Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 92, 1–22.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP818

Limit theorems and ergodicity for general bootstrap random walks*

Andrea Collevecchio[†] Kais Hamza[‡] Meng Shi[§] Ruth J. Williams[¶]

Abstract

Given the increments of a simple symmetric random walk $(X_n)_{n\geq 0}$, we characterize all possible ways of recycling these increments into a simple symmetric random walk $(Y_n)_{n\geq 0}$ adapted to the filtration of $(X_n)_{n\geq 0}$. We study the long term behavior of a suitably normalized two-dimensional process $((X_n, Y_n))_{n\geq 0}$. In particular, we provide necessary and sufficient conditions for the process to converge to a two-dimensional Brownian motion (possibly degenerate). We also discuss cases in which the limit is not Gaussian. Finally, we provide a simple necessary and sufficient condition for the ergodicity of the recycling transformation, thus generalizing results from Dubins and Smorodinsky (1992) and Fujita (2008), and solving the discrete version of the open problem of the ergodicity of the general Lévy transformation (see Mansuy and Yor, 2006).

Keywords: functional limit theorems; ergodicity; Lévy transformation; random walks; long memory.
MSC2020 subject classifications: 60G50; 60F17; 28D05.
Submitted to EJP on July 1, 2021, final version accepted on June 30, 2022.

1 Introduction

Simple symmetric random walks, stochastic integrals and measure preserving transformations are ubiquitous to the theory of probability as well as to a wide range of applications. This paper investigates the properties of the most general discrete-time setting that combines all three; that is the most general measure-preserving stochastic integral of a simple symmetric random walk.

More specifically, consider a one-dimensional simple symmetric random walk $(X_n)_n$, with $X_0 = 0$. Let $\xi_n = X_n - X_{n-1}$ be the independent increments of the random walk, with

*This research was supported by the Australian Research Council Grant DP180100613. RJ Williams gratefully acknowledges the support of NSF grant DMS-1712974.

[†]Monash University, Australia. E-mail: andrea.collevecchio@monash.edu
[‡]Monash University, Australia. E-mail: kais.hamza@monash.edu
[§]Monash University, Australia. E-mail: meng.shi@monash.edu
[¶]Mathematics Department, University of California, San Diego. E-mail: rjwilliams@ucsd.edu

common distribution $\mathbb{P}(\xi_1 = -1) = \mathbb{P}(\xi_1 = +1) = 1/2$. We study the measure-preserving, non-anticipative (i.e. adapted) bootstrapping (i.e. recycling) of $(\xi_n)_{n\geq 1}$ and obtain a complete description of all functions ϕ_n such that the sequence $\eta_n = \phi_n(\xi_1, \ldots, \xi_n)$ replicates the law of the original sequence: $(\eta_n)_{n\geq 1} \stackrel{d}{=} (\xi_n)_{n\geq 1}$. Such a sequence defines a new simple symmetric random walk

$$Y_n = \sum_{k=1}^{n} \eta_k,\ n \geq 1,\ \text{and } Y_0 = 0.$$

Equipped with such a representation, we study the limiting behaviour of a suitably normalized pair (X_n, Y_n) as well as the ergodicity of the recycling transformation.

Seen from the point of view of the random walks themselves, rather than their increments, we aim to study the long term behaviour of (X_n, Y_n), where $(Y_n)_{n\geq 0}$ is also a simple symmetric random walk adapted to the natural filtration of $(X_n)_{n\geq 0}$. We call such a two-dimensional process General Bootstrap Random Walk (GBRW). Properly normalized, each of $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ converges to a standard Brownian motion. Furthermore, the classical theory (see for example Theorem VIII.3.11 of [10]) describes the long term behaviour of the normalized pair $(X_{\lfloor nt \rfloor}/\sqrt{n}, Y_{\lfloor nt \rfloor}/\sqrt{n})$ in terms of the asymptotic behaviour of its quadratic variation process. In this paper, using a representation property (see Section 2), we characterise those cases that lead to a Gaussian limit. For example, despite the strong dependence that exists between the pre-limit processes, it is shown in [3] that when $\eta_n = \prod_{k=1}^{n} \xi_k$, the limit (in the sense of the weak topology) is a two-dimensional Brownian motion with independent components. In this paper we seek to characterise all processes (X_n, Y_n) for which this is true. We also ask whether other limits are possible and in particular whether and under what circumstances a correlated two-dimensional Brownian motion can be obtained as a limiting process. More generally, we attempt to describe long term behaviours that result from combining two strongly dependent simple symmetric random walks. Of particular interest will be the case $\eta_n = \text{sgn}(X_{n-1})\xi_n$ and other similar settings.

By the martingale representation property (see for example 15.1 of [18]),

$$\eta_n = Y_n - Y_{n-1} = H_{n-1}(X_n - X_{n-1}) = H_{n-1}\xi_n,$$

for some adapted process $(H_n)_{n\geq 0}$. This, combined with the fact that $(Y_n)_{n\geq 0}$ is a simple random walk, or equivalently that $\eta_n \in \{-1, 1\}$, enables the characterisation and parametrisation of the processes $(H_n)_{n\geq 0}$ – see Section 2. In Section 3, we give necessary and sufficient conditions for a suitably normalized (X_n, Y_n) to converge weakly to a two-dimensional Brownian motion (possibly degenerate). Two classes of GBRW are then studied. The Extended Bootstrap Random Walk focuses on the case $\eta_n = \left(\prod_{k=1}^{\lfloor R(n) \rfloor} \xi_k\right) \xi_n$, a setting that generalises that of [3], and provides sufficient conditions on the asymptotic behaviour of the real function R to achieve convergence to a pair of independent Brownian motions. The next setting demonstrates how convergence to a correlated Brownian motion can be achieved. Section 4 explores the non-Gaussian case and more specifically that of $\eta_n = f(X_{n-1}/\sqrt{n})\xi_n$, of which $\eta_n = \text{sgn}(X_{n-1})\xi_n$ is one example, and shows that a non-Gaussian limit results.

Besides describing the asymptotic behaviour of the pair (X_n, Y_n), the paper examines the ergodicity of the recycling transformation and provides a simple necessary and sufficient condition for the latter to hold. This is detailed in Section 5.

Literature review

The case $\eta_n = \prod_{k=1}^{n} \xi_k$ has been the object of a number of investigations in a variety of contexts. It was referred to as the Bootstrap Random Walk (BRW) and described in great details in [3] (see also [4]). In mathematical finance, it was used to discuss the continuity of utility maximization under weak convergence (see [1]). Within the context of noise sensitivity, [15] compares the effects on the sequences $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ of "Poisson switches" in the sequence $(\xi_n)_{n\geq 1}$. Owing to the fact that every switch in the sequence $(\xi_n)_{n\geq 1}$ results in multiple concurrent switches in the sequence $(\eta_n)_{n\geq 1}$, it is shown in [15] that the noise sensitivity of $(Y_n)_{n\geq 0}$ is greater than that of $(X_n)_{n\geq 0}$.

GBRW bear some relationship to the celebrated elephant random walk introduced by Schütz and Trimper [17]. This is a model for long memory within a random walk setting. In it, the cumulative effect of BRW $\eta_n = \eta_{n-1}\xi_n$ is replaced with $\eta_n = \eta^*\xi_n$, where η^* is selected at random from $\{\eta_1, \dots, \eta_{n-1}\}$. The process is shown in [17] to undergo a phase transition at the critical value $p_c = 1/2$, from a weakly localized regime to an escape regime. We refer the reader to [9], [11] and [2], and the reference lists within.

Finally, [6] and [7] investigate a time-dependent biased bootstrap random walk and show that as the "turning", that is $\mathbb{P}(\xi_n = -1)$, gets slower, the Central Limit Theorem and then the Law of Large Numbers break down.

The case $\eta_n = \mathrm{sgn}(X_{n-1})\xi_n$ is a discrete version of the celebrated Lévy transformation

$$B_\bullet \to \int_0^\bullet \mathrm{sgn}(B_s)dB_s,$$

where B is a Brownian motion. The ergodicity of the latter is to this date an open problem (see for example [16]). The ergodicity of more general, Lévy-type transformations, $B_\bullet \to \int_0^\bullet H_s dB_s$, for a predictable process $H_t \in \{-1, 1\}$, is also of some interest. [14] explicitly asks to find a characterisation of the predictable processes H for which ergodicity holds true. [14] adds that 'this seems to be an extremely difficult question, the solution to which has escaped so far both Brownian motion and ergodic theory experts'. We solve, completely, this question in the discrete setting. A previous work by Fujita [8] showed that the discrete Lévy transformation is not ergodic. We refer the interested reader to [5] for a modified version for which ergodicity does hold. We note here that in [5], the definition of the discrete Lévy transformation naturally requires $\mathrm{sgn}(0) = 0$. Since the predictable process H is not allowed to take the value 0 (only ± 1 are allowed), our setting is different to that in [5]. A consequence of our results is that if we pick a recycling rule uniformly at random, then it is almost surely non-ergodic.

2 Non-anticipative bootstrapping – two representations

Let $\xi_1, \xi_2, \dots$ be independent and identically distributed random variables with

$$\mathbb{P}(\xi_1 = -1) = \mathbb{P}(\xi_1 = +1) = \frac{1}{2}.$$

Given a sequence of functions $\phi_n : \{-1, +1\}^n \longrightarrow \{-1, +1\}$, we define

$$\eta_n = \phi_n(\xi_1, \dots, \xi_n).$$

As mentioned above, a direct consequence of the martingale representation property for the simple symmetric random walk $(X_n)_{n\geq 0}$ yields a first representation of η_n in terms of the sequence $(\xi_n)_{n\geq 1}$.

Proposition 2.1. *$(\eta_n)_{n\geq 1} \overset{d}{=} (\xi_n)_{n\geq 1}$ if and only if $\phi_n(u_1, \dots, u_n)$ is of the following form:*

$$\phi_n(u_1, \dots, u_n) = \psi_{n-1}(u_1, \dots, u_{n-1})u_n, \tag{2.1}$$

where $\psi_{n-1}(u_1, \dots, u_{n-1})$ is any function of $(u_1, \dots, u_{n-1})$ taking values in $\{-1, +1\}$ and $\psi_0 \in \{-1, +1\}$.

We note that the map on $\{-1,+1\}^n$ defined by (2.1) is invertible; that is for any $(v_1,\dots,v_n) \in \{-1,+1\}^n$, there exists a unique $(u_1,\dots,u_n) \in \{-1,+1\}^n$ such that $v_1 = \psi_0 u_1$ and for any $1 < k \leq n$, $v_k = \psi_{k-1}(u_1,\dots,u_{k-1})u_k$. It follows that the filtrations generated by $(\eta_n)_{n\geq 1}$ and $(\xi_n)_{n\geq 1}$ are identical.

Next, we parametrise the functions ψ_n in terms of the $\max$ function as a "building block". To this end, we introduce the following notations. $\mathbb{K}(n) = \mathcal{P}(\{1,\dots,n\})$ denotes the power set of $\{1,\dots,n\}$ and for $(u_1,\dots,u_n) \in \{-1,+1\}^n$:

$$u_{[\emptyset]} = -1 \text{ and for } K \in \mathbb{K}(n) \setminus \{\emptyset\}, u_{[K]} = \max_{k\in K} u_k.$$

Proposition 2.2. *A function ψ on $\{-1,+1\}^n$ takes values in $\{-1,+1\}$ if and only if it can be written as*

$$\psi(u_1,\dots,u_n) = \prod_{K\in\mathbb{K}(n)} u_{[K]}^{\beta_K} \tag{2.2}$$

where $\beta_K \in \{0,1\}$. Furthermore, this representation is unique.

Proof. We first establish the uniqueness of the representation. To this end we must solve the system of equations

$$\prod_{K\in\mathbb{K}(n)} u_{[K]}^{\beta_K} = \prod_{K\in\mathbb{K}(n)} u_{[K]}^{\beta'_K},\ \forall u_1,\dots,u_n. \tag{2.3}$$

Solving the equation for $u_1 = \dots = u_n = 1$ immediately yields $\beta_\emptyset = \beta'_\emptyset$.

Let $\mathbb{B} = \{K \in \mathbb{K}(n) : \beta_K \neq \beta'_K\}$. We proceed by contradiction. Assume that $\mathbb{B} \neq \emptyset$ and let K^* be a minimal element in $\mathbb{B}$ (i.e. $K^* \in \mathbb{B}$ and either K^* is a singleton or for any $K \subsetneq K^*$, $\beta_K = \beta'_K$). We choose $u = (u_1,\dots,u_n)$ such that $u_k = -1$ for $k \in K^*$ and $u_k = 1$ for $k \notin K^*$. For such u, $u_{[K]} = 1$ unless $K \subset K^*$, and (2.3) reduces to

$$\prod_{\substack{K\in\mathbb{K}(n)\\ K\subset K^*}} u_{[K]}^{\beta_K} = \prod_{\substack{K\in\mathbb{K}(n)\\ K\subset K^*}} u_{[K]}^{\beta'_K}.$$

Since K^* is minimal in $\mathbb{B}$, we deduce that we must have $\beta_{n,K^*} = \beta'_{n,K^*}$, which contradicts the assumption that $K^* \in \mathbb{B}$.

To establish that every function on $\{-1,+1\}^n$ taking values in $\{-1,+1\}$ is of the form (2.2), we show that the two sets of functions have the same cardinality. Indeed, there are 2^{2^n} functions $\psi : \{-1,+1\}^n \longrightarrow \{-1,+1\}$ and there are as many choices of β_K, $K \in \mathbb{K}(n)$. ☐

While the use of the $\max$ function in (2.2) is natural (and in some way canonical), it is not the only "building block" one can use to represent functions $\psi : \{-1,+1\}^n \longrightarrow \{-1,+1\}$. We present here a generic way for constructing a representation of the form (2.2). We start by labeling all elements of $\{-1,+1\}^n$. We do so with the use of the sets $K \in \mathbb{K}(n)$ so that elements of $\{-1,+1\}^n$ are written as u_K, $K \in \mathbb{K}(n)$. Next we choose a partial order, $\prec$, on $\mathbb{K}(n)$ (which in turn induces a partial order on $\{-1,+1\}^n$). Finally, we let g_K, $K \in \mathbb{K}(n)$, be the family of functions on $\{-1,+1\}^n$ taking values in $\{-1,+1\}$, such that for any $K, K' \in \mathbb{K}(n)$, $g_K(u_{K'}) = -1$ if and only if $K \prec K'$. Then any function ψ on $\{-1,+1\}^n$ taking values in $\{-1,+1\}$ can be written as

$$\psi(u) = \prod_{K\in\mathbb{K}(n)} g_K(u)^{\beta_K},$$

where $\beta_K \in \{0,1\}$. This representation is unique in the sense that there is a one-to-one correspondence between the functions ψ and the sequences β_K. The proof of this statement is an immediate adaptation of the proof of Proposition 2.2.

In other words, any partial order on $\mathbb{K}(n)$ (and labelling of $\{-1,+1\}^n$) defines a new set of building blocks g_K, $K \in \mathbb{K}(n)$, that can be used to produce a representation of the form (2.2).

The $\max$ function is obtained by labeling $u \in \{-1,+1\}^n$ with the set of indices carrying the value -1, $K = \{k : u_k = -1\}$, and using the usual inclusion as a partial order on $\mathbb{K}(n)$.

The $\min$ function can also be used as can easily be seen by replacing u with $-u$ in the $\max$ representation, and making all necessary adjustments.

Another example can be constructed by thinking of $\mathbb{K}(n)$ (or equivalently of $\{-1,+1\}^n$) as a totally unordered set. Then one can simply use the functions

$$g_K(u) = \begin{cases} -1 & \text{if } u = u_K \\ 1 & \text{if } u \neq u_K \end{cases}$$

The remainder of this paper uses the $\max$ function as the building block: $g_K(u) = u_{[K]} = \max_{k\in K} u_k$.

The dependence between the random variables $\xi_{[K]} = \max_{k\in K} \xi_k$, for various K's, renders the computation of quantities such as $\mathbb{E}\Big[\prod_{K\in\mathbb{K}(n)} \xi_{[K]}^{\beta_K}\Big]$ cumbersome. The next proposition enables a linearisation of the product $\prod_{K\in\mathbb{K}(n)} \xi_{[K]}^{\beta_K}$ and therefore a better handle on its expectation.

Proposition 2.3. *For any collection of sets of integers $M_1, \dots, M_m$,*

$$\prod_{k=1}^{m} u_{[M_k]} = \frac{1}{2}(-1)^m - \frac{1}{2}\sum_{k=0}^{m}(-2)^k \sum_{\substack{K\in\mathbb{K}(m)\\ |K|=k}} u_{[M_{[K]}]},$$

where $M_{[K]} = \bigcup_{j\in K} M_j$ and $M_{[\emptyset]} = \emptyset$.

Proof. The identity is clearly true for $m=1$: $-\frac{1}{2}+\frac{1}{2}-\frac{1}{2}(-2)u_{[M_1]} = u_{[M_1]}$. Suppose the identity true for $m-1$. Then,

$$\begin{aligned}
&\prod_{k=1}^{m} u_{[M_k]} = \left(\prod_{k=1}^{m-1} u_{[M_k]}\right) u_{[M_m]} \\
=\;& \frac{1}{2}(-1)^{m-1}u_{[M_m]} - \frac{1}{2}\sum_{k=0}^{m-1}(-2)^k \sum_{\substack{K\in\mathbb{K}(m-1)\\ |K|=k}} u_{[M_{[K]}]}u_{[M_m]} \\
=\;& \frac{1}{2}(-1)^{m-1}u_{[M_m]} - \frac{1}{2}\sum_{k=0}^{m-1}(-2)^k \sum_{\substack{K\in\mathbb{K}(m-1)\\ |K|=k}} \Big(1 + u_{[M_{[K]}]} + u_{[M_m]} - 2u_{[M_{[K]}\cup M_m]}\Big).
\end{aligned}$$

Using the facts that

$$\sum_{k=0}^{m-1}(-2)^k\binom{m-1}{k} = (-1)^{m-1}$$

and any $K \in \mathbb{K}(m)$ is either in $\mathbb{K}(m-1)$ or is of the form $K' \cup \{m\}$, where $K' \in \mathbb{K}(m-1)$, we see that

$$\begin{aligned}
\prod_{k=1}^{m} u_{[M_k]} &= -\frac{1}{2}(-1)^m - \frac{1}{2}\sum_{k=0}^{m-1}(-2)^k \sum_{\substack{K\in\mathbb{K}(m-1)\\|K|=k}} \left(u_{[M_{[K]}]} - 2u_{[M_{[K]}\cup M_m]}\right)\\
&= -\frac{1}{2}(-1)^m - \frac{1}{2}\sum_{K\in\mathbb{K}(m-1)}(-2)^{|K|}u_{[M_{[K]}]} - \frac{1}{2}\sum_{\substack{K\in\mathbb{K}(m-1)\\K'=K\cup\{m\}}}(-2)^{|K'|}u_{[M_{[K']}]}\\
&= -\frac{1}{2}(-1)^m - \frac{1}{2}\sum_{K\in\mathbb{K}(m)}(-2)^{|K|}u_{[M_{[K]}]}\\
&= \frac{1}{2}(-1)^m - \frac{1}{2}\sum_{k=0}^{m}(-2)^k \sum_{\substack{K\in\mathbb{K}(m)\\|K|=k}} u_{[M_{[K]}]}. \quad \square
\end{aligned}$$

Combining Propositions 2.1, 2.2 and 2.3, we obtain the following theorem.

Theorem 2.4. *Let $\eta_n = \phi_n(\xi_1,\ldots,\xi_n)$, then $(\eta_n)_{n\geq 1} \stackrel{d}{=} (\xi_n)_{n\geq 1}$ if and only if for each n and each $K\in\mathbb{K}(n-1)$, there exists $\beta_{n,K}\in\{0,1\}$ such that*

$$\phi_n(u_1,\ldots,u_n) = \left(\prod_{K\in\mathbb{K}(n-1)} u_{[K]}^{\beta_{n,K}}\right)u_n. \tag{2.4}$$

Furthermore, for such functions

$$\phi_n(u_1,\ldots,u_n) = \left(\frac{1}{2}(-1)^{|\mathcal{B}(n)|} - \frac{1}{2}\sum_{H\in\mathcal{P}(\mathcal{B}(n))}(-2)^{|H|}u_{[\langle H\rangle]}\right)u_n$$

where $\mathcal{B}(n) = \{K\in\mathbb{K}(n-1) : \beta_{n,K}=1\}$ and for $H=\{K_1,\ldots,K_h\}\in\mathcal{P}(\mathcal{B}(n))$, $\langle H\rangle = \bigcup_{j=1}^h K_j$.

For the remainder of the paper,

$$\eta_n = \xi_n\prod_{K\in\mathbb{K}(n-1)}\xi_{[K]}^{\beta_{n,K}} \text{ and } \zeta_{n-1} = \prod_{K\in\mathbb{K}(n-1)}\xi_{[K]}^{\beta_{n,K}} = \eta_n\xi_n.$$

Note that these quantities can be expressed in terms of the sets $\mathcal{B}(n)$ instead of the sets $\mathbb{K}(n-1)$ as follows

$$\eta_n = \xi_n\prod_{K\in\mathcal{B}(n)}\xi_{[K]} \text{ and } \zeta_{n-1} = \prod_{K\in\mathcal{B}(n)}\xi_{[K]} = \eta_n\xi_n.$$

For any set of integers M, we let $q(M) = \mathbb{P}(\xi_{[M]} = -1) = 2^{-|M|}$. Then $\mathbb{E}[\xi_{[M]}] = 1-2q(M)$.

Corollary 2.5. *For any $k\geq 2$,*

$$\mathbb{E}[\zeta_{k-1}] = \sum_{H\in\mathcal{P}(\mathcal{B}(k))}(-2)^{|H|}q(\langle H\rangle)$$

and any $k\neq\ell$,

$$\mathbb{E}[\zeta_{k-1}\zeta_{\ell-1}] = \sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\J\in\mathcal{P}(\mathcal{B}(\ell))}}(-2)^{|H|+|J|}q(\langle H\rangle\cup\langle J\rangle).$$

3 The Gaussian case

Let $U_t^{(n)} = \frac{1}{\sqrt{n}} X_{\lfloor nt \rfloor}$ and $V_t^{(n)} = \frac{1}{\sqrt{n}} Y_{\lfloor nt \rfloor}$. It is well-known that $U_t^{(n)}$, $t \in [0,1]$, converges weakly to a standard Brownian motion, and the same is also trivially true for $V_t^{(n)}$, $t \in [0,1]$. In this section we give necessary and sufficient conditions on the parameters of the GBRW for the two-dimensional process $W_t^{(n)} = (U_t^{(n)}, V_t^{(n)})$, $t \in [0,1]$, to converge to a two-dimensional Brownian motion.

3.1 The main result

Theorem 3.1. *$W^{(n)}$ converges weakly to a two-dimensional Brownian motion (possibly degenerate) with correlation ρ if and only if*

(A) $\rho = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{H \in \mathcal{P}(\mathcal{B}(k))} (-2)^{|H|} q(\langle H \rangle)$ *exists;*

(B) $\lim_{n\to\infty} \frac{1}{n^2} \sum_{\ell=1}^{n} \sum_{k=1}^{n} \sum_{\substack{H \in \mathcal{P}(\mathcal{B}(k)) \\ J \in \mathcal{P}(\mathcal{B}(\ell))}} (-2)^{|H|+|J|} q(\langle H \rangle \cup \langle J \rangle) = \rho^2.$

Proof. For each n, $U^{(n)}$ and $V^{(n)}$ are martingales with respect to $\mathcal{F}_t^{(n)} = \sigma(\xi_1, \ldots, \xi_{\lfloor nt \rfloor})$. Furthermore

$$\Delta U_t^{(n)} = \begin{cases} 0 & nt \notin \mathbb{N} \\ \frac{1}{\sqrt{n}} \xi_{nt} & nt \in \mathbb{N} \end{cases}, \; |\Delta U_t^{(n)}| \le \frac{1}{\sqrt{n}} |\xi_{\lfloor nt \rfloor}| \le 1$$

and similarly for $V_t^{(n)}$.

Using Theorem VIII.3.11 of [10], we see that $W^{(n)}$ converges weakly to a two-dimensional Brownian motion with correlation ρ if and only if $[U^{(n)}, V^{(n)}]_t$ converges in probability to ρt, where

$$[U^{(n)}, V^{(n)}]_t = \sum_{s \le t} \Delta U_s^{(n)} \Delta V_s^{(n)} = \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \xi_k \eta_k = \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} \zeta_{k-1}.$$

Suppose $[U^{(n)}, V^{(n)}]_t$ converges in probability to ρt. Since $|[U^{(n)}, V^{(n)}]_t| \le \lfloor nt \rfloor / n \le 1$, we must have

$$\rho = \lim_n \mathbb{E}\big[[U^{(n)}, V^{(n)}]_1\big] = \lim_n \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[\zeta_{k-1}] = \lim_n \frac{1}{n} \sum_{k=1}^{n} \sum_{H \in \mathcal{P}(\mathcal{B}(k))} (-2)^{|H|} q(\langle H \rangle),$$

from which we deduce that (A) must hold. We must also have

$$\begin{aligned} \rho^2 &= \lim_n \mathbb{E}\big[[U^{(n)}, V^{(n)}]_1^2\big] = \lim_n \frac{1}{n^2} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \mathbb{E}[\zeta_{k-1} \zeta_{\ell-1}] \\ &= \lim_n \frac{1}{n^2} \sum_{k=1}^{n} \sum_{\ell=1}^{n} \sum_{\substack{H \in \mathcal{P}(\mathcal{B}(k)) \\ J \in \mathcal{P}(\mathcal{B}(\ell))}} (-2)^{|H|+|J|} q(\langle H \rangle \cup \langle J \rangle) \end{aligned}$$

and (B) must also hold proving the necessity of these conditions.

Next we show sufficiency. In fact we shall prove that (A) and (B) lead to an L^2 convergence of $[U^{(n)}, V^{(n)}]_t$ to ρt.

Let

$$\rho_k = \mathbb{E}[\zeta_{k-1}] = \sum_{H \in \mathcal{P}(\mathcal{B}(k))} (-2)^{|H|} q(\langle H \rangle)$$

and

$$\theta_{k,\ell} = \mathbb{E}[\zeta_{k-1}\zeta_{\ell-1}] = \sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\ J\in\mathcal{P}(\mathcal{B}(\ell))}} (-2)^{|H|+|J|} q(\langle H\rangle \cup \langle J\rangle).$$

Then, assuming (A) and (B),

$$\begin{aligned}
&\mathbb{E}\left[\left(\frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor}\zeta_{k-1} - \rho t\right)^2\right]\\
=\;& \frac{1}{n^2}\sum_{\ell=1}^{\lfloor nt\rfloor}\sum_{k=1}^{\lfloor nt\rfloor}\mathbb{E}[\zeta_{k-1}\zeta_{\ell-1}] - \frac{2\rho t}{n}\sum_{k=1}^{\lfloor nt\rfloor}\mathbb{E}[\zeta_{k-1}] + \rho^2 t^2\\
=\;& \frac{1}{n^2}\sum_{\ell=1}^{\lfloor nt\rfloor}\sum_{k=1}^{\lfloor nt\rfloor}\theta_{k,\ell} - \frac{2\rho t}{n}\sum_{k=1}^{\lfloor nt\rfloor}\rho_k + \rho^2 t^2 \underset{n\uparrow\infty}{\longrightarrow} \rho^2t^2 - 2\rho^2t^2 + \rho^2t^2 \;=\; 0 \quad \square
\end{aligned}$$

The asymptotics of $W^{(n)}$ is dependent on the (asymptotic) behaviour of the families $\mathcal{B}(n)$. The following example provides settings in which the latter is easily described.

Example 3.2. Suppose $\eta_k = \psi_{k-2}(\xi_1,\dots,\xi_{k-2})\xi_{k-1}\xi_k$, where ψ_{k-2} is any function on $\{-1,+1\}^{k-2}$ taking values in $\{-1,+1\}$. In this case, for any $k<\ell$,

$$\mathbb{E}[\zeta_{k-1}] = 0 \text{ and } \mathbb{E}[\zeta_{k-1}\zeta_{\ell-1}] = \mathbb{E}[\psi_{k-2}(\xi_1,\dots,\xi_{k-2})\xi_{k-1}\psi_{\ell-2}(\xi_1,\dots,\xi_{\ell-2})]\mathbb{E}[\xi_{\ell-1}] = 0.$$

It immediately follows that $W^{(n)}$ converges weakly to a pair of independent Brownian motions.

3.2 The extended Bootstrap random walk

We extend the model introduced in [3] and [4] to the case $\eta_n = \xi_n\prod_{k\in M_n}\xi_k,\ n\geq 1$, where M_n is not necessarily the entire set $\{1,\dots,n-1\}$, but any subset thereof. We investigate the convergence of $W^{(n)}$ in terms of the behaviour of the sets M_n. Of particular interest is the case of consecutive indexes, $M_n = \{1,\dots,\lfloor R(n)\rfloor\}$, for some real function R. Note that the case $M_n = \{\lfloor r(n)\rfloor,\dots,n-1\}$ is covered by Example 3.2.

Proposition 3.3. *Suppose $\mathcal{B}(n)$ is made up of disjoint sets of equal cardinality ($\forall K_1,K_2\in\mathcal{B}(n)$, $|K_1|=|K_2|$ and $K_1\cap K_2=\emptyset$, whenever $K_1\neq K_2$). Call κ the cardinality.*

1. *If $\kappa=1$ or if $\kappa>1$ and $\lim_n|\mathcal{B}(n)|=+\infty$, then (A) holds true with $\rho=0$.*
2. *If $\kappa>1$ and $\lim_n|\mathcal{B}(n)|=m<+\infty$, then (A) holds true with $\rho=\left(1-2^{1-\kappa}\right)^m$.*

Proof. For $H\in\mathcal{P}(\mathcal{B}(n))$, $|\langle H\rangle| = \sum_{K\in H}|K| = \kappa|H|$. It follows that for such an H, $q(\langle H\rangle) = 2^{-|\langle H\rangle|} = 2^{-\kappa|H|}$ and

$$\begin{aligned}
&\sum_{H\in\mathcal{P}(\mathcal{B}(n))}(-2)^{|H|}q(\langle H\rangle)\\
=\;& \sum_{H\in\mathcal{P}(\mathcal{B}(n))}(-2)^{|H|}2^{-\kappa|H|} \;=\; \sum_{H\in\mathcal{P}(\mathcal{B}(n))}\left(-2^{1-\kappa}\right)^{|H|}\\
=\;& \sum_{h=0}^{|\mathcal{B}(n)|}\sum_{\substack{H\in\mathcal{P}(\mathcal{B}(n))\\ |H|=h}}\left(-2^{1-\kappa}\right)^h \;=\; \sum_{h=0}^{|\mathcal{B}(n)|}\binom{|\mathcal{B}(n)|}{h}\left(-2^{1-\kappa}\right)^h \;=\; \left(1-2^{1-\kappa}\right)^{|\mathcal{B}(n)|}
\end{aligned}$$

The result immediately follows. $\square$

Next we focus on the case $\kappa=1$. We call the resulting process the Extended Bootstrap Random Walk.

Corollary 3.4. *Suppose $\zeta_{k-1} = \prod_{j\in M_k} \xi_j$, where $M_k \in \mathbb{K}(k-1)$ (i.e. $\kappa = 1$).*

1. *If $\lim_n \frac{1}{n}\sum_{k=1}^n 1_{M_k=M_n} = 0$, then $W^{(n)}$ converges weakly to a pair of independent Brownian motions.*
2. *Suppose further that $M_k \subseteq M_{k+1}$ and let $N(n) = \min\{k : M_k = M_n\}$.*

   If $\lim_n N(n)/n = 1$, then $W^{(n)}$ converges weakly to a pair of independent Brownian motions.

Proof. 1. Proposition 3.3 guarantees that (A) holds true with $\rho = 0$. We check (B):

$$\frac{1}{n^2}\sum_{\ell=1}^n\sum_{k=1}^n \mathbb{E}[\zeta_{k-1}\zeta_{\ell-1}] = \frac{1}{n} + \frac{2}{n^2}\sum_{\ell=1}^n\sum_{k=1}^\ell 1_{M_k=M_\ell} \le \frac{1}{n} + \frac{2}{n}\sum_{\ell=1}^n \frac{1}{\ell}\sum_{k=1}^\ell 1_{M_k=M_\ell}.$$

Under the assumption that $\lim_n \frac{1}{n}\sum_{k=1}^n 1_{M_k=M_n} = 0$,

$$\lim_n \frac{1}{n^2}\sum_{\ell=1}^n\sum_{k=1}^n \mathbb{E}[\zeta_{k-1}\zeta_{\ell-1}] = 0 = \rho^2$$

and $W^{(n)}$ converges weakly to a pair of independent Brownian motions.

2. Since $M_k \subseteq M_{k+1} \subseteq \{1,\dots,k\}$, $\frac{1}{n}\sum_{k=1}^n 1_{M_k=M_n} = (n - N(n) + 1)/n$. Therefore (B) holds true as soon as $\lim_n N(n)/n = 1$. □

The next example shows that linear growth leads to a two-dimensional Brownian motion, while logarithmic growth does not. The corollary that follows makes these observations precise.

Example 3.5. 1. Fix $\lambda \in (0,1)$ and suppose $M_k = \{1,\dots,\lfloor \lambda k\rfloor\}$, for $k > 1/\lambda$. Then, for $n > 1/\lambda$, $N(n) > (\lfloor \lambda n\rfloor - 1)/\lambda$, and $\lim_n N(n)/n = 1$.

2. Suppose $M_k = \{1,\dots,m(k)\}$, where $m(k) = \lfloor \ln k\rfloor$ for $k \ge 3$. Then, for any $\ell \in \{\lfloor e^{n-1}\rfloor + 1,\dots,\lfloor e^n\rfloor\}$, $m(\ell) = n-1$. It follows that for any $\ell \in \{\lfloor e^{n-1}\rfloor + 1,\dots,\lfloor e^n\rfloor\}$, $N(\ell) = \lfloor e^{n-1}\rfloor + 1$ and, for $n \ge 2$,

$$\frac{N(\lfloor e^n\rfloor)}{\lfloor e^n\rfloor} = \frac{\lfloor e^{n-1}\rfloor + 1}{\lfloor e^n\rfloor} < \frac{e^{n-1}+1}{e^n - 1} \le \frac{e+1}{e^2-1} < 1.$$

However, $m(\lfloor e^n\rfloor + 1) = n$ and $N(\lfloor e^n\rfloor + 1) = \lfloor e^n\rfloor + 1$. It follows that $N(n)/n$ does not converge. Furthermore, for $\overline{n} = \lfloor e^{n+1}\rfloor$ and $\underline{n} = \lfloor e^n\rfloor + 1$,

$$\frac{2}{\overline{n}^2}\sum_{\ell=2}^{\overline{n}}\sum_{k=1}^{\ell-1} 1_{M_k=M_\ell} = \frac{2}{\overline{n}^2}\sum_{\ell=2}^{\overline{n}}(\ell - N(\ell)) \ge \frac{2}{\overline{n}^2}\sum_{\ell=\underline{n}}^{\overline{n}}(\ell - \underline{n}) = \frac{(\overline{n} - \underline{n} + 1)(\overline{n} - \underline{n})}{\overline{n}^2}.$$

Since the last term approaches $1 - 2e^{-1} + e^{-2} > 0$, (B) fails and $W^{(n)}$ does not converge to a two-dimensional Brownian motion. Recall that Proposition 3.3 guarantees that (A) holds true with $\rho = 0$.

Corollary 3.6. *Suppose $\zeta_{k-1} = \prod_{j=1}^{\lfloor R(k)\rfloor} \xi_j$, where R is a strictly increasing, continuous, unbounded and regularly varying function with the property that $1 \le R(z) < z$. If the regular variation index α is positive ($\alpha > 0$), then $W^{(n)}$ converges weakly to a pair of independent Brownian motions.*

Proof. Here $M_k = \{1,\dots,\lfloor R(k)\rfloor\}$. Let $k = \lfloor R^{-1}(\lfloor R(n)\rfloor)\rfloor + 1$. Then $R^{-1}(\lfloor R(n)\rfloor) < k$, $\lfloor R(n)\rfloor < g(k)$ and $\lfloor R(k)\rfloor = \lfloor R(n)\rfloor$. Therefore, $N(n) \le \lfloor R^{-1}(\lfloor R(n)\rfloor)\rfloor + 1$.

Now, let $k = \lfloor R^{-1}(\lfloor R(n)\rfloor)\rfloor - 1$. Then $k+1 \le R^{-1}(\lfloor R(n)\rfloor)$, $R(k) < R(k+1) \le \lfloor R(n)\rfloor$ and $\lfloor R(k)\rfloor < \lfloor R(n)\rfloor$. Therefore, $N(n) > \lfloor R^{-1}(\lfloor R(n)\rfloor)\rfloor - 1$.

In total,

$$\lfloor R^{-1}(\lfloor R(n) \rfloor) \rfloor - 1 < N(n) \leq \lfloor R^{-1}(\lfloor R(n) \rfloor) \rfloor + 1.$$

Since $\alpha > 0$, $\forall\, \varepsilon > 0$, $\exists n^*$ such that $\forall n \geq n^*$,

$$\frac{R((1-\varepsilon)n)}{R(n)} + \frac{1}{R(n)} < 1$$

or equivalently that

$$1 - \varepsilon < \frac{1}{n} R^{-1}(R(n) - 1) < \frac{1}{n} R^{-1}(\lfloor R(n) \rfloor) \leq 1.$$

It immediately follows that $\lim_n N(n)/n = 1$ and therefore that $W^{(n)}$ converges weakly to a pair of independent Brownian motions. $\square$

Remark 3.7. Note that if the function R is slowly varying ($\alpha = 0$), then convergence of $N(n)/n$ is not guaranteed as can be seen from the case $R(z) = \ln z$.

3.3 The bounded case

Here we look at the case where the number of non-empty sets in $\mathbb{K}(k-1)$ used in the recycling is fixed and equal to say m. In other words, $|\mathcal{B}(n) \setminus \{\emptyset\}| = m$.

Corollary 3.8. *Suppose* $\zeta_{k-1} = \varepsilon_k \prod_{i=1}^m \xi_{[M_k^{(i)}]}$, *where* $M_k^{(i)} \in \mathbb{K}(k-1) \setminus \{\emptyset\}$, $\varepsilon_k \in \{-1, 1\}$, *and, writing* $M_k^{[K]}$ *for* $\bigcup_{i \in K} M_k^{(i)}$,

(A$_b$) for any $K \in \mathbb{K}(m)$, $\gamma(K) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n \varepsilon_k q\big(M_k^{[K]}\big)$ *exists;*

(B$_b$) for any $K_1, K_2 \in \mathbb{K}(m)$, $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^n q\big(M_k^{[K_1]} \cap M_{n+1}^{[K_2]}\big) = 1.$

Then $(U^{(n)}, V^{(n)})$ *converges weakly to a two-dimensional Brownian motion (possibly degenerate) with correlation* $\sum_{K \in \mathbb{K}(m)} (-2)^{|K|} \gamma(K)$.

Proof. Note that $\varepsilon_k = \xi_{[\emptyset]}^{\beta_{k,\emptyset}}$ and $\varepsilon_k = -1$ if and only if $\emptyset \in \mathcal{B}(k)$. Therefore, $\mathcal{B}(k) = \{M_k^{(1)}, \ldots, M_k^{(m)}\}$ or $\{\emptyset, M_k^{(1)}, \ldots, M_k^{(m)}\}$ depending on whether $\varepsilon_k = 1$ or -1. It follows that selecting $H \in \mathcal{P}(\mathcal{B}(k))$ reduces to choosing $K \in \mathbb{K}(m)$. If $\varepsilon_k = 1$, then

$$\sum_{H \in \mathcal{P}(\mathcal{B}(k))} (-2)^{|H|} q(\langle H \rangle) = \sum_{K \in \mathbb{K}(m)} (-2)^{|K|} q\big(M_k^{[K]}\big).$$

If $\varepsilon_k = -1$, then

$$\begin{aligned} & \sum_{H \in \mathcal{P}(\mathcal{B}(k))} (-2)^{|H|} q(\langle H \rangle) \\ = \quad & \sum_{K \in \mathbb{K}(m)} (-2)^{|K|} q\big(M_k^{[K]}\big) + \sum_{K \in \mathbb{K}(m)} (-2)^{|K|+1} q\big(M_k^{[K]}\big) \\ = \quad & - \sum_{K \in \mathbb{K}(m)} (-2)^{|K|} q\big(M_k^{[K]}\big), \end{aligned}$$

and (A$_b$) clearly implies (A) with $\rho = \sum_{K \in \mathbb{K}(m)} (-2)^{|K|} \gamma(K)$.

To show that (B$_b$) implies (B), we note that, assuming (A$_b$) holds,

$$\lim_n \frac{1}{n^2}\sum_{\ell=1}^{n}\sum_{k=1}^{n}\sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\ J\in\mathcal{P}(\mathcal{B}(\ell))}}(-2)^{|H|+|J|}q(\langle H\rangle)q(\langle J\rangle)$$
$$= \left(\lim_n \frac{1}{n}\sum_{k=1}^{n}\sum_{H\in\mathcal{P}(\mathcal{B}(k))}(-2)^{|H|}q(\langle H\rangle)\right)^2 = \rho^2$$

and we write

$$\frac{1}{n^2}\sum_{\ell=1}^{n}\sum_{k=1}^{n}\sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\ J\in\mathcal{P}(\mathcal{B}(\ell))}}(-2)^{|H|+|J|}\big(q(\langle H\rangle\cup\langle J\rangle)-q(\langle H\rangle)q(\langle J\rangle)\big)$$
$$= \frac{1}{n^2}\sum_{\ell=1}^{n}\sum_{k=1}^{n}\sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\ J\in\mathcal{P}(\mathcal{B}(\ell))}}(-2)^{|H|+|J|}q(\langle H\rangle\cup\langle J\rangle)\big(1-q(\langle H\rangle\cap\langle J\rangle)\big),$$

where we use the fact that $q(M_1\cup M_2)q(M_1\cap M_2)=q(M_1)q(M_2)$. We distinguish three cases: $\varepsilon_k=\varepsilon_\ell=1$, $\varepsilon_k=\varepsilon_\ell=-1$ and $\varepsilon_k\varepsilon_\ell=-1$. However, $\langle H\rangle$ and $\langle J\rangle$ are not affected by the inclusion of the empty set, and the only change is the increase by one of $|H|$ and $|J|$. It follows that

$$\sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\ J\in\mathcal{P}(\mathcal{B}(\ell))}}(-2)^{|H|+|J|}q(\langle H\rangle\cup\langle J\rangle)\left(1-q(\langle H\rangle\cap\langle J\rangle)\right)$$
$$= \varepsilon_k\varepsilon_\ell\sum_{K_1,K_2\in\mathbb{K}(m)}(-2)^{|K_1|+|K_2|}q\big(M_k^{[K_1]}\cup M_\ell^{[K_2]}\big)\Big(1-q\big(M_k^{[K_1]}\cap M_\ell^{[K_2]}\big)\Big)$$

and that

$$\Big|\frac{1}{n^2}\sum_{\ell=1}^{n}\sum_{k=1}^{n}\sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\ J\in\mathcal{P}(\mathcal{B}(\ell))}}(-2)^{|H|+|J|}\left(q(\langle H\rangle\cup\langle J\rangle)-q(\langle H\rangle)q(\langle J\rangle)\right)\Big|$$
$$\leq 2^{2m}\sum_{K_1,K_2\in\mathbb{K}(m)}\frac{1}{n^2}\sum_{\ell=1}^{n}\sum_{k=1}^{n}\Big(1-q\big(M_k^{[K_1]}\cap M_\ell^{[K_2]}\big)\Big)$$

If we let $a_n=\sum_{\ell=2}^{n}\sum_{k=1}^{\ell-1}\Big(1-q\big(M_k^{[K_1]}\cap M_\ell^{[K_2]}\big)\Big)$, then

$$\frac{a_{n+1}-a_n}{(n+1)^2-n^2}=\frac{1}{2n+1}\sum_{k=1}^{n}\Big(1-q\big(M_k^{[K_1]}\cap M_{n+1}^{[K_2]}\big)\Big)\longrightarrow 0,$$

and the result follows immediately by application of Stolz-Cesàro Theorem. □

Example 3.9. Suppose $\eta_k=\xi_{[\emptyset]}^{\beta_{k,\emptyset}}\xi_k=\varepsilon_k\xi_k$ and let $E_n=\{k\leq n:\varepsilon_k=-1\}=\{k\leq n:\beta_{k,\emptyset}=1\}$. Although this setting is not that of Corollary 3.8, similar and in fact more straightforward calculations, show that if $p=\lim_n|E_n|/n$ exists, then $W^{(n)}$ converges weakly to a two-dimensional Brownian motion with correlation $\rho=1-2p$. In particular and since clearly p can be made to take any value in $[0,1]$, given any $\rho\in[-1,1]$, one can construct a GBRW that, when suitably normalized, converges to a two-dimensional Brownian motion with correlation ρ.

Corollary 3.10. *Suppose $\zeta_{k-1} = \xi_{[M_k]}$, where $M_k \in \mathbb{K}(k-1)$.*

1. *If $\lim_n |M_n| = +\infty$, then $(U^{(n)}, V^{(n)})$ converges to a degenerate Brownian motion.*
2. *If $\lim_n |M_n|$ exists and is finite, and $\limsup_n M_n = \emptyset$ (no index appears infinitely many times) then $(U^{(n)}, V^{(n)})$ converges to a (proper) Brownian motion with correlation $\rho = 1 - 2^{1-m}$, where $m = \lim_n |M_n|$.*

Proof. 1. We check that (A) of Theorem 3.1 holds true with $\rho = 1$. Indeed, $\mathcal{B}(k) = \{M_k\}$ and

$$\begin{aligned}\lim_n \frac{1}{n}\sum_{k=1}^n \sum_{H\in\mathcal{P}(\mathcal{B}(k))} (-2)^{|H|} q(\langle H\rangle) &= \lim_n \frac{1}{n}\sum_{k=1}^n \sum_{H\in\{\emptyset, M_k\}} (-2)^{|H|} q(\langle H\rangle)\\ &= \lim_n \frac{1}{n}\sum_{k=1}^n (1-2q(M_k)) = 1.\end{aligned}$$

Similarly,

$$\begin{aligned}&\frac{1}{n^2}\sum_{\ell=1}^n\sum_{k=1}^n \sum_{\substack{H\in\mathcal{P}(\mathcal{B}(k))\\ J\in\mathcal{P}(\mathcal{B}(\ell))}} (-2)^{|H|+|J|} q(\langle H\rangle \cup \langle J\rangle)\\ &= \frac{1}{n^2}\sum_{\ell=1}^n\sum_{k=1}^n (1 - 2q(M_k) - 2q(M_\ell) + 4q(M_k\cup M_\ell))\\ &= 1 - \frac{4}{n}\sum_{k=1}^n q(M_k) + \frac{4}{n^2}\sum_{\ell=1}^n\sum_{k=1}^n q(M_k\cup M_\ell),\end{aligned}$$

and (B) follows. Applying Theorem 3.1 completes the proof.

2. In this case $\lim_n \sum_{H\in\mathcal{P}(\mathcal{B}(n))} (-2)^{|H|} q(\langle H\rangle) = \lim_n (1 - 2^{1-|M_n|})$ exists and is strictly smaller than 1. Next we check that (B) holds. Indeed,

$$\begin{aligned}&\left|\frac{1}{n^2}\sum_{k,\ell=1}^n \sum_{\substack{K\in\mathcal{P}(\mathcal{B}(k))\\ L\in\mathcal{P}(\mathcal{B}(\ell))}} (-2)^{|K|+|L|} q(\langle K\rangle\cup\langle L\rangle) - \left(\frac{1}{n}\sum_{k=1}^n \sum_{H\in\mathcal{P}(\mathcal{B}(k))} (-2)^{|H|} q(\langle H\rangle)\right)^2\right|\\ &= \frac{1}{n^2}\sum_{k,\ell=1}^n \big(1 - 2q(M_k) - 2q(M_\ell) + 4q(M_k\cup M_\ell) - (1-2q(M_k)(1-2q(M_\ell))\big)\\ &= \frac{4}{n^2}\sum_{k,\ell=1}^n \big(q(M_k\cup M_\ell) - q(M_k)q(M_\ell)\big)\\ &= \frac{4}{n^2}\sum_{k,\ell=1}^n q(M_k\cup M_\ell)(1-q(M_k\cap M_\ell))\\ &\le \frac{4}{n^2}\sum_{k,\ell=1}^n (1-q(M_k\cap M_\ell))\\ &= \frac{4}{n^2}\sum_{k=1}^n (1-q(M_k)) + \frac{8}{n^2}\sum_{\ell=2}^n\sum_{k=1}^{\ell-1}(1-q(M_k\cap M_\ell))\end{aligned}$$

The first term clearly converges to 0. To obtain the limit of the second term, we let $a_n = \sum_{\ell=2}^n \sum_{k=1}^{\ell-1}(1-q(M_k\cap M_\ell))$ and observe that

$$\frac{a_{n+1}-a_n}{(n+1)^2-n^2} = \frac{1}{2n+1}\sum_{k=1}^n (1-q(M_k\cap M_{n+1})).$$

Since $|M_n|$ becomes constant (for n large enough), applying Lemma 5.4, we see that

$$\lim_n \frac{1}{n}\sum_{k=1}^{n}(1-q(M_k\cap M_{n+1})) \le \lim_n \frac{1}{n}\sum_{k=1}^{n}|M_k\cap M_{n+1}| = 0.$$

The result immediately follows. □

Suppose the dependence of η_n on $\xi_1,\ldots,\xi_{n-1}$ reduces to a dependence on the last m terms, where m is a fixed integer. At one end of the spectrum of settings is the case $\eta_n = \prod_{k=n-m}^{n}\xi_k$, or even $\eta_n = \psi_{n-2}(\xi_{n-m},\ldots,\xi_{n-2})\xi_{n-1}\xi_n$. Here, as we have seen in Example 3.2, the limiting process is a pair of independent Brownian motions. In the next example we look in some sense at the other end of the spectrum.

Example 3.11. Suppose $\eta_n = \max(\xi_{n-m},\ldots,\xi_{n-1})\xi_n$ so that ξ_{n-1} is tied to the successive increments $\xi_{n-m},\ldots,\xi_{n-2}$ with the use of the $\max$ function. As the latter is very sensitive to any of each component taking the value $+1$, the correlation increases to 1 as m increases to infinity. More specifically, the conditions of Corollary 3.10 are satisfied with $M_n = \{n-m,\ldots,n-1\}$ so that $|M_n| = m$ and $\limsup_n M_n = \emptyset$. It follows that $(U^{(n)},V^{(n)})$ converges to a Brownian motion with correlation $\rho = 1-2^{1-m}$.

4 The non-Gaussian case

It is not difficult to imagine situations where $(U^{(n)},V^{(n)})$ converges to a non-Gaussian limit or even fails to converge. In this section we look at a particular setting that leads to non-degenerate non-Gaussian limits.

First, we make the observation that $(U^{(n)})_{n\ge 0}$ and $(V^{(n)})_{n\ge 0}$ are C-tight (on $D_{\mathbb{R}}$) and therefore so is $((U^{(n)},V^{(n)}))_{n\ge 0}$ (on $D_{\mathbb{R}\times\mathbb{R}}$) – see [10], Corollary VI.3.33. The question we ask is what settings lead to convergence (of the whole sequence). Amongst those is the case of symmetric functions. A function is said to be symmetric if it is unchanged by any permutation of the coordinates. When dealing with functions on $\{-1,1\}^n$, these can be described in a succinct manner.

Lemma 4.1. *A function ψ on $\{-1,1\}^n$ is symmetric if and only if there exists a function f such that*

$$\psi(u_1,\ldots,u_n) = f(s_n),$$

where $s_n = \sum_{k=1}^{n} u_k$.

Proof. Let ν_n be the number of components equal to -1 in $(u_1,\ldots,u_n)$: $\nu_n = |\{k : u_k = -1\}|$. Since ψ is symmetric,

$$\psi(u_1,\ldots,u_n) = \psi(\underbrace{-1,\ldots,-1}_{\nu_n},\underbrace{+1,\ldots,+1}_{n-\nu_n}).$$

Therefore $\psi(u)$ depends on $(u_1,\ldots,u_n)$ only through ν_n. The result follows from the observation that $s_n = n-2\nu_n$. □

For example, in the special case of the product (studied in [3] and [4]), the ψ_n's are symmetric and $\prod_{k=1}^{n} u_k = (-1)^{(n-s_n)/2}$.

We wish to investigate the asymptotics of $(U^{(n)},V^{(n)})$ in the case of symmetric ψ_n's. More specifically, we assume that

$$\psi_{k-1}(u_1,\ldots,u_{k-1}) = f((u_1+\ldots+u_{k-1})/\sqrt{k}),$$

for some left-continuous function f taking values in $\{-1,1\}$. In other words, we assume that

$$\eta_k = \phi_k(\xi_1,\ldots,\xi_k) = \psi_{k-1}(\xi_1,\ldots,\xi_{k-1})\xi_k = f(X_{k-1}/\sqrt{k})\xi_k.$$

Since, for $t \in [0,1]$,

$$\Delta U_t^{(n)} = U_t^{(n)} - U_{t-}^{(n)} = \begin{cases} \xi_k/\sqrt{n} & t = k/n,\ k = 1,\dots,n \\ 0 & \text{otherwise} \end{cases}$$

the above assumptions lead to

$$V_t^{(n)} = \frac{1}{\sqrt{n}} \sum_{k=1}^{\lfloor nt \rfloor} f(X_{k-1}/\sqrt{k})\xi_k = \int_{(0,t]} f(U_{s-}^{(n)}/\sqrt{s}) dU_s^{(n)}.$$

We shall further assume that the set $\Xi_f = \{z \in \mathbb{R} : \Xi f(z) = f(z) - \lim_{z' \downarrow z} f(z') \neq 0\}$ is non-empty and has positive minimum gap; that is $\delta_0 = \inf_{a \in \Xi_f} \inf_{b \in \Xi_f \setminus \{a\}} |b-a| > 0$. Note that if $a \in \Xi_f$, $\Xi f(a) = 2f(a)$ and that if $a, b \in \Xi_f$ and $(a,b) \cap \Xi_f = \emptyset$, then $f(b) = -f(a)$.

For $a \in \Xi_f$ and $0 < \delta < \delta_0$, let

$$f_\delta^{(a)}(z) = f(a) \left(\frac{1}{\delta}(z-a) 1_{|z-a|<\delta} + \operatorname{sgn}(z-a) 1_{|z-a| \geq \delta} \right).$$

Next, we introduce a family, indexed by $\delta \in (0, \delta_0)$, of continuous piecewise linear functions that approximate f:

$$f_\delta(z) = \begin{cases} f_\delta^{(a)}(z) & z \in (a-\delta, a+\delta),\ a \in \Xi_f \\ f(z) & z \notin \bigcup_{a \in \Xi_f} (a-\delta, a+\delta) \end{cases}$$

Lemma 4.2. *$\forall \delta \in (0,\delta_0)$, $\forall z \in \mathbb{R}$, $(f_\delta(z) - f(z))^2 \leq \sum_{a \in \Xi_f} 1_{(a-\delta, a+\delta)}(z)$. In particular, $\forall z \notin \Xi_f$, $\lim_{\delta \to 0} f_\delta(z) = f(z)$. Furthermore, $|f_\delta - f|$ is a continuous function.*

The proof is trivial, hence omitted.

Theorem 4.3. *If $\eta_k = f(X_{k-1}/\sqrt{k})\xi_k$ where the left-continuous function f has positive minimum gap, then $W^{(n)}$ converges weakly to the non-Gaussian process*

$$W = \left(B_t, \int_0^t f(B_s/\sqrt{s}) dB_s \right)_{t \in [0,1]}.$$

Proof. As already observed, the first and second components of the limiting process W, assuming it exists, are one-dimensional Brownian motions. However, the two-dimensional process is not a two-dimensional Brownian motion because the co-variation process

$$\left\langle B, \int_0^{\cdot} f(B_s/\sqrt{s}) dB_s \right\rangle_t = \int_0^t f(B_s/\sqrt{s}) ds \tag{4.1}$$

is not a constant times t, as it would be for a two-dimensional Brownian motion.

By the Skorokhod representation theorem, the convergence of $U^{(n)}$ to a Brownian motion can be assumed to be almost sure, uniformly on $[0,1]$. Call B the limiting process. We shall establish that for each $t \geq 0$,

$$\int_{(0,t]} f(U_{s-}^{(n)}/\sqrt{s}) dU_s^{(n)} \underset{n \uparrow \infty}{\overset{\text{prob.}}{\longrightarrow}} \int_0^t f(B_s/\sqrt{s}) dB_s.$$

This will immediately imply the joint convergence in probability for any collection of times $t_1 < \ldots < t_d$, thus establishing that the limiting finite-dimensional distributions are those of the process $\left(B_t, \int_0^t f(B_s/\sqrt{s}) dB_s \right)$.

Fix $t \geq 0$ and $\varepsilon > 0$. By the L^2-isometry for stochastic integrals with respect to the Brownian motion,

$$\begin{aligned} & \mathbb{E}\left[\left(\int_0^t f_\delta(B_s/\sqrt{s}) dB_s - \int_0^t f(B_s/\sqrt{s}) dB_s \right)^2 \right] \\ = \;& \mathbb{E}\left[\int_0^t \left(f_\delta(B_s/\sqrt{s}) - f(B_s/\sqrt{s}) \right)^2 ds \right] = \int_0^t \mathbb{E}\left[\left(f_\delta(B_s/\sqrt{s}) - f(B_s/\sqrt{s}) \right)^2 \right] ds \end{aligned}$$

where the right member goes to zero as $\delta \to 0$, by the bounded convergence theorem and the fact that for each $s \in (0,t]$, $\mathbb{P}(B_s/\sqrt{s} \in \Xi_f) = 0$ so that $f_\delta(B_s/\sqrt{s}) \longrightarrow f(B_s/\sqrt{s})$ a.s. as $\delta \to 0$. Since convergence in L^2 implies convergence in probability, there is $\delta_{\varepsilon,1} \in (0,\delta_0)$ such that for all $0 < \delta \le \delta_{\varepsilon,1}$,

$$\mathbb{P}\left(\left|\int_0^t f_\delta(B_s/\sqrt{s})dB_s - \int_0^t f(B_s/\sqrt{s})dB_s\right| > \varepsilon\right) < \varepsilon. \tag{4.2}$$

Since almost surely, the amount of time that the Brownian motion B spends in Ξ_f has Lebesgue measure zero, there is $\delta_{\epsilon,2} \in (0,\delta_0)$ such that for all $0 < \delta \le \delta_{\epsilon,2}$

$$\sum_{a\in\Xi_f} \mathbb{E}\left[\int_0^t 1_{(a-\delta,a+\delta)}(B_s/\sqrt{s})ds\right] < \varepsilon^3. \tag{4.3}$$

Let $\delta^* = \min(\delta_{\varepsilon,1},\delta_{\varepsilon,2})$. By the continuity of f_{δ^*}, almost surely, $f_{\delta^*}(U_s^{(n)}/\sqrt{s})$ converges uniformly on $[0,t]$ to $f_{\delta^*}(B_s/\sqrt{s})$. Thus, almost surely, $(U_s^{(n)}, f_{\delta^*}(U_s^{(n)}/\sqrt{s}))$ converges uniformly on $[0,t]$ to $(B_s, f_{\delta^*}(B_s/\sqrt{s}))$. Since $U^{(n)}$ is a martingale and has uniformly bounded jumps (recall that $|\Delta U_s^{(n)}| \le 1/\sqrt{n}$), it immediately follows from Theorem 2.2 of [12] that $\int_{(0,t]} f_{\delta^*}(U_{s-}^{(n)}/\sqrt{s})dU_s^{(n)}$ converges in probability to $\int_0^t f_{\delta^*}(B_s/\sqrt{s})dB_s$ as $n \to \infty$.

Let $n_\epsilon > 0$ such that for all $n \ge n_\epsilon$,

$$\mathbb{P}\left(\left|\int_{(0,t]} f_{\delta^*}(U_{s-}^{(n)}/\sqrt{s})dU_s^{(n)} - \int_0^t f_{\delta^*}(B_s/\sqrt{s})dB_s\right| > \varepsilon\right) < \varepsilon. \tag{4.4}$$

Combining the above we have that for all $n \ge n_\epsilon$,

$$\begin{aligned}
&\mathbb{P}\left(\left|\int_{(0,t]} f(U_{s-}^{(n)}/\sqrt{s})dU_s^{(n)} - \int_0^t f(B_s/\sqrt{s})dB_s\right| > 3\varepsilon\right) \\
&\le \mathbb{P}\left(\left|\int_{(0,t]} \left(f(U_{s-}^{(n)}/\sqrt{s}) - f_{\delta^*}(U_{s-}^{(n)}/\sqrt{s})\right) dU_s^{(n)}\right| > \varepsilon\right) \\
&\quad + \mathbb{P}\left(\left|\int_{(0,t]} f_{\delta^*}(U_{s-}^{(n)}/\sqrt{s})dU_s^{(n)} - \int_0^t f_{\delta^*}(B_s/\sqrt{s})dB_s\right| > \varepsilon\right) \\
&\quad + \mathbb{P}\left(\left|\int_0^t f_{\delta^*}(B_s/\sqrt{s})dB_s - \int_0^t f(B_s/\sqrt{s})dB_s\right| > \varepsilon\right) \\
&\le \frac{1}{\varepsilon^2}\mathbb{E}\left[\left|\int_{(0,t]} \left(f(U_{s-}^{(n)}/\sqrt{s}) - f_{\delta^*}(U_{s-}^{(n)}/\sqrt{s})\right) dU_s^{(n)}\right|^2\right] + 2\varepsilon \\
&= \frac{1}{\varepsilon^2}\mathbb{E}\left[\int_{(0,t]} \left|f(U_{s-}^{(n)}/\sqrt{s}) - f_{\delta^*}(U_{s-}^{(n)}/\sqrt{s})\right|^2 d[U^{(n)},U^{(n)}]_s\right] + 2\varepsilon
\end{aligned}$$

where we used Markov's inequality, (4.2) and (4.4) for the second inequality, and the Itô isometry for the equality. Recall that

$$[U^{(n)},U^{(n)}]_t = \sum_{0<s\le t} (\Delta U_s^{(n)})^2 = \sum_{k=1}^{\lfloor nt\rfloor} (\xi_k/\sqrt{n})^2 = \frac{\lfloor nt\rfloor}{n}$$

and denote by g the function $(f - f_{\delta^*})^2$. Then

$$\begin{aligned}\int_{(0,t]} g(U^{(n)}_{s-}/\sqrt{s})d[U^{(n)},U^{(n)}]_s &= \sum_{0<s\le t} g(U^{(n)}_{s-}/\sqrt{s})(\Delta U^{(n)}_s)^2 = \frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor} g(X_{k-1}/\sqrt{k}) \\ &= \int_0^{\lfloor nt\rfloor/n} g\left(\left(U^{(n)}_{s-}/\sqrt{s}\right)\sqrt{ns/(\lfloor ns\rfloor+1)}\right)ds \\ &\le \int_0^t g\left(\left(U^{(n)}_{s-}/\sqrt{s}\right)\sqrt{ns/(\lfloor ns\rfloor+1)}\right)ds\end{aligned}$$

and

$$\begin{aligned}&\mathbb{P}\left(\left|\int_{(0,t]} f(U^{(n)}_{s-}/\sqrt{s})dU^{(n)}_s - \int_0^t f(B_s/\sqrt{s})dB_s\right| > 3\varepsilon\right) \\ \le\ & \frac{1}{\epsilon^2}\mathbb{E}\left[\int_0^t g\left(\left(U^{(n)}_{s-}/\sqrt{s}\right)\sqrt{ns/(\lfloor ns\rfloor+1)}\right)ds\right] + 2\varepsilon \qquad (4.5)\end{aligned}$$

By the continuity of g and the bounded convergence theorem, as $n \to \infty$,

$$\mathbb{E}\left[\int_0^t g\left(\left(U^{(n)}_{s-}/\sqrt{s}\right)\sqrt{ns/(\lfloor ns\rfloor+1)}\right)ds\right] \longrightarrow \mathbb{E}\left[\int_0^t g(B_s/\sqrt{s})ds\right]. \qquad (4.6)$$

The right-hand side can further be evaluated:

$$\mathbb{E}\left[\int_0^t g(B_s/\sqrt{s})ds\right] \le \sum_{a\in\Xi_f}\mathbb{E}\left[\int_0^t 1_{(a-\delta^*,a+\delta^*)}(B_s/\sqrt{s})ds\right] < \varepsilon^3,$$

where we used Lemma 4.2 for the first inequality and (4.3) for the second.

It follows that there is an $n'_\varepsilon \ge n_\varepsilon$ such that for all $n \ge n'_\varepsilon$,

$$\mathbb{E}\left[\int_0^t g\left(\left(U^{(n)}_{s-}/\sqrt{s}\right)\sqrt{ns/(\lfloor ns\rfloor+1)}\right)ds\right] \le 2\varepsilon^3. \qquad (4.7)$$

Substituting this in (4.5), we find that for all $n \ge n'_\varepsilon$,

$$\mathbb{P}\left(\left|\int_{(0,t]} f(U^{(n)}_{s-}/\sqrt{s})dU^{(n)}_s - \int_0^t f(B_s/\sqrt{s})dB_s\right| > 3\varepsilon\right) \le 2\varepsilon + 2\varepsilon = 4\varepsilon. \qquad (4.8)$$

Since $\varepsilon > 0$ was arbitrary, the desired convergence in probability follows. □

Corollary 4.4. *Suppose* $V^{(n)}_t = \frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt\rfloor}\mathrm{sgn}(X_{k-1})\xi_k$*, where* $\mathrm{sgn}(0) = -1$*. Then* $W^{(n)}$ *converges weakly to the two-dimensional non-Gaussian process*

$$\left(B_t, \int_0^t \mathrm{sgn}(B_s)dB_s\right)_{t\in[0,1]}.$$

5 Ergodicity of the GBRW transformation – a characterisation

In this section we study the ergodicity of the measure-preserving transformation $(X_n)_{n\ge0} \to (Y_n)_{n\ge0}$. More specifically, let Ω be the space of integer-valued sequences $x = (x_n)_{n\ge0}$ such that $x_0 = 0$ and for any $n \ge 1$, $x_n - x_{n-1} \in \{-1,+1\}$, endowed with the product sigma-field and the probability measure under which the coordinate map is a simple symmetric random walk started at 0. Given a sequence of functions

$\psi_n : \{-1,+1\}^n \to \{-1,+1\}$ (ψ_0 is a constant taking value -1 or $+1$), we let T be the transformation on Ω such that, with $y = T(x)$, $y_n = \sum_{k=1}^n v_k$, where $v_k = \psi_{k-1}(u_1,\dots,u_{k-1})u_k$ and $u_k = x_k - x_{k-1}$. We have seen that T is measure-preserving. Here we ask whether T is ergodic.

We have seen that the following discrete-time version of the Lévy transformation

$$\psi_{k-1}(x_1,\dots,x_{k-1}) = \text{sgn}(x_1 + \ldots + x_{k-1})$$

is not ergodic (see [8]). Dubins and Smorodinsky [5] modified this transformation by first letting sgn$(0) = 0$ and then skipping any flat portions of the path thus produced, and showed that such a transformation is ergodic. While such a map is measure-preserving, it does not satisfy the non-anticipative property enjoyed by the general bootstrap random walk. Consequently, the two settings are distinct and the results in one cannot translate to results in the other. Next we give a necessary and sufficient condition, in terms of the sequence $(\psi_n)_{n\geq 0}$, for the GBRW to be ergodic.

We start by observing that the ergodicity of T is equivalent to that of $\tau = \Delta \circ T \circ \Delta^{-1}$ defined on the space $\Theta = \{-1,+1\}^{\mathbb{N}}$, where $\Delta : \Omega \to \Theta$ is the difference operator; that is, for $x \in \Omega$, $u = \Delta(x)$ is the sequence in $\{-1,+1\}^{\mathbb{N}}$ defined by $u_k = x_k - x_{k-1}$ ($u_1 = x_1$). The measure that τ preserves is the Bernoulli measure μ defined as

$$\mu\big(\{u \in \Theta :\ u_1 = \varepsilon_1,\dots,u_n = \varepsilon_n\}\big) = \frac{1}{2^n},$$

for any $\varepsilon_1,\dots,\varepsilon_n \in \{-1,+1\}$. For each $n \geq 1$, we let μ_n be the measure induced from μ by the projection π_n on Θ_n: $\mu_n(A_n) = \mu(\pi_n^{-1}(A_n))$. We also let τ_n be the mapping $\tau_n : (u_1,\dots,u_n) \to (v_1,\dots,v_n)$ such that $v_k = \psi_{k-1}(u_1,\dots,u_{k-1})u_k$. We observe that $\tau_n \circ \pi_n = \pi_n \circ \tau$, that τ and τ_n are bijections, and deduce that

$$\tau_n^{-1}(\pi_n(A)) = \tau_n^{-1}(\pi_n(\tau(\tau^{-1}(A)))) = \tau_n^{-1}(\tau_n(\pi_n(\tau^{-1}(A)))) = \pi_n(\tau^{-1}(A)).$$

Lemma 5.1. *The transformation τ is ergodic if and only if, for any $n \geq 1$, whenever $\tau_n^{-1}(A_n) = A_n$, we must have $\mu_n(A_n) = 0$ or $\mu_n(A_n) = 1$.*

Proof. Let A be be such that $\tau^{-1}(A) = A$. Then we must have that, for any $n \geq 1$, $\tau_n^{-1}(\pi_n(A)) = \pi_n(\tau^{-1}(A)) = \pi_n(A)$. Since $\pi_{n+1}^{-1}(\pi_{n+1}(A)) = \pi_{n+1}(A) \times \Theta \subset \pi_n(A) \times \Theta = \pi_n^{-1}(\pi_n(A))$, the sequence $(\mu_n(\pi_n(A)))_{n\geq 1}$ is non-increasing and must be constant, equal to 0 or 1, for n large enough. As $A = \lim_n \pi_n(A) \times \Theta$, we must have that $\mu(A) = 0$ or $\mu(A) = 1$, and that τ must be ergodic.

Conversely, if τ is ergodic and for $n \geq 1$, A_n is such that $\tau_n^{-1}(A_n) = A_n$, then

$$\tau^{-1}(\pi_n^{-1}(A_n)) = \pi_n^{-1}(\tau_n^{-1}(A_n)) = \tau_n^{-1}(A_n) \times \Theta = A_n \times \Theta = \pi_n^{-1}(A_n).$$

It follows that $A_n \times \Theta = \pi_n^{-1}(A_n)$ has measure 0 or 1. It immediately follows that A_n itself has measure 0 or 1. ☐

We are now ready to state the main result of this section.

Theorem 5.2. *Let T be the measure-preserving transformation associated with a sequence of functions $(\psi_n)_{n\geq 0}$. T is ergodic (i.e. τ is ergodic) if and only if either of the following equivalent statements holds true*

1. $\psi_0 = -1$ *and* $\forall n \geq 1$, $\prod_{(u_1,\dots,u_n)\in\Theta_n} \psi_n(u_1,\dots,u_n) = -1$,
2. $\psi_0 = -1$ *and* $\forall n \geq 1$, $\max(u_1,\dots,u_n)$ *appears in representation* (2.4)*; that is* $\beta_{n+1,\{1,\dots,n\}} = 1$.

Proof. The key idea in the proof is the fact τ is ergodic if and only if, for any $n \geq 1$, orbits of τ_n have period 2^n as otherwise a size $m < 2^n$ orbit O defines a proper subset of Θ_n such that $\tau_n^{-1}(O) = O$. For $n = 1$, we clearly require that $\psi_0 = -1$. Next we build τ_2 by constructing the orbit started at $(-1,-1)$: $(-1,-1) \to (+1, -\psi_1(-1)) \to (-1, -\psi_1(-1)\psi_1(+1)) \to (+1, -\psi_1(+1))$. This orbit is of period 2^2 if and only if $\psi_1(-1)\psi_1(+1) = -1$. We reason by induction and assume that each of $\tau_1, \ldots, \tau_n$ has a single orbit. Let θ_1 be the vector in Θ_n made up of n (-1)'s and $(\theta_1, \ldots, \theta_{2^n})$ be the orbit of τ_n started at θ_1. Note that $\tau_n(\theta_{2^n}) = \theta_1$. Let ϑ_1 be the vector in Θ_{n+1} made up of $n+1$ (-1)'s. Then $\vartheta_2 = (\theta_2, -\psi_n(\theta_1))$, $\vartheta_3 = (\theta_3, -\psi_n(\theta_1)\psi_n(\theta_2))$,$\ldots$,$\vartheta_{2^n} = \left(\theta_{2^n}, -\prod_{k=1}^{2^n-1}\psi_n(\theta_k)\right)$ and $\vartheta_{2^n+1} = \left(\theta_1, -\prod_{k=1}^{2^n}\psi_n(\theta_k)\right)$. The requirement that $\vartheta_{2^n+1} \neq \vartheta_1$ translates to

$$\prod_{(u_1,\ldots,u_n)\in\Theta_n} \psi_n(u_1,\ldots,u_n) = \prod_{k=1}^{2^n} \psi_n(\theta_k) = -1,$$

which establishes the first statement. The second statement is a direct application of representation (2.2):

$$\begin{aligned}
-1 &= \prod_{\Theta_n} \psi_n(u_1,\ldots,u_n) = \prod_{\Theta_n}\prod_{K\in\mathbb{K}(n)} u_{[K]}^{\beta_K} = \prod_{K\in\mathbb{K}(n)}\prod_{\Theta_n} u_{[K]}^{\beta_{n+1,K}} \\
&= \prod_{K\in\mathbb{K}(n)} \left(\prod_{\Theta_n} u_{[K]}\right)^{\beta_{n+1,K}} = \prod_{K\in\mathbb{K}(n)} \left((-1)^{2^{n-|K|}}\right)^{\beta_{n+1,K}} = (-1)^{\beta_{n+1,\{1,\ldots,n\}}}. \quad \square
\end{aligned}$$

We observe that to each transformation τ on Θ (T on Ω) corresponds an ergodic transformation $\tilde{\tau}$ on Θ ($\tilde{T}$ on Ω) such that these transformations time-shifted (and space-shifted) coincide with high probability; that is

$$\mu\big(\{u:\ \omega_N(\tau(u)) = \omega_N(\tilde{\tau}(u))\}\big) \geq 1 - 2^{-N},$$

where $\omega_N(u) = (u_{n+N})_{n\geq 1}$. The transformation $\tilde{\tau}$ is simply built from τ by changing the value of $\beta_{n+1,\{1,\ldots,n\}}$ to 1 whenever required.

An ergodic (slightly) modified discrete Lévy transformation.

As already pointed out the discrete Lévy transformation is not ergodic but a modification that skips flat portions of the path created by an adjustment to the definition of the sgn function ($\mathrm{sgn}(0) = 0$) is ergodic. Here we suggest another (simpler) modification of the discrete Lévy transformation.

We start by investigating representation (2.2) for $\mathrm{sgn}(u_1 + \ldots + u_n)$. As a latter is a symmetric function, the $\beta_{n,K}$'s must be equal for all subsets of equal size; that is, with a slight modification of the notation, writing $\beta_{n,|K|}$ for $\beta_{n,K}$,

$$\mathrm{sgn}(u_1 + \ldots + u_n) = \prod_{K\in\mathbb{K}(n)} u_{[K]}^{\beta_{n,K}} = \prod_{k=0}^{n}\left(\prod_{|K|=k} u_{[K]}\right)^{\beta_{n,k}}.$$

Recall that our definition of the sgn function assumes that $\mathrm{sgn}(0) = -1$. Now suppose $n \geq 2m+1$ and let $(u_1,\ldots,u_n)$ be such that $\nu_n = |\{k : u_k = -1\}| = \ell$ for $\ell \leq m$, and $s_n = u_1 + \ldots + u_n = n - 2\nu_n \geq 1$. Letting $\ell = 0$ yields

$$1 = \prod_{k=0}^{n}\left(\prod_{|K|=k} u_{[K]}\right)^{\beta_{n,k}} = (-1)^{\beta_{n,0}};\ \text{that is } \beta_{n,0} = 0.$$

For $\ell = 1$, $\prod_{|K|=1} u_{[K]} = -1$ and, for $k \geq 2$, $\prod_{|K|=k} u_{[K]} = 1$. Therefore for such $(u_1, \ldots, u_n)$,

$$1 = \left(\left(\prod_{|K|=1} u_{[K]} \right)^{\beta_{n,1}} \right) \left(\prod_{k=2}^{n} \left(\prod_{|K|=k} u_{[K]} \right)^{\beta_{n,k}} \right) = (-1)^{\beta_{n,1}}; \text{ that is } \beta_{n,1} = 0.$$

This can be repeated and yields the fact that $\beta_{n,m} = 0$ whenever $n \geq 2m+1$.

For $m \leq n \leq 2m$, again choosing $(u_1, \ldots, u_n)$ such that $\nu_n = |\{k : u_k = -1\}| = m$, and $s_n = n - 2m \leq 0$, yields

$$-1 = \left(\prod_{k=0}^{\ell_n} \left(\prod_{|K|=k} u_{[K]} \right)^{\beta_{n,k}} \right) \left(\prod_{k=\ell_n+1}^{m} \left(\prod_{|K|=k} u_{[K]} \right)^{\beta_{n,k}} \right) \left(\prod_{k=m+1}^{n} \left(\prod_{|K|=k} u_{[K]} \right)^{\beta_{n,k}} \right),$$

where $\ell_n = \lfloor (n-1)/2 \rfloor$. Since $\beta_{n,k} = 0$ for $k \leq \ell_n$ and $\prod_{|K|=k} u_{[K]} = 1$ for $k \geq m+1$, the above identity becomes

$$-1 = \prod_{k=\ell_n+1}^{m} \left(\prod_{|K|=k} u_{[K]} \right)^{\beta_{n,k}} = \prod_{k=\ell_n+1}^{m} \left((-1)^{\binom{m}{k}} \right)^{\beta_{n,k}} = (-1)^{\sum_{k=\ell_n+1}^{m} \binom{m}{k} \beta_{n,k}};$$

that is,

$$\sum_{k=\ell_n+1}^{m} \binom{m}{k} \beta_{n,k} = 1 \mod 2 \text{ or equivalently } \beta_{n,m} = 1 + \sum_{k=\ell_n+1}^{m-1} \binom{m}{k} \beta_{n,k} \mod 2.$$

The figure below describes the array $\beta_{n,k}$ where an orange cell indicates that $\beta_{n,k} = 0$ and a blue cell that $\beta_{n,k} = 1$.

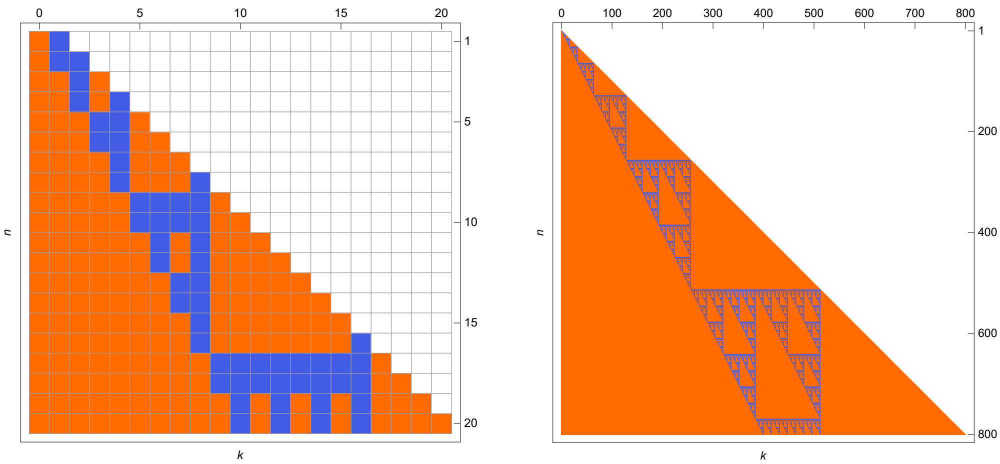

Figure 1: The $\beta_{n,k}$ array for $n \leq 20$ (left) and $n \leq 800$ (right)

For example,

$$\operatorname{sgn}(u_1+u_2) = u_1 u_2 \max(u_1, u_2) \text{ and } \operatorname{sgn}(u_1+u_2+u_3) = \max(u_1, u_2)\max(u_1, u_3)\max(u_2, u_3).$$

Corollary 5.3. *The following slight adaptation of the discrete Lévy transformation is ergodic:* $\psi_0 = -1$ *and for* $n \geq 1$, $\psi_n(u_1, \ldots, u_n) = \operatorname{sgn}_n(u_1 + \ldots + u_n)$, *where*

$$\operatorname{sgn}_n(s) = \begin{cases} \operatorname{sgn}(s) & n \text{ is a power of 2 or } s > -n \\ 1 & n \text{ is not a power of 2 and } s \leq -n \end{cases}$$

or equivalently that $\psi_0 = -1$ *and for* $n \geq 1$*,*

$$\psi_n(u_1,\ldots,u_n) = \begin{cases} \mathrm{sgn}(u_1+\ldots+u_n) & n \text{ is a power of 2} \\ \max(u_1,\ldots,u_n)\mathrm{sgn}(u_1+\ldots+u_n) & n \text{ is not a power of 2} \end{cases}$$

Proof. The proof relies on Theorem 5.2 and the ergodicity criterion thereof: $\beta_{n+1,\{1,\ldots,n\}} = 1$.

The case $n=1$ is trivial. Fix $n \geq 2$,

$$\prod_{(u_1,\ldots,u_n)\in\Theta_n} \mathrm{sgn}(u_1+\ldots+u_n) = \prod_{k=0}^{n} \left(\mathrm{sgn}(2k-n)\right)^{\binom{n}{k}} = \begin{cases} (-1)^{\sum_{k=0}^{(n-1)/2}\binom{n}{k}} & n \text{ odd} \\ (-1)^{\sum_{k=0}^{n/2}\binom{n}{k}} & n \text{ even} \end{cases}$$

However, for n odd, $\sum_{k=0}^{(n-1)/2}\binom{n}{k} = \sum_{k=(n+1)/2}^{n}\binom{n}{k} = 2^{n-1}$ is even. And for n even, $\sum_{k=0}^{n/2}\binom{n}{k}$ is even if and only if $\frac{1}{2}\binom{n}{n/2}$ is even. We use Lucas' Theorem [13] to show that $\frac{1}{2}\binom{2n}{n} = \binom{2n-1}{n-1}$ is odd if and only if n is a power of 2.

Suppose that $n = 2^\ell$ for $\ell \in \mathbb{N}$. Then

$$n-1 = \sum_{k=0}^{\ell-1} 2^k \text{ and } 2n-1 = \sum_{k=0}^{\ell} 2^k$$

and every digit in the base 2 expansion of $n-1$ is less than or equal to the corresponding digit in the base 2 expansion of $2n-1$, thus proving that $\binom{2n-1}{n-1}$ is not divisible by 2.

Now, suppose that $2^\ell < n < 2^{\ell+1}$ for $\ell \in \mathbb{N}$ so that $n = 2^\ell + \sum_{k=\kappa+1}^{\ell-1}\alpha_k 2^k + 2^\kappa$, for $\alpha_{\kappa+1},\ldots,\alpha_{\ell-1} \in \{0,1\}$ and $0 \leq \kappa < \ell$.

If $\kappa > 0$, then

$$n-1 = 2^\ell + \sum_{k=\kappa+1}^{\ell-1}\alpha_k 2^k + \sum_{k=0}^{\kappa-1} 2^k \text{ and } 2n-1 = 2^{\ell+1} + \sum_{k=\kappa+1}^{\ell-1}\alpha_k 2^{k+1} + \sum_{k=0}^{\kappa} 2^k$$

so that the digits in the base 2 expansions of $n-1$ and $2n-1$ are:

k	$\ell+1$	ℓ	$\ell-1$	$\ldots$	$\kappa+2$	$\kappa+1$	κ	$\kappa-1$	$\ldots$	0
$2n-1$	1	$\alpha_{\ell-1}$	$\alpha_{\ell-2}$	$\ldots$	$\alpha_{\kappa+1}$	0	1	1	$\ldots$	1
$n-1$	0	1	$\alpha_{\ell-1}$	$\ldots$	$\alpha_{\kappa+2}$	$\alpha_{\kappa+1}$	0	1	$\ldots$	1

If $\alpha_{\kappa+1} = \ldots = \alpha_{\ell-1} = 1$, then the digit of order $\kappa+1$ in the base 2 expansion of $n-1$ is greater than the corresponding digit in the base 2 expansion of $2n-1$, and $\binom{2n-1}{n-1}$ is divisible by 2.

Suppose at least one of $\alpha_{\kappa+1},\ldots,\alpha_{\ell-1}$ equals 0 and let $k^* = \max\{k \in \{\kappa+1,\ldots,\ell-1\} : \alpha_k = 0\}$. Then the digit of order k^*+1 in the base 2 expansion of $n-1$ is greater than the corresponding digit in the base 2 expansion of $2n-1$, and $\binom{2n-1}{n-1}$ is divisible by 2.

If $\kappa = 0$, then

$$n-1 = 2^\ell + \sum_{k=1}^{\ell-1}\alpha_k 2^k \text{ and } 2n-1 = 2^{\ell+1} + \sum_{k=1}^{\ell-1}\alpha_k 2^{k+1} + 1$$

so that the digits in the base 2 expansions of $n-1$ and $2n-1$ are:

k	$\ell+1$	ℓ	$\ell-1$	$\ldots$	2	1	0
$2n-1$	1	$\alpha_{\ell-1}$	$\alpha_{\ell-2}$	$\ldots$	α_1	0	1
$n-1$	0	1	$\alpha_{\ell-1}$	$\ldots$	α_2	α_1	0

The same argument as before shows that $\binom{2n-1}{n-1}$ is divisible by 2. □

Another example for which $\beta_{n+1,\{1,\dots,n\}} = 1$ and ergodicity holds, is the following modification of the discrete Lévy transformation:

$$\psi_n(u_1,\dots,u_n) = \max(u_1,\dots,u_n)\mathrm{sgn}(u_1+\dots+u_{n-1}).$$

Appendix

Lemma 5.4. *Let $(M_k)_{k\geq 1}$ be a sequence of subsets of $\mathbb{N}$. Suppose that the cardinality of M_k is finite and constant. If $\limsup_n M_n = \emptyset$ then*

$$\lim_n \frac{1}{n}\sum_{k=1}^{n} |M_k \cap M_{n+1}| = 0.$$

Proof. We prove the contrapositive statement that if $\lim_n \frac{1}{n}\sum_{k=1}^{n} |M_k \cap M_{n+1}| > 0$, then $\limsup_n M_n \neq \emptyset$.

We label, say in the increasing order, the elements of M_k, $u_k^{(1)},\dots,u_k^{(m)}$, where $m = |M_k|$. Then

$$M_k \cap M_{n+1} = \left(\bigcup_{i=1}^{m}\{u_k^{(i)}\}\right) \cap \left(\bigcup_{j=1}^{m}\{u_{n+1}^{(i)}\}\right) = \bigcup_{i,j=1}^{m}\left(\{u_k^{(i)}\}\cap\{u_{n+1}^{(j)}\}\right)$$

and

$$\sum_{i,j=1}^{m} \lim_n \frac{1}{n}\sum_{k=1}^{n} |\{u_k^{(i)}\}\cap\{u_{n+1}^{(j)}\}| = \lim_n \frac{1}{n}\sum_{k=1}^{n} |M_k \cap M_{n+1}| > 0,$$

from which we deduce that for at least one pair (i,j), $\lim_n \frac{1}{n}\sum_{k=1}^{n} \delta(u_k^{(i)}, u_{n+1}^{(j)}) > 0$. Here $\delta(k,\ell)$ denotes the Kronecker delta function: $\delta(k,\ell) = 1$ if $k=\ell$ and 0 otherwise.

Let $A_n = \{k \leq n; u_k^{(i)} = u_{n+1}^{(j)}\}$. Then for such a pair (i,j), for $\varepsilon > 0$ and n large enough

$$\frac{1}{n}|A_n| = \frac{1}{n}\sum_{k=1}^{n} \delta(u_k^{(i)}, u_{n+1}^{(j)}) > \varepsilon; \text{ i.e. } |A_n| > n\varepsilon.$$

Fix such n after which this inequality is satisfied and let $N = \min\{\ell > n; u_\ell^{(j)} \notin \{u_1^{(i)},\dots,u_\ell^{(i)}\}\}$. If N is finite, then $A_{N-1} = \emptyset$ and does not satisfy the requirement that $|A_{N-1}| > (N-1)\varepsilon$. It follows that N is infinite, that at least one integer in $\{u_1^{(i)},\dots,u_n^{(i)}\}$ is repeated infinitely many times and that $\limsup_n M_n \neq \emptyset$. ☐

References

[1] Bayraktar, E., Dolinsky, Y. & Guo, J. (2020) *Continuity of utility maximization under weak convergence*, Math. Finan. Econ. 14, 725–757. MR4135703

[2] Bertoin J. (2020) *Counterbalancing steps at random in a random walk.* https://arxiv.org/pdf/2011.14069.pdf

[3] Collevecchio A., Hamza K. & Shi M. (2016) *Bootstrap random walks*, Stochastic Processes and their Applications, Volume 126, Issue 6, 1744–1760. MR3483735

[4] Collevecchio A., Hamza K. & Liu Y. (2019) *Invariance principle for biased bootstrap random walks*, Stochastic Processes and their Applications, Volume 129, Issue 3, 860–877. MR3913271

[5] Dubins L.E. & Smorodinsky M. (1992) *The Modified, Discrete, Lévy-Transformation Is Bernoulli*, Séminaire de Probabilités, XXVI, Lecture Notes in Math., 1526, 157–161. MR1231991

[6] Engländer, J. & Volkov, S. (2018) *Turning a coin over instead of tossing it.* Journal of Theoretical Probability volume 31, 1097–1118. MR3803925

[7] Engländer, J., Volkov, S. & Wang, Z. (2021) *The coin-turning walk and its scaling limit.* Electron. J. Probab. 25, 1–38. MR4053903

[8] Fujita T. (2008) *A random walk analogue of Lévy's Theorem*, Studia Sci. Math. Hungar. 45, no. 2, 223–233. MR2417970

[9] Gut A. & Stadtmüller U. *Variations of the elephant random walk* https://arxiv.org/abs/1812.01915

[10] Jacod J. & Shiryaev A. (1987) *Limit Theorems for Stochastic Processes. Springer Science & Business Media*, Springer Berlin Heidelberg MR0959133

[11] Kubota N. & Takei M. (2019) *Gaussian Fluctuation for Superdiffusive Elephant Random Walks*, Journal of Statistical Physics. 177, pages 1157–1171. MR4034803

[12] Kurtz T. G. & Protter P. (1991), *Weak limit theorems for stochastic integrals and stochastic differential equations*, Annals of Probability, 19, 1035–1070. MR1112406

[13] Lucas E. (1878) *Théorie des Fonctions Numériques Simplement Périodiques*, American Journal of Mathematics, Vol.1 No.3 197–240. MR1505164

[14] Mansuy, R. & Yor, M. (2005) Random Times and Enlargements of Filtrations in a Brownian Setting. Lecture Notes in Mathematics (Springer, volume 1873). MR2200733

[15] Prigent, M. & Roberts, M.I. (2020) *Noise sensitivity and exceptional times of transience for a simple symmetric random walk in one dimension*, Probab. Theory Relat. Fields 178, 327–367. MR4146540

[16] Prokaj V. (2012) *Some Sufficient Conditions for the Ergodicity of the Lévy-Transformation*, Séminaire de Probabilités, XLV, 93–121. MR3185910

[17] Schütz, G.M. & Trimper, S. (2004) *Elephants can always remember: Exact long-range memory effects in a non-Markovian random walk*, Phys. Rev. E 70, no. 4, (045101)1–4.

[18] Williams D. (1991) *Probability with Martingales. Cambridge University Press.* MR1155402

Acknowledgments. The authors are indebted to two anonymous referees whose suggestions and comments greatly improved the presentation and flow of the paper. RJ Williams thanks the Center for Modeling of Stochastic Systems at Monash University for their generous hospitality during her sabbatical visit there.