. Writing down the two moment generating functions by the help of the moments a_p and b_p, respectively, it follows that

$$\mathbf{E}(\exp(u\mathcal{I}^{-1})) = \frac{c}{\nu} \mathbf{E}(\exp(uZ_{2\nu}/2))$$

in a neighborhood of the origin. Substituting $u = 0$, one obtains $c = \mathbf{E}(\mathcal{I}^{-1}) = \nu$ and this proves (b).

Finally, again, $Y_m \to \mathcal{I}$ a.s. by Lemma 9. If p is an integer such that $1 \leq p < 2\nu$, by Lemma 10, using the moments of the gamma distribution, and by (b), we have $\lim_{m\to\infty} \mathbf{E}(Y_m^p) = \mathbf{E}(2^p Z_{2\nu}^{-p}) = \mathbf{E}(\mathcal{I}^p)$. This proves (c). □

3.4.2 Properties of the exponential functional process

Theorem 12 says that the one dimensional marginal distributions of the process $X(\nu) = \frac{1}{\mathcal{I}(\nu)}$, $0 < \nu$ are gamma distributed with index 2ν as that of a gamma process. In spite of the equality of the one dimensional marginals X is completely different from the gamma process as the next proposition shows.

Proposition 4. *The process*

$$X(\nu) := \frac{1}{\int_0^\infty \mathrm{e}^{B_s - \nu s}\, \mathrm{d}s} \tag{3.39}$$

is almost surely continuous so it is not a gamma process.

Proof. Clearly, it is enough to prove that $\mathcal{I}(\cdot)$ is continuous. Fix ν_0 and let $\nu < \nu_0$ so we have

$$I(\nu) - I(\nu_0) = \int_0^\infty \mathrm{e}^{B_s - \nu s} \left(1 - \mathrm{e}^{(\nu - \nu_0)s}\right) \mathrm{d}s$$

Letting ν tends to ν_0 by the Lebesgue dominated convergence theorem one gets that the integral on the right tends to 0. So $I(\nu) \to I(\nu_0)$ almost surely as $\nu \to \nu_0$ if $\nu < \nu_0$. The proof of the case $\nu > \nu_0$ is likewise. Therefore, we get that $\mathcal{I}$ and so X are almost surely continuous. □

An interesting consequence of the technique used in the proof of Lemma 11 is the following lemma. It says that every multi indices moment of the process $(\mathcal{I}^{-1}(\nu))_{\nu>0}$ can be derived from some other multi indices moments with bigger individual moments. We remark that this statement seems to be insufficient to determine all multi indices moments.

For simplicity, we introduce some notations.
Fix a positive integer d and vectors $\underline{k} \in {\mathbb{Z}_+}^d$, $\underline{\nu} \in \mathbb{R}_+^d$. Let

$$M((\underline{\nu}, \underline{k})) = \mathbf{E}\left(\mathcal{I}(\nu_1)^{-k_1} \ldots \mathcal{I}(\nu_d)^{-k_d}\right).$$

Finally, let $(\underline{\nu}, \underline{k})_i = (\underline{\nu}, (k_1, \ldots, k_{i-1}, k_i + 1, k_{i+1}, \ldots k_d))$.

For the multi indices moment $M((\underline{\nu}, \underline{k}))$ we have the following

Lemma 12.

$$M((\underline{\nu}, \underline{k})) = \left(\frac{k_1 + \cdots + k_d}{2} + \frac{k_1\nu_1 + \cdots + k_d\nu_d}{k_1 + \cdots + k_d}\right)^{-1} \sum_{i=1}^d \frac{k_i}{k_1 + \cdots + k_d} M((\underline{\nu}, \underline{k})_i).$$

Proof. Basically, we use the technique introduced in the proof of Lemma 11 from (3.34) till (3.38).

By using (3.35), we can determine the multidimensional version of (3.34) for higher moments

$$\xi_m^{-k_1}(\nu_1)Y_m^{-k_1}(\nu_1)\cdots\xi_m^{-k_d}(\nu_d)Y_m^{-k_d}(\nu_d) \stackrel{\mathsf{d}}{=} \prod_{i=1}^d \left(\sum_{n_i=k_i}^\infty \binom{n_i-1}{k_i-1} Y_m^{-n_i}(\nu_i) 2^{-2m(n_i-k_i)} \right)$$

$$= \sum_{n_1=k_1}^\infty \cdots \sum_{n_d=k_d}^\infty \left(\binom{n_1-1}{k_1-1} \cdots \binom{n_d-1}{k_d-1} \cdot Y_m^{-n_1}(\nu_1)\cdots Y_m^{-n_d}(\nu_d)\cdot 2^{-2m\sum_{i=1}^d (n_i-k_i)} \right) \quad (3.40)$$

Let $\mu_m((\underline{\nu},\underline{k})) = \mathbf{E}\, \xi_m^{-k_1}(\nu_1)\cdots\xi_m^{-k_d}(\nu_d)$ and $z_m((\underline{\nu},\underline{n})) = \mathbf{E}\, Y_m^{-n_1}(\nu_1)\cdots Y_m^{-n_d}(\nu_d)$. Using the definition of $\xi_m(\nu)$ and the following two equations

$$\cosh\left((k_1+\cdots+k_d)\frac{1}{2^m}\right) = 1 + \frac{(k_1+\cdots+k_d)^2}{2}\frac{1}{2^m} + o\left(\frac{1}{2^m}\right),$$

$$e^{(\nu_1 k_1+\ldots\nu_d k_d)\frac{1}{2^{2m}}} = 1 + (\nu_1 k_1 + \ldots \nu_d k_d)\frac{1}{2^{2m}} + o\left(\frac{1}{2^m}\right)$$

simple calculation shows that

$$\mu_m((\underline{\nu},\underline{k})) = \cosh\left((k_1+\cdots+k_d)\frac{1}{2^m}\right) e^{(\nu_1 k_1+\ldots\nu_d k_d)\frac{1}{2^{2m}}}$$

$$= 1 + (k_1+\cdots+k_d)2^{-2m}\left(\frac{k_1+\cdots+k_d}{2} + \frac{k_1\nu_1+\cdots+k_d\nu_d}{k_1+\cdots+k_d}\right) + o(2^{-2m}).$$

Now, taking expectation in both sides of (3.40) then using the independence of $\xi_m(\nu)$ and $Y_m(\nu)$ we get

$$\mu_m((\underline{\nu},\underline{k}))z_m((\underline{\nu},\underline{k})) = z_m((\underline{\nu},\underline{k})) + \sum_{i=1}^d k_i 2^{-2m} z_m((\underline{\nu},\underline{k})_i) + a_m((\underline{\nu},\underline{k})) \quad (3.41)$$

where

$$a_m((\underline{\nu},\underline{k})) = \mathbf{E}\left(\sum_{\upharpoonleft \underline{n}-\underline{k}\upharpoonright \geq 2}\!\!\cdots\sum \binom{n_1-1}{k_1-1}\cdots\binom{n_d-1}{k_d-1}\cdot Y_m^{-n_1}(\nu_1)\cdots Y_m^{-n_d}(\nu_d)\cdot 2^{-2m\sum_{i=1}^d(n_i-k_i)}\right)$$

where $\upharpoonleft \underline{n}-\underline{k} \upharpoonright = \sum_{i=1}^d (n_i-k_i)$, that is, we sum those terms whose vector $\underline{n}$ is either bigger than $\underline{k}$ in at least two entries or the difference is bigger than 2 in at least one entry. Using the estimate (3.37), the remark and the estimation afterwards one can obtain

$$a_m((\underline{\nu},\underline{k})) = \sum_{i=1}^d z_m((\underline{\nu},\underline{k})_i)$$

$$\cdot\left((1-2^{1-m})^{k_1+\cdots+k_d-k_i} 2k(k+1)2^{-3m} + \sum_{j=1}^d (1-2^{1-m})^{k_1+\cdots+k_d-k_i-k_j} 2^{2-2m} k_j 2^{1-m} \right)$$

$$\leq C_{\underline{k}}\, 2^{-3m} \sum_{i=1}^d z_m((\underline{\nu},\underline{k})_i)$$

for appropriate constant $C_{\underline{k}}$. From (3.41), we have

$$z_m((\underline{\nu},\underline{k})) = \frac{k_1 2^{-2m}}{1-\mu_m((\underline{\nu},\underline{k}))} z_m((\underline{\nu},\underline{k})_1) + \cdots + \frac{k_1 2^{-2m}}{1-\mu_m((\underline{\nu},\underline{k}))} z_m((\underline{\nu},\underline{k})_1) + O(2^{-m}).$$

Applying the convergence results of Theorem 12, that is, $z_m((\underline{\nu},\underline{k})) \to M((\underline{\nu},\underline{k}))$ as m tends to infinity, we get

$$M((\underline{\nu},\underline{k})) = \left(\frac{k_1+\cdots+k_d}{2} + \frac{k_1\nu_1+\cdots+k_d\nu_d}{k_1+\cdots+k_d}\right)^{-1} \sum_{i=1}^{d} \frac{k_i}{k_1+\cdots+k_d} M((\underline{\nu},\underline{k})_i)$$

as it was required.

□

Bibliography

[1] Carmona, P., Petit, F. and Yor, M. (1997) On the distribution and asymptotic results for exponential functionals of Lévy processes. In: Yor, M., ed. *Exponential functionals and principal values related to Brownian motion*, pp. 73-121. Biblioteca de la Revista Matemática Iberoamericana.

[2] Csörgő, M. (1999) Random walking around financial mathematics. In: Révész, P. and Tóth, B., eds. *Random walks, Bolyai Society Mathematical Studies, 9*, pp. 59-111, J. Bolyai Math. Soc., Budapest.

[3] Dambis, K.E. (1965) On the decomposition of continuous martingales. *Theor. Prob. Appl.* **10**, 401-410.

[4] Dubins, L. E.; Émery, Michel and Yor, M. (1993) On the Lévy transformation of Brownian motions and continuous martingales. In: *Séminaire de Probabilités, XXVII, Lecture Notes in Math., 1557*, pp. 122–132, Springer, Berlin.

[5] Dubins, L. and Schwarz, G. (1965) On continuous martingales. *Proc. Nat. Acad. Sci. USA* **53**, 913-916.

[6] Dubins, L. E. and Smorodinsky, M. (1992) The Modified, Discrete, Lévy-Transformation is Bernoulli. In: *Séminaire de Probabilités, XXVI, Lecture Notes in Math., 1526*, Springer.

[7] Dufresne, D. (1990) The distribution of a perpetuity, with applications to risk theory and pension funding. *Scand. Actuarial J.*, 39-79.

[8] Falconer, K. (1990) *Fractal geometry. Mathematical foundations and applications.* Wiley, Chichester.

[9] Grincevičius, A.K. (1974) On the continuity of the distribution of a sum of dependent variables connected with independent walks on lines. *Theory Prob. Appl.* **19**, 163-168.

[10] Hoeffding, W. (1963) Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13-30.

[11] Jacod, J. and Shiryaev, A. N. (1987) *Limit Theorems for Stochastic Processes*, Springer-Verlag, Berlin

[12] Karandikar, R.L. (1983) On the quadratic variation process of continuous martingales. *Illinois J. Math.* **27**, 178-181.

[13] Karandikar, L. Rajeeva (1995) On pathwise stochastic integration. *SPA* **57**, 11-18.

[14] Kiefer, J. (1969) On the deviation in the Skorokhod-Strassen approximation scheme. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **13** 321-332.

[15] Knight, F.B. (1962) On the random walk and Brownian motion. *Trans. Amer. Math. Soc.* **103**, 218-228.

[16] Knuth, D.E. (1973) The Art of Computer Programming. Vol. 3. Sorting and searching. Addison-Wesley, Reading, Mass.

[17] R. S. Lipster and A. N. Shiryaev (1977) *Statistic of Random Process* Vol I: General Theory, Springer-Verlag, New York.

[18] MacMahon, P.A. (1915) Combinatory Analysis. Vol. 1. Cambridge Univ. Press, Cambridge, England.

[19] Marchal, P. (2003) Constructing a sequence of random walks strongly converging to Brownian motion, Preprint

[20] Ocone, D.L. (1993) A symmetry characterization of conditionally independent increment martingales. In: D. Nualart and M. Sanz-Solé, eds. *Barcelona Seminar on Stochastic Analysis, Progress in Probability, 32*, pp. 147-167, Birkhäuser, Basel.

[21] Pitman, J. (2003) Combinatorial Stochastic Processes, Preprint

[22] Révész, P. (1990) *Random Walk in Random and Non-Random Environments.* World Scientific, Singapore.

[23] Revuz, D. and Yor, M. (1999) *Continuous Martingales and Brownian Motion.* Third edition, Springer, Berlin.

[24] Rudin, W. (1970) *Real and complex analysis.* McGraw-Hill, New York.

[25] Simon, B. (1998) The classical moment problem as a self-adjoint finite difference operator. *Adv. Math.* **137**, 82-203.

[26] Simon, K., Solomyak, B., and Urbański, M. (2001) Invariant measures for parabolic IFS with overlaps and random continued fractions. *Trans. Amer. Math. Soc.* **353**, 5145-5164.

[27] Stanley, R. (1986) Enumerative Combinatorics. Vol. 1. Wadsworth and Brooks/Cole Mathematics Series, Monterey, Calif.

[28] Szabados, T. (1996) An elementary introduction to the Wiener process and stochastic integrals. *Studia Sci. Math. Hung.* **31**, 249-297.

[29] Szabados, T. (2001) Strong approximation of fractional Brownian motion by moving averages of simple random walks. *Stochastic Process. Appl.* **92**, 31-60.

[30] Szabados, T. and Székely, B. (2003) An exponential functional of random walks. *J. Appl. Prob.* **40**, 413-426.

[31] Szabados, T. and Székely, B. (2004) Moments of an exponential functional of random walks and permutations with given descent sets, *Periodicam Math. Hung.*, to appear

[32] Székely, B. and Szabados, T (2004) Strong approximation of continuous local martingales by simple random walks, *Studia Sci. Math. Hung.* **41**, 101-126.

[33] Székely, G.J. (1975) On the polynomials of independent random variables. In: *Limit theorems of probability theory (Keszthely, 1974), Colloq. Math. Soc. János Bolyai, 11*, pp. 365-371, North-Holland, Amsterdam.

[34] Van Zanten, Herry (2002) Continuous Ocone martingales as weak limits of rescaled martingales. *Elect. Comm. in Probab.* **7**, 205-212.

[35] Vervaat, W. (1979) On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Prob.* **11**, 750-783.

[36] Vostrikova, L. and Yor, M. (2000) Some invariance properties (of the laws) of Ocone's martingales. In: *Séminaire de Probabilités, XXXIV, Lecture Notes in Math., 1729*, pp. 417–431, Springer, Berlin.

[37] P. Walters (1982) *An Introduction to Ergodic Theory.* Spriger-Verlag, New York.

[38] Yor, M. (1992) On certain exponential functionals of real-valued Brownian motion. *J. Appl. Prob.*, **29**, 202-208.

[39] Yor, M. (2001) *Exponential functionals of Brownian motion and related processes.* Springer, Berlin.

[40] Zabrocki, M. (2001) Integer sequence A060351. *On-line encyclopedia of integer sequences.* `http://www.research.att.com/~njas/sequences/Seis.html`

DECLARATION

Alulírott Székely Balázs kijelentem, hogy ezt a doktori értekezést magam készítettem és abban csak a megadott forrásokat használtam fel. Minden olyan részt, amelyet szó szerint, vagy azonos tartalomban, de átfogalmazva más forrásból átvettem, egyértelműen, a forrás megadásával megjelöltem.

Budapest, 2004. április 29.

Székely Balázs