PREFERENTIAL ATTACHMENT WHEN STABLE

SVANTE JANSON,* *Uppsala University*

SUBHABRATA SEN,** *Massachusetts Institute of Technology*

JOEL SPENCER,*** *New York University*

Abstract

We study an urn process with two urns, initialized with a ball each. Balls are added sequentially, the urn being chosen independently with probability proportional to the αth power ($\alpha > 1$) of the existing number of balls. We study the (rare) event that the urn compositions are balanced after the addition of $2n - 2$ new balls. We derive precise asymptotics of the probability of this event by embedding the process in continuous time. Quite surprisingly, fine control of this probability may be leveraged to derive a lower-tail large deviation principle (LDP) for $L = \sum_{i=1}^{n} (S_i^2/i^2)$, where $\{S_n : n \geq 0\}$ is a simple symmetric random walk started at zero. We provide an alternative proof of the LDP via coupling to Brownian motion, and subsequent derivation of the LDP for a continuous-time analog of L. Finally, we turn our attention back to the urn process conditioned to be balanced, and provide a functional limit law describing the trajectory of the urn process.

Keywords: Urn model; large deviations

2010 Mathematics Subject Classification: Primary 60F10; 60F17
Secondary 60C05

1. Model and summary of results

Consider two urns, each of which initially has one ball. Balls come sequentially. When there are i balls in the first urn and j balls in the second urn, the next ball is placed in the first urn with probability $i^\alpha/(i^\alpha + j^\alpha)$ and in the second urn with probability $j^\alpha/(i^\alpha + j^\alpha)$. Here α is a constant. In this work we shall consider only $\alpha > 1$, though we note that the case $\alpha = 1$ is the classic Pólya urn model [10, 23]; see Remark 1.1. Equivalently, we define a Markov chain $\{(X_n, Y_n) : n \geq 2\}$ with states $\mathbb{N} \times \mathbb{N}$ and initial state $(X_2, Y_2) = (1, 1)$, and transition probabilities

$$\mathbb{P}[(X_{n+1}, Y_{n+1}) = (i + 1, j) \mid (X_n, Y_n) = (i, j)] = \frac{i^\alpha}{i^\alpha + j^\alpha}, \tag{1.1}$$

$$\mathbb{P}[(X_{n+1}, Y_{n+1}) = (i, j + 1) \mid (X_n, Y_n) = (i, j)] = \frac{j^\alpha}{i^\alpha + j^\alpha}. \tag{1.2}$$

(We have chosen the notation such that $X_n + Y_n = n$; we thus start with $n = 2$.)

In this model the rich get richer. It is highly unstable, as demonstrated by Theorems 2.1 and 2.3 below. We examine the (rare) event that the urn populations remain stable. For definiteness we concentrate on the event that state (n, n) is reached, i.e. $(X_{2n}, Y_{2n}) = (n, n)$, which we denote BINGO (n, n).

This work has three interrelated goals. First, we find the asymptotic probability of the urn population remaining stable, more precisely that it reaches the state (n, n) (see Theorem 2.3). Second, we study the typical behavior of the urn process $\{(X_n, Y_n): n \geq 2\}$, conditional on the rare event that state (n, n) is reached. We find that the typical walk from $(1, 1)$ to (n, n) is given asymptotically by a distorted Brownian bridge (see Theorem 1.2 below) and locally, suitably scaled, by an Ornstein–Uhlenbeck process, stated in Theorem 1.4. The technique of continuous-time embedding, as described in Section 2, is crucially utilized in deriving these results. Finally, as an independent application of these ideas, we provide a lower-tail large deviation principle (LDP) for a quadratic functional of simple symmetric random walk, defined formally in (1.3).

For the convenience of the reader, we start with a high-level description of the connection of our urn process to the random walk statistic L defined below. Note that a sequence $\xi_1, \ldots, \xi_{2n-2} = \pm 1$ corresponds bijectively to a path **P** of length $2n - 2$ on $\mathbb{N} \times \mathbb{N}$ starting at $(1, 1)$, with ± 1 representing horizontal and vertical moves respectively. $\mathbb{P}[\textbf{P}]$ represents the probability the Markov chain follows the path **P**. $\mathbb{P}[\text{BINGO}(n, n)]$ is then the sum of $\mathbb{P}[\textbf{P}]$ over all paths from $(1, 1)$ to (n, n). The values $\mathbb{P}[\textbf{P}]$ vary considerably. The further **P** is from the main diagonal the lower is $\mathbb{P}[\textbf{P}]$. Set $S_t = \sum_{i=1}^{t} \xi_i$. S_t then measures how far the path is from the main diagonal after t steps. The exact relationship between the S_t and $\mathbb{P}[\textbf{P}]$ is given by a somewhat complex formula. Perhaps surprisingly, the asymptotic relationship is well captured by a single statistic, $L := \sum_{t=1}^{2n-2} t^{-2} S_t^2$. This relationship is captured in Section 4. $\mathbb{P}[\textbf{P}]$ is given by a term BASE, common to all paths, and an expression FIT, defined in (4.3) and (4.5). FIT, in turn, will be approximated by $\exp[-\lambda L]$, λ given by (4.2). $\mathbb{P}[\text{BINGO}(n, n)]$ is then given by the number of paths times BASE times the expected value of FIT. The expectation of FIT, critically, is approximated by the Laplace transform $\mathbb{E}[\exp[-\lambda L]]$. An approximation for $\mathbb{P}[\text{BINGO}(n, n)]$ thus provides tight estimates of the Laplace transform of L, which in turn governs the lower-tail large deviations of L, given by Theorem 1.1 below.

We return to a more formal presentation. Let $\xi_1, \ldots, \xi_n = \pm 1$, uniformly and independently, and set $S_t := \sum_{i=1}^{t} \xi_i$. That is, S_t is the position of the simple symmetric random walk at time t. (Note that we are not conditioning on $\sum \xi_i = 0$.) We set

$$L = L_n := \sum_{i=1}^{n} \frac{S_i^2}{i^2}. \tag{1.3}$$

As $\mathbb{E}[S_i^2] = i$,

$$\mathbb{E}[L_n] = \sum_{i=1}^{n} i^{-1} = \ln n + \mathrm{O}(1),$$

and, as will be seen later, L_n is typically about $\ln n$.

Our concern is with the lower tail of the distribution of L_n, and we prove the following. (See Theorem 8.4 for a corresponding result for the upper tail.)

Theorem 1.1. *For any fixed* $c \in (0, 1)$,

$$\mathbb{P}[L_n \leq c \ln n] = \mathrm{e}^{-(K(c)+\mathrm{o}(1)) \ln n}$$

with

$$K(c) = \frac{(1-c)^2}{8c}. \tag{1.4}$$

The connection between the urn process and the random walk statistic L, alluded to above, provided the original motivation for our investigations. In pursuing these investigations we have found L_n to have quite subtle properties.

In studying L_n it is helpful to parametrize $S_i = \sqrt{i} N_i$; the N_i are asymptotically (in i) standard Gaussian and $L = \sum N_i^2 / i$. The harmonic series suggests a logarithmic scaling, $t = \ln i$. Note that under this scaling we have strong correlation when t, t' are close, which fades as the distance increases. That is, $S_i, S_{i\lambda}$ are closely correlated when λ is close to one and have positive asymptotic correlation for any fixed λ, but that correlation approaches zero as λ approaches infinity.

We give two very different arguments for Theorem 1.1. In Sections 4–6 we employ the Markov process (X_n, Y_n) and the continuous-time argument for it in Section 2 to derive the Laplace transform of L, and from that deduce the large deviation Theorem 1.1. In Sections 7–8 we provide a more traditional proof, which turns out to be quite challenging. We couple the random walk to a standard Brownian motion via the celebrated KMT (Komlós–Major–Tusnády) coupling, and establish that it suffices to derive the corresponding LDP for a Brownian analog of L. Subsequently, we derive the LDP by the general theory for quadratic functionals of Brownian motion.

Theorem 1.1 establishes a rigorous lower-tail LDP for a quadratic functional of the simple symmetric random walk. Large deviations for non-linear functions of $\{\pm 1\}$ variables have been an active research area in recent years. In a breakthrough paper, Chatterjee and Dembo [7] initiated a systematic study of LDPs of non-linear functionals of $\{\pm 1\}$ variables. The theory was subsequently extended by Eldan [11], and has been applied to numerous problems in probability and combinatorics (see e.g. [3], [4], [12], [13], and [20]). We emphasize that Theorem 1.1 does not follow using the general theory established in these prior works, and that our approaches are entirely different. We elaborate more on this in Remark 8.2.

In Section 9 we return to the Markov process (X_k, Y_k) defined above, conditioned on the rare event BINGO (n, n), and examine the typical path from $(1, 1)$ to (n, n). We define $\Delta_k := X_k - Y_k$, so that

$$(X_k, Y_k) = \left(\frac{k + \Delta_k}{2}, \frac{k - \Delta_k}{2} \right), \quad k \geq 2, \tag{1.5}$$

and define, for completeness, $\Delta_0 = \Delta_1 := 0$. (For typographical reasons, we sometimes write $\Delta(k)$.) Note that the event BINGO (n, n) can be written $\Delta_{2n} = 0$. We provide a functional limit law which shows that conditioned on BINGO (n, n), Δ_k is typically of order $\sqrt{n}$ for $2 < k < 2n$, and that suitably rescaled, $(\Delta_k)_{k=2}^{2n}$ converges to a distorted Brownian bridge.

Theorem 1.2. *Let $G_\alpha(t)$, $t \in [0, 1]$, be the continuous Gaussian process with mean 0 and covariance function*

$$\text{Cov}(G_\alpha(s), G_\alpha(t)) = \frac{2}{2\alpha - 1} s^\alpha (t^{1-\alpha} - t^\alpha), \quad 0 \leq s \leq t \leq 1. \tag{1.6}$$

Then, as $n \to \infty$, conditioned on BINGO (n, n) *(i.e. $\Delta_{2n} = 0$),*

$$n^{-1/2} \Delta_{\lfloor 2nt \rfloor} \xrightarrow{d} G_\alpha(t), \quad t \in [0, 1], \tag{1.7}$$

in $\mathcal{D}[0, 1]$ with the Skorokhod topology.

$G_\alpha(t)$ can be constructed from a standard Brownian bridge $\text{Br}(t)$ as

$$G_\alpha(t) := (\alpha - 1/2)^{-1/2} t^{1-\alpha} \text{Br}(t^{2\alpha-1}), \tag{1.8}$$

with $G_\alpha(0) := 0$. Related constructions from a Brownian motion are given in (9.16)–(9.17).

We give also a version of this theorem for $k = \mathrm{o}(n)$. Now a Brownian motion $B(t)$ appears instead of a Brownian bridge.

Theorem 1.3. *Let $m_n \to \infty$ be real numbers with $m_n = \mathrm{o}(n)$. Then, as $n \to \infty$, conditioned on* BINGO (n, n) *(i.e. $\Delta_{2n} = 0$),*

$$m_n^{-1/2} \Delta_{\lfloor m_n t \rfloor} \xrightarrow{\text{d}} H_\alpha(t) := (2\alpha - 1)^{-1/2} t^{1-\alpha} B(t^{2\alpha-1}), \quad t \in [0, \infty), \tag{1.9}$$

in $\mathcal{D}[0, \infty)$ with the Skorokhod topology.

In particular, taking m_n integers and $t = 1$, it follows that for any integers $m = m_n \to \infty$ with $m = \mathrm{o}(n)$, conditioned on BINGO (n, n),

$$m^{-1/2} \Delta_m \xrightarrow{\text{d}} N\left(0, \frac{1}{2\alpha - 1}\right).$$

For $m < n$ with $m = \Theta(n)$, we obtain from Theorem 1.2 a similar result with a correction factor for the variance. Thus, Δ_m is typically of order $\sqrt{m}$ for $m < n$.

Under a suitable logarithmic scaling, we have in the limit a stationary Ornstein–Uhlenbeck process, defined by (1.11) below.

Theorem 1.4. *Fix any sequence t_n such that $t_n \to \infty$ and $\log n - \log t_n \to \infty$. Then we have, as $n \to \infty$, conditional on* BINGO (n, n),*

$$e^{-(s+t_n)/2} \Delta(\lfloor e^{s+t_n} \rfloor) \xrightarrow{\text{d}} Z(s), \quad -\infty < s < \infty, \tag{1.10}$$

in $\mathcal{D}(-\infty, \infty)$ with the Skorokhod topology, where $Z(s)$ is a centered Gaussian process with covariance function

$$\mathbb{E}[Z(s)Z(t)] = \frac{1}{2\alpha - 1} e^{-(\alpha - 1/2)|s-t|}, \quad s, t \in \mathbb{R}. \tag{1.11}$$

Remark 1.1. As said above, we consider in this paper only $\alpha > 1$, which is necessary for Theorem 2.1, for example. However, it would also be interesting to study $\alpha \in [0, 1]$, when (X_k, Y_k) behaves quite differently. Note that $\alpha = 0$ yields Δ_k as a simple symmetric random walk, and then it is well known that (1.7) holds with $G_0(t) := \sqrt{2}\text{Br}(t)$ (see e.g. [5, Theorem 24.1]). (The factor $\sqrt{2}$ is because of our choice of normalization.) Furthermore, for $\alpha = 1$, when, as mentioned above, (X_k, Y_k) is the classical Pólya urn, it is well known that the increments are exchangeable, and thus, conditioned on BINGO (n, n), all paths to (n, n) have the same probability. This can be seen from (4.5)–(4.7) below, noting that for $\alpha = 0$ or $\alpha = 1$, FIT_i in (4.3) is constant 1. Thus, conditioned on BINGO (n, n), $\alpha = 1$ and $\alpha = 0$ coincide, and therefore (1.7) holds for $\alpha = 1$ too, with $G_1(t) = G_0(t) = \sqrt{2}\text{Br}(t)$. Note further that this agrees with (1.6). However, for (1.8), it holds for $\alpha = 1$ but not for $\alpha = 0$. Similarly, (1.9) and (1.10)–(1.11) hold for $\alpha = 1$, with $H_1(t) = B(t)$. It would be interesting to find an analog of Theorem 1.2 for $0 < \alpha < 1$.

2. Continuous time

In this section we examine the Markov chain defined in Section 1 (1.1)–(1.2) with initial state $(1, 1)$.

Definition 2.1. BINGO (i, j) denotes the event that state (i, j) is reached.

In this section, we study asymptotic probabilities of various BINGO events. These probabilities will turn out to be extremely crucial in our proof of Theorem 1.1 in Sections 4, 5, and 6.

This preferential attachment model is best tackled (see Remark 2.1) via continuous time. Let V_i, W_i, $i \geq 1$, denote exponential distributions with rate parameter i^α. That is, V_i, W_i have probability density function $\lambda\, e^{-\lambda x}$ with $\lambda = i^\alpha$. The V_i, W_i are all chosen mutually independently. Begin the urn model, as before, with each urn having one ball. Begin time at zero. When an urn has i balls it waits time V_i until it receives its next ball. The forgetfulness property of the exponential distribution (plus a little calculus) gives that when the bins have i, j balls, respectively, the probability that $V_i < W_j$ is $i^\alpha / (i^\alpha + j^\alpha)$ as desired. This leads to a remarkable theorem ([8], but see Remark 2.1) with what is surely a Proof from The Book.

Theorem 2.1. *With probability* 1, *one of the bins gets all but a finite number of the balls.*

Proof. Let $V = \sum_{i=1}^{\infty} V_i$, $W = \sum_{i=1}^{\infty} W_i$. As $\sum i^{-\alpha}$ is finite (here using that $\alpha > 1$), both V and W are finite a.s. (almost surely). As the distribution is non-atomic, $V \neq W$ a.s. Say $V < W$. Then bin one receives all its balls before bin two does. When bin one has all its balls the process stops (a countable number of balls have been placed) and bin two has only a finite number of balls. $\qquad\square$

Corollary 2.1.

$$\lim_{M \to \infty} \lim_{k \to \infty} \sum_{i=M}^{k-M} \mathbb{P}[\text{BINGO}\,(i, k-i)] = 0. \qquad (2.1)$$

Proof. For M fixed let FENCE (k) denote the union of the BINGO $(i, k-i)$ over $M \leq i \leq k - M$; this is the event that at the time when there are k balls in the urns, there are at least M balls in each urn. As these BINGO $(i, k-i)$ are disjoint (a path can only hit one state with a given sum of coefficients), $\Pr[\text{FENCE}\,(k)]$ is given by the sum in (2.1). Further, FENCE (k) implies FENCE (k') for all $k' \geq k$, as once a path hits the fence at k it cannot escape the fence at k'. Thus the union of all FENCE (k) has probability $\lim_{k \to \infty}$ of the sum. But the union is the event that both bins eventually get at least M balls. From Theorem 2.1, this has limiting value (in M) of zero. $\qquad\square$

Remark 2.1. The use of continuous time appears to be due to Herman Rubin, as attributed by Burgess Davis in [8]. A thorough study of preferential attachment (in a far more general setting) via continuous time was given in the PhD thesis of Roberto Oliveira. Many of the results of Oliveira's thesis are given in [22]. Theorem 2.1 and Corollary 2.1 provided the original motivation for our current research. The senior author (JS) searched for a combinatorial proof, appropriately counting paths with their respective probabilities, for Corollary 2.1. This in turn led to attempts to estimate BINGO (i, j) without using continuous time. Somewhat surprisingly, one result is in the opposite direction. The estimates on BINGO (i, j) given by continuous time have given a fairly roundabout argument for the large deviation results for the random variable L given by Theorem 1.1.

Continuous time gives us excellent asymptotics on BINGO. We first provide the bridge between continuous time and BINGO.

Theorem 2.2. *Set*

$$\Xi := \sum_{s=1}^{i-1} V_i - \sum_{t=1}^{j-1} W_j.$$

Then BINGO (i, j) *occurs if and only if either* $0 \leq \Xi < W_j$ *or* $0 \leq -\Xi < V_i$.

Proof. Ξ is the time difference between when urn one receives its ith ball and urn two receives its jth ball. Suppose $\Xi \leq 0$. At time $T = \sum_{s=1}^{i-1} V_i$, urn one receives its ith ball. Urn two will receive its jth ball at time $T - \Xi$. BINGO (i, j) occurs when urn one has not yet received its $(i + 1)$th ball, which it does at time $T + V_i$. This occurs if and only if $V_i > -\Xi$. The case $\Xi \geq 0$ is similar. □

Theorem 2.3. *There is a positive constant* β, *dependent only on* α, *so that when* $i, j \to \infty$

$$\mathbb{P}[\text{BINGO}\, (i, j)] \sim \beta[i^{-\alpha} + j^{-\alpha}].$$

In particular,

$$\mathbb{P}[\text{BINGO}\, (n, n)] \sim 2\beta n^{-\alpha}. \tag{2.2}$$

Remark 2.2. Set $\Xi^{\dagger} := \sum_{i=1}^{\infty} (V_i - W_i)$. Basically Ξ is estimated by $\Xi^{\dagger}$ and β is the probability density function of $\Xi^{\dagger}$ at 0. W_j is almost always o(1) (as $j \to \infty$) so that $0 \leq \Xi < W_j$ should occur with asymptotic probability $\beta\mathbb{E}[W_j] = \beta j^{-\alpha}$. However, the validity of the approximation is non-trivial and has forced our somewhat technical calculations.

Proof of Theorem 2.3. We analyze $\mathbb{P}[0 \leq \Xi < W_j]$ as $i, j \to \infty$. The analysis of the other term is similar, and is thus omitted. Note that $\Xi = \Xi_{ij}$ is the sum of independent random variables, each with a density with respect to Lebesgue measure. Thus Ξ_{ij} has a probability density function, which we denote as f_{ij}.

We will use characteristic functions to study the density f_{ij}. We set $\phi_{ij}(t) = \mathbb{E}[\exp(\mathrm{i}t\Xi_{ij})]$. Upon direct computation, we have

$$\phi_{ij}(t) = \prod_{k=1}^{i-1} \frac{k^{\alpha}}{k^{\alpha} - \mathrm{i}t} \prod_{k=1}^{j-1} \frac{k^{\alpha}}{k^{\alpha} + \mathrm{i}t}.$$

Using the Fourier inversion theorem, the density may be related to the characteristic function. Thus, provided $i + j \geq 4$, $\phi_{ij}(t)$ is integrable with respect to Lebesgue measure and then

$$f_{ij}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}tx} \phi_{ij}(t)\, \mathrm{d}t = \frac{1}{\pi} \int_{0}^{\infty} \cos(tx) \phi_{ij}(t)\, \mathrm{d}t.$$

In particular,

$$f_{ij}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_{ij}(t)\, \mathrm{d}t.$$

Further, note that for each $t \in \mathbb{R}$ as $i, j \to \infty$,

$$\phi_{ij}(t) \to \prod_{k=1}^{\infty} \frac{k^{2\alpha}}{k^{2\alpha} + t^2},$$

and thus by dominated convergence,

$$f_{ij}(x) \to \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos(tx) \prod_{k=1}^{\infty} \frac{k^{2\alpha}}{k^{2\alpha} + t^2} \, dt =: f_\infty(x).$$

This establishes the pointwise convergence of the density. Moreover, the same argument shows that for any convergent sequence $x_{ij} \to x$, $f_{ij}(x_{ij}) \to f_\infty(x)$. In particular, we define $\beta := f_\infty(0)$. We have, since W_j is independent of Ξ_{ij}, and has the same distribution as $j^{-\alpha} W_1$,

$$\mathbb{P}[0 \le \Xi_{ij} < W_j] = \mathbb{E}\left[\int_0^{W_j} f_{ij}(z) \, dz\right] = \mathbb{E}\left[j^{-\alpha} \int_0^{W_1} f_{ij}(j^{-\alpha} z) \, dz\right]$$

and hence, by dominated convergence,

$$j^\alpha \mathbb{P}[0 \le \Xi_{ij} < W_j] \to \mathbb{E}\left[\int_0^{W_1} f_\infty(0) \, dz\right] = f_\infty(0)$$

as $i, j \to \infty$. This establishes the required asymptotics of $\mathbb{P}[0 \le \Xi_{ij} < W_j]$. $\qquad\square$

3. More continuous time

We generalize BINGO to allow for arbitrary initial states. These results will be used crucially to establish finite-dimensional convergence in Section 9.

Definition 3.1. BINGO $(a, b; c, d)$ denotes the event that the Markov chain (X_k, Y_k) given by (1.1)–(1.2) with initial state (a, b) reaches state (c, d).

As before, let V_i, W_j be exponentials at rate i^α, j^α but now restrict to $i \ge a$, $j \ge b$. Again we have a bridge.

Theorem 3.1. *Let $a \le c$ and $b \le d$, and set*

$$\Xi := \sum_{i=a}^{c-1} V_i - \sum_{j=b}^{d-1} W_j.$$

BINGO $(a, b; c, d)$ occurs if and only if either $0 \le -\Xi < V_c$ or $0 \le \Xi < W_d$.

Proof. The same as for Theorem 2.2. $\qquad\square$

Continuous time gives the asymptotics of BINGO for a wide variety of the parameters. We derive accurate estimates for various BINGO events, which will be used in our subsequent discussions.

We begin with the simplest case, starting and ending on the diagonal. (See (2.2), when starting at $(1, 1)$.)

Theorem 3.2. *For any sequence $A = A(n) \to \infty$ with $A(n) = \mathrm{o}(n)$, as $n \to \infty$,*

$$\mathbb{P}[\text{BINGO}\,(A, A; n, n)] \sim \left(\frac{2\alpha - 1}{\pi}\right)^{1/2} A^{(2\alpha-1)/2} n^{-\alpha}. \tag{3.1}$$

In the proof of Theorem 1.1, we will use only the weaker

$$\Pr[\text{BINGO}(A, A; n, n)] = n^{-\alpha+o(1)}, \quad A = n^{o(1)}. \tag{3.2}$$

Before proving Theorem 3.2 we state a generalization, where we allow initial and final points that are off the diagonal (but not too far away; we consider only what will turn out to be the typical cases: see Theorems 1.2 and 1.3). Also, for later use in Section 9, we allow $A = \Theta(n)$ as long as $n - A = \Theta(n)$ (and in this connection we change the notation from n to B).

Theorem 3.3. *Fix $M > 0$, and $\theta > 1$. Then, uniformly for all $A, B, \Lambda, \Gamma \in \frac{1}{2}\mathbb{Z}$ with $A \pm \Lambda \in \mathbb{Z}$, $B \pm \Gamma \in \mathbb{Z}$ such that $A > 0$, $B \geq \theta A$, $|\Lambda| \leq M\sqrt{A}$, and $|\Gamma| \leq M\sqrt{B}$,*

$$\mathbb{P}[\text{BINGO}(A + \Lambda, A - \Lambda; B + \Gamma, B - \Gamma)]$$

$$= (1 + o_A(1))\sqrt{\frac{2\alpha - 1}{\pi}} \frac{A^{\alpha - 1/2}}{B^\alpha \sqrt{1 - (A/B)^{2\alpha - 1}}}$$

$$\times \exp\left(-\frac{2\alpha - 1}{1 - (A/B)^{2\alpha - 1}}\left(\frac{\Lambda}{A^{1/2}} - \frac{\Gamma}{B^\alpha A^{1/2 - \alpha}}\right)^2\right), \tag{3.3}$$

where $o_A(1)$ is a quantity that tends to 0 as $A \to \infty$, uniformly in the other variables; that is, $|o_A(1)| \leq \varepsilon(A)$ for some function $\varepsilon(A) \to 0$ as $A \to \infty$.

Remark 3.1. The right-hand side of (3.3), omitting the o_A term, is the density function at Γ for a normal distribution $N(\mu, \sigma^2)$ with parameters

$$\mu = (B/A)^\alpha \Lambda,$$
$$\sigma^2 = \frac{A^{1-2\alpha} B^{2\alpha}(1 - (A/B)^{2\alpha - 1})}{2(2\alpha - 1)}.$$

The two main cases of interest to us are $A \ll B = n$, as in Theorem 3.2, and B/A constant (at least up to rounding errors). For convenience, we state immediate corollaries covering these cases.

Corollary 3.1. *Suppose $A = A(n) \to \infty$ with $A = o(n)$. Then, for all $\Lambda = \Lambda(n)$ and $\Gamma = \Gamma(n)$ with $\Lambda = O(\sqrt{A})$ and $\Gamma = O(\sqrt{n})$,*

$$\mathbb{P}[\text{BINGO}(A + \Lambda, A - \Lambda; n + \Gamma, n - \Gamma)] \sim \sqrt{\frac{2\alpha - 1}{\pi}} \frac{A^{\alpha - 1/2}}{n^\alpha} \exp\left(-(2\alpha - 1)\frac{\Lambda^2}{A}\right).$$

In Section 5 we use only the rougher asymptotics, extending (3.2):

$$\Pr[\text{BINGO}(A + \Lambda, A - \Lambda; n + \Gamma, n - \Gamma)] = n^{-\alpha+o(1)}, \quad A = n^{o(1)}. \tag{3.4}$$

Corollary 3.2. *Suppose $A = A(n)$ and $B = B(n)$ with $A \to \infty$ and $B/A \to \theta > 1$ as $n \to \infty$. Then, for all $\Lambda = \Lambda(n)$ and $\Gamma = \Gamma(n)$ with $\Lambda = O(\sqrt{A})$ and $\Gamma = O(\sqrt{B})$,*

$$\mathbb{P}[\text{BINGO}(A + \Lambda, A - \Lambda; B + \Gamma, B - \Gamma)]$$

$$\sim \sqrt{\frac{2\alpha - 1}{\pi}} \frac{A^{\alpha - 1/2}}{B^\alpha \sqrt{1 - \theta^{1-2\alpha}}} \exp\left(-\frac{(2\alpha - 1)}{(1 - \theta^{1-2\alpha})A}(\Lambda - \theta^{-\alpha}\Gamma)^2\right).$$

The proofs of Theorems 3.2 and 3.3 are similar to the proof of Theorem 2.3. We begin with a proof of the simpler Theorem 3.2, to show the main features of the proof. We then show the modifications needed for the more general Theorem 3.3.

Remark 3.2. Theorem 3.2 is basically a local CLT for Ξ; see Remark 2.2. Set

$$\Xi_n := \sum_{k=A}^{n-1} (V_k - W_k).$$

Then Ξ_n is asymptotically Gaussian with mean $\mu = 0$ and variance

$$\sigma_n^2 := \mathrm{Var}[\Xi_n] = 2 \sum_{k=A}^{n-1} k^{-2\alpha} \sim \frac{2}{2\alpha - 1} A^{1-2\alpha}. \tag{3.5}$$

Approximating Ξ_n by this Gaussian, its probability density function at zero is asymptotically

$$(2\pi)^{-1/2} \sigma^{-1} = ((2\alpha - 1)/4\pi)^{1/2} A^{(2\alpha-1)/2}.$$

The probability that $0 \leq -\Xi_n < V_n$ is then $\sim \mathbb{E}[V_n] = n^{-\alpha}$ times this, and $\mathbb{P}[\text{BINGO}]$ is twice that.

Note that in Corollary 3.1, the probability has an extra factor of $\exp[-(2\alpha - 1)(\Lambda/\sqrt{A})^2]$ over the basic $\lambda = 0$ case of Theorem 3.2. Roughly, while Ξ is still asymptotically Gaussian, the mean has moved $\sim 2\Lambda A^{-\alpha}$ from zero.

As before, the validity of the approximations is non-trivial and has forced our somewhat technical calculations.

Proof of Theorem 3.2. We assume $A \leq n - 1$ and define, as in Remark 3.2,

$$\Xi_n := \sum_{k=A}^{n-1} (V_k - W_k).$$

We note that Ξ_n is centered with variance σ_n^2 given by (3.5), and, by the same argument as in the proof of Theorem 2.3, it has a probability density function, which we denote as f_n.

We set the characteristic function $\phi_n(t) := \mathbb{E}[\exp(it\,\Xi_n)]$ and note that, by direct computation,

$$\phi_n(t) = \prod_{k=A}^{n-1} \frac{k^{2\alpha}}{k^{2\alpha} + t^2}. \tag{3.6}$$

As in our earlier analysis, we use the Fourier inversion formula to conclude that

$$f_n(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi_n(t)\, dt. \tag{3.7}$$

This again implies that $f_n(x) \leq f_n(0)$. Note that (3.7) implies

$$\sigma_n f_n(0) = \frac{\sigma_n}{\pi} \int_0^{\infty} \phi_n(t)\, dt.$$

Using the change of variables $v = \sigma_n t$, we have, by (3.6),

$$\sigma_n f_n(0) = \frac{1}{\pi} \int_0^{\infty} \phi_n\left(\frac{v}{\sigma_n}\right) dv = \frac{1}{\pi} \int_0^{\infty} \prod_{k=A}^{n-1} \frac{k^{2\alpha}}{k^{2\alpha} + v^2/\sigma_n^2}\, dv. \tag{3.8}$$

We use the dominated convergence theorem, and begin by noting that for $k \geq A$, we have, using (3.5) and letting C and c denote unspecified positive constants,

$$k^{-2\alpha} \frac{v^2}{\sigma_n^2} \leq A^{-2\alpha} \frac{v^2}{\sigma_n^2} = \mathrm{O}(v^2/A). \tag{3.9}$$

In particular, for any fixed real v, recalling again (3.6) and (3.5),

$$\ln \phi_n \left(\frac{v}{\sigma_n} \right) = -\sum_{k=A}^{n-1} \ln \left(1 + \frac{v^2/\sigma_n^2}{k^{2\alpha}} \right) \sim -\sum_{k=A}^{n-1} \frac{v^2/\sigma_n^2}{k^{2\alpha}} = -\frac{v^2}{2} \tag{3.10}$$

and thus

$$\phi_n \left(\frac{v}{\sigma_n} \right) \to \mathrm{e}^{-v^2/2}. \tag{3.11}$$

Furthermore, if $|v| \leq \sqrt{A}$, then (3.9) shows $k^{-2\alpha} v^2/\sigma_n^2 = \mathrm{O}(1)$ and thus, for some $c > 0$, $\ln(1 + k^{-2\alpha} v^2/\sigma_n^2) \geq ck^{-2\alpha} v^2/\sigma_n^2$ and, similarly to (3.10), $\ln \phi_n(v/\sigma_n) \leq -c(v^2/2)$ and thus

$$\phi_n \left(\frac{v}{\sigma_n} \right) \leq \mathrm{e}^{-cv^2/2}, \quad |v| \leq \sqrt{A}. \tag{3.12}$$

If $|v| > \sqrt{A}$, we instead have, when $k \leq 2A$, using again (3.5),

$$\frac{k^{2\alpha}}{k^{2\alpha} + v^2/\sigma_n^2} \leq \frac{(2A)^{2\alpha}}{(2A)^{2\alpha} + A/\sigma_n^2} \leq \frac{2^{2\alpha}}{2^{2\alpha} + c_1} = c_2 < 1.$$

Thus, crudely, by (3.6) and (3.5), for large enough n and $|v| > \sqrt{A}$,

$$\phi_n \left(\frac{v}{\sigma_n} \right) \leq \prod_A^{2A} \frac{k^{2\alpha}}{k^{2\alpha} + v^2/\sigma_n^2} \leq \frac{A^{2\alpha}}{A^{2\alpha} + v^2/\sigma_n^2} \prod_{A+1}^{2A} c_2 \leq \frac{A^{2\alpha} \sigma_n^2}{v^2} c_2^A. \tag{3.13}$$

For convenience, we combine (3.12) and (3.13) into the (far from sharp) estimate, valid for large n and all v,

$$\phi_n \left(\frac{v}{\sigma_n} \right) = \mathrm{O} \left(\frac{1}{1 + v^2} \right). \tag{3.14}$$

Consequently, dominated convergence yields, using (3.8), (3.11), and (3.14),

$$\sigma_n f_n(0) = \frac{1}{\pi} \int_0^\infty \phi_n \left(\frac{v}{\sigma_n} \right) \mathrm{d}v \to \frac{1}{\pi} \int_0^\infty \mathrm{e}^{-v^2/2} \, \mathrm{d}v = \frac{1}{\sqrt{2\pi}}. \tag{3.15}$$

Moreover, for any sequence $x_n = \mathrm{o}(\sigma_n)$, we obtain in the same way from (3.7)

$$\sigma_n f_n(x_n) = \frac{1}{2\pi} \int_{-\infty}^\infty \mathrm{e}^{-\mathrm{i}vx_n/\sigma_n} \phi_n \left(\frac{v}{\sigma_n} \right) \mathrm{d}v \to \frac{1}{2\pi} \int_{-\infty}^\infty \mathrm{e}^{-v^2/2} \, \mathrm{d}v = \frac{1}{\sqrt{2\pi}}. \tag{3.16}$$

We again use the fact that W_j has the same distribution as $j^{-\alpha} W_1$, and obtain

$$\mathbb{P}[0 \leq \Xi_n < W_n] = \mathbb{E} \int_0^{W_n} f_n(x) \, \mathrm{d}x = n^{-\alpha} \mathbb{E} \int_0^{W_1} f_n \left(\frac{y}{n^\alpha} \right) \mathrm{d}y. \tag{3.17}$$

Recall also that $f_n(x) \leq f_n(0)$ for every x, and thus (3.15) implies that $\sigma_n f_n(x)$ is uniformly bounded for all n and x.

Since, for every fixed y, $y/n^\alpha = \mathrm{o}(\sigma_n)$ by (3.5), we can use (3.16) and dominated convergence (twice) in (3.17) and obtain

$$n^\alpha \sigma_n \mathbb{P}[0 \leq \Xi_n < W_n] = \mathbb{E} \int_0^{W_1} \sigma_n f_n\left(\frac{y}{n^\alpha}\right) \mathrm{d}y \to \mathbb{E}\left[\frac{W_1}{\sqrt{2\pi}}\right] = \frac{1}{\sqrt{2\pi}}. \tag{3.18}$$

Thus,

$$\mathbb{P}[0 \leq \Xi_n < W_n] \sim \frac{1}{\sqrt{2\pi}} \sigma_n^{-1} n^{-\alpha} = \left(\frac{2\alpha - 1}{4\pi}\right)^{1/2} A^{(2\alpha-1)/2} n^{-\alpha}. \tag{3.19}$$

The probability $\mathbb{P}[0 \leq -\Xi_n < V_n]$ is the same, and (3.1) follows. $\qquad \square$

Proof of Theorem 3.3. The proof is similar to that of Theorem 3.2 and thus we detail only the novelties and omit some parts which are similar to the earlier proof.

Let $p(A, B, \Lambda, \Gamma)$ denote the left-hand side of (3.3), and let $q(A, B, \Lambda, \Gamma)$ denote the right-hand side without the factor $1 + o_A(1)$. First, note that if the asserted uniform estimate does not hold, then there exist $\varepsilon > 0$ and $A = A(n) \to \infty$, $B = B(n)$, $\Lambda = \Lambda(n)$, and $\Gamma = \Gamma(n)$ that satisfy the conditions such that $|p(A, B, \Lambda, \Gamma)/q(A, B, \Lambda, \Gamma) - 1| > \varepsilon$ for every n. By selecting a subsequence, we may furthermore assume that

$$A/B \to \zeta, \quad \Lambda/\sqrt{A} \to \lambda, \quad \Gamma/\sqrt{B} \to \gamma, \tag{3.20}$$

for some $\zeta \in [0, 1)$ and $\lambda, \gamma \in \mathbb{R}$. Hence, to obtain the desired contradiction, it suffices to prove that $p(A, B, \Lambda, \Gamma) \sim q(A, B, \Lambda, \Gamma)$ under the extra assumption (3.20). (This assumption is convenient below, but not essential.)

We assume (3.20) and define

$$\Xi_n = \sum_{k=A+\Lambda}^{B+\Gamma-1} V_k - \sum_{k=A-\Lambda}^{B-\Gamma-1} W_k.$$

From Theorem 3.1, BINGO $(A + \Lambda, A - \Lambda; B + \Gamma, B - \Gamma)$ occurs if either $\{0 \leq -\Xi_n < V_{B+\Gamma}\}$ or if $\{0 \leq \Xi_n < W_{B-\Gamma}\}$. We analyze $\mathbb{P}[0 \leq \Xi_n < W_{B-\Gamma}]$.

We first compute the variance of Ξ_n and observe that, using (3.20),

$$\begin{aligned}
\sigma_n^2 &:= \mathrm{Var}(\Xi_n) \\
&= \sum_{k=A+\Lambda}^{B+\Gamma-1} \mathrm{Var}V_k + \sum_{k=A-\Lambda}^{B-\Gamma-1} \mathrm{Var}W_k \\
&= \sum_{k=A+\Lambda}^{B+\Gamma-1} k^{-2\alpha} + \sum_{k=A-\Lambda}^{B-\Gamma-1} k^{-2\alpha} \\
&\sim \frac{2}{2\alpha - 1}(A^{-(2\alpha-1)} - B^{-(2\alpha-1)}) \\
&\sim \frac{2}{2\alpha - 1} A^{-(2\alpha-1)}(1 - \zeta^{2\alpha-1}).
\end{aligned} \tag{3.21}$$

We continue with the characteristic function

$$\phi_n(t) := \mathbb{E}[e^{it\Xi_n}] = \prod_{k=A+\Lambda}^{B+\Gamma-1} \frac{k^\alpha}{k^\alpha - it} \prod_{k=A-\Lambda}^{B-\Gamma-1} \frac{k^\alpha}{k^\alpha + it}.$$

This is no longer real, but we can still estimate its absolute value as in (3.12) and (3.13), with minor modifications, and obtain (3.14). Furthermore,

$$\ln \phi_n(t) = -\sum_{k=A+\Lambda}^{B+\Gamma-1} \ln\left(1 - \frac{it}{k^\alpha}\right) - \sum_{k=A-\Lambda}^{B-\Gamma-1} \ln\left(1 + \frac{it}{k^\alpha}\right).$$

We consider $t = v/\sigma_n$ for a fixed real v, and obtain by Taylor expansions, recalling (3.9) (with a trivial modification), (3.21) and (3.20),

$$\ln \phi_n\left(\frac{v}{\sigma_n}\right) = \left(\sum_{k=A+\Lambda}^{B+\Gamma-1} - \sum_{k=A-\Lambda}^{B-\Gamma-1}\right)\frac{iv/\sigma_n}{k^\alpha} - \left(\sum_{k=A+\Lambda}^{B+\Gamma-1} + \sum_{k=A-\Lambda}^{B-\Gamma-1}\right)\frac{v^2/\sigma_n^2}{2k^{2\alpha}}(1 + o(1))$$

$$= -2\Lambda\frac{iv/\sigma_n}{A^\alpha} + 2\Gamma\frac{iv/\sigma_n}{B^\alpha} - \frac{v^2}{2} + o(1)$$

$$\rightarrow \sqrt{\frac{2(2\alpha-1)}{1-\zeta^{2\alpha-1}}}(-\lambda + \zeta^{\alpha-1/2}\gamma)v\,i - \frac{v^2}{2}. \tag{3.22}$$

It follows by Fourier inversion and dominated convergence, using (3.22) and (3.14), that, for any sequence $x_n = o(\sigma_n)$, with $c_1 := \sqrt{2(2\alpha-1)/(1-\zeta^{2\alpha-1})}$,

$$\sigma_n f_n(x_n) = \frac{\sigma_n}{2\pi} \int_{-\infty}^{\infty} e^{-ix_n t}\phi_n(t)\,dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ivx_n/\sigma_n}\phi_n\left(\frac{v}{\sigma_n}\right)dv$$

$$\rightarrow \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-c_1(\lambda-\zeta^{\alpha-1/2}\gamma)v\,i - v^2/2}\,dv$$

$$= \frac{1}{\sqrt{2\pi}} e^{-c_1^2(\lambda-\zeta^{\alpha-1/2}\gamma)^2/2}. \tag{3.23}$$

Furthermore, using (3.14) again, we have the uniform bound, for all real x,

$$\sigma_n f_n(x) \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \left|\phi_n\left(\frac{v}{\sigma_n}\right)\right| dv \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{C}{1+v^2}\,dv \leq C.$$

We complete the proof as in (3.17)–(3.19), now obtaining

$$\mathbb{P}[0 \leq \Xi_n < W_{B-\Gamma}] \sim \frac{1}{\sqrt{2\pi}}\sigma_n^{-1}B^{-\alpha}\,e^{-c_1^2(\lambda-\zeta^{\alpha-1/2}\gamma)^2/2}.$$

$\mathbb{P}[0 < -\Xi_n < V_{B+\Gamma}]$ is similar, and thus

$$p(A, B, \Lambda, \Gamma) \sim \frac{2}{\sqrt{2\pi}}\sigma_n^{-1}B^{-\alpha}\,e^{-c_1^2(\lambda-\zeta^{\alpha-1/2}\gamma)^2/2}.$$

If we use (3.20) in (3.3), a simple calculation shows that the same asymptotics hold for $q(A, B, \Lambda, \Gamma)$, and thus $p(A, B, \Lambda, \Gamma) \sim q(A, B, \Lambda, \Gamma)$. As explained at the beginning of the proof, this implies the theorem. $\qquad\square$

We end this section with two less precise estimates that are useful because they do not require the condition $B \geq \theta A$ in Theorem 3.3.

Lemma 3.1. *Suppose that $0 < A < B$, and that Λ, $\Gamma \geq 0$ with $\Lambda < A$, $B - \Gamma \geq A$, and $\Gamma/B \leq \frac{1}{8}\Lambda/A$. Then, for some constant $c > 0$ depending on α only,*

$$\mathbb{P}\left[\bigcup_{\ell \leq \Gamma} \text{BINGO}\,(A + \Lambda, A - \Lambda; B + \ell, B - \ell)\right] \leq e^{-c\Lambda^2/A}. \tag{3.24}$$

Proof. Denote the event on the left-hand side of (3.24) by $\mathcal{E}$. We may assume $B + \Gamma \geq A + \Lambda$, since otherwise $\mathcal{E}$ is empty. By an argument similar to the proof of Theorem 2.2, the event $\mathcal{E}$ occurs if and only if

$$\sum_{i=A+\Lambda}^{B+\Gamma} V_i \geq \sum_{j=A-\Lambda}^{B-\Gamma-1} W_j.$$

Note that this event is monotonically decreasing in Λ; hence it suffices to prove (3.24) for $\Lambda \leq A/2$ (and $\Gamma/B \leq \frac{1}{4}\Lambda/A$), since we may otherwise decrease Λ to $A/2$ (changing c); we make these assumptions.

By Markov's inequality and independence, for every $t \geq 0$,

$$\Pr[\mathcal{E}] = \mathbb{P}\left[\sum_{i=A+\Lambda}^{B+\Gamma} V_i - \sum_{j=A-\Lambda}^{B-\Gamma-1} W_j \geq 0\right]$$

$$\leq \mathbb{E}\,e^{t\sum_{i=A+\Lambda}^{B+\Gamma} V_i - t\sum_{j=A-\Lambda}^{B-\Gamma-1} W_j}$$

$$= \prod_{i=A+\Lambda}^{B+\Gamma} \mathbb{E}\,e^{tV_i} \prod_{j=A-\Lambda}^{B-\Gamma-1} \mathbb{E}\,e^{-tW_j}. \tag{3.25}$$

Furthermore, when $-\infty < t < i^\alpha$,

$$\mathbb{E}\,e^{tV_i} = \mathbb{E}\,e^{tW_i} = \frac{1}{1 - ti^{-\alpha}}.$$

Consequently, (3.25) implies, for $0 \leq t \leq \frac{1}{2}A^{-\alpha}$ and some constant $C \geq 1$ (depending on α), using the convexity of $j \mapsto j^{-\alpha}$ in the fourth inequality,

$$\ln \Pr[\mathcal{E}] \leq -\sum_{i=A+\Lambda}^{B+\Gamma} \ln(1 - ti^{-\alpha}) - \sum_{j=A-\Lambda}^{B-\Gamma-1} \ln(1 + tj^{-\alpha})$$

$$\leq \sum_{i=A+\Lambda}^{B+\Gamma} (ti^{-\alpha} + t^2 i^{-2\alpha}) - \sum_{j=A-\Lambda}^{B-\Gamma-1} (tj^{-\alpha} - t^2 j^{-2\alpha})$$

$$\leq -\sum_{j=A-\Lambda}^{A+\Lambda-1} tj^{-\alpha} + \sum_{i=B-\Gamma}^{B+\Gamma} ti^{-\alpha} + 2\sum_{i=A-\Lambda}^{\infty} t^2 i^{-2\alpha}$$

$$\leq -2\Lambda t A^{-\alpha} + (2\Gamma + 1)t(B - \Gamma)^{-\alpha} + Ct^2 A^{1-2\alpha}$$

$$\leq -t\Lambda A^{-\alpha} + Ct^2 A^{1-2\alpha}, \tag{3.26}$$

where the last inequality follows because the assumptions imply $(2\Gamma + 1)/(B - \Gamma) \leq \Lambda/A$. Now choose $t := (2C)^{-1}\Lambda A^{\alpha-1}$; then (3.26) yields (3.24), with $c = 1/4C$. $\qquad\square$

Lemma 3.2. *For any $A < n$ and Λ with $|\Lambda| \leq A - 1$, for some universal constants C, C',*

$$\mathbb{P}[\text{BINGO}\,(A + \Lambda, A - \Lambda; n, n)] \leq \frac{C}{\sqrt{n - A}}\, \mathrm{e}^{-\Lambda^2/(n-A)} \leq \frac{C'}{\Lambda}\, \mathrm{e}^{-\Lambda^2/2(n-A)}. \tag{3.27}$$

Proof. Consider the Markov chain $(X_k, Y_k)_{2A}^{\infty}$, started at $(X_{2A}, Y_{2A}) = (A + \Lambda, A - \Lambda)$. We couple the chain with a simple random walk $(X_k^*, Y_k^*)_{2A}^{\infty}$, also started at $(A + \Lambda, A - \Lambda)$, such that, for every $k \geq 2A$,

$$|X_k - Y_k| \geq |X_k^* - Y_k^*|. \tag{3.28}$$

This can be achieved as follows. If strict inequality holds in (3.28), so $|X_k - Y_k| \geq |X_k^* - Y_k^*| + 2$ since both sides have the same parity, we may couple the next steps for the two chains arbitrarily. The same holds if $|X_k - Y_k| = |X_k^* - Y_k^*| = 0$. Finally, if $|X_k - Y_k| = |X_k^* - Y_k^*| > 0$, we have to couple such that if $|X_{k+1}^* - Y_{k+1}^*| = |X_k^* - Y_k^*| + 1$, then $|X_{k+1} - Y_{k+1}| = |X_k - Y_k| + 1$; this is always possible, since the first event has probability $1/2$, and the second has probability $\max\{X_k^\alpha, Y_k^\alpha\}/(X_k^\alpha + Y_k^\alpha) > 1/2$.

Using this coupling, (3.28) shows $|X_{2n} - Y_{2n}| \geq |X_{2n}^* - Y_{2n}^*|$, and thus

$$\begin{aligned}
\mathbb{P}[\text{BINGO}\,(A + \Lambda, A - \Lambda; n, n)] &= \mathbb{P}[X_{2n} = Y_{2n} = n] \\
&\leq \mathbb{P}[X_{2n}^* = Y_{2n}^*] \\
&= \mathbb{P}[\text{Bin}(2n - 2A, 1/2) = n - A - \Lambda] \\
&= 2^{-(2n-2A)}\binom{2n - 2A}{n - A - \Lambda},
\end{aligned}$$

and (3.27) follows by standard calculations using Stirling's formula. $\qquad\square$

4. A basic case

Here we prove a modified version of Theorem 1.1. Initially $\xi_1, \ldots, \xi_{2n-2} = \pm 1$ are uniform and independent. We set

$$A = \lfloor \ln^{10} n \rfloor. \tag{4.1}$$

We shall be splitting the walk into an initial part, until time $2A - 2$, and the main part, in the time interval $2A - 2 \leq i \leq 2n - 2$. See Remark 4.1 for further comments.

We *condition* on $S_{2A-2} = 0$ *and* $S_{2n-2} = 0$. We may and shall consider the S_i in two regimes. For $0 \leq i \leq 2A - 2$ the S_i form a random excursion, beginning ($i = 0$) and ending ($i = 2A - 2$) at zero. For $2A - 2 \leq i \leq 2n - 2$ the S_i form a random excursion beginning ($i = 2A - 2$) and ending ($i = 2n - 2$) at zero. We note, importantly, that the two sides of the walk are mutually independent excursions. Let COND denote this condition. The function $L = L_{2n-2}$ splits naturally into two parts:

$$L^{\text{init}} = \sum_{i=1}^{2A-2} \frac{S_i^2}{i^2},$$

$$L^{\text{main}} = \sum_{i=2A-1}^{2n-2} \frac{S_i^2}{i^2}.$$

Theorem 4.1. *Under* COND, *for $c \in (0, 1)$,*

$$\mathbb{P}[L^{\text{main}} \leq c \ln n \mid \text{COND}] = \mathrm{e}^{-(K(c)+\mathrm{o}(1))\ln n}$$

with (as in Theorem 1.1)

$$K(c) = \frac{(1-c)^2}{8c}.$$

Remark 4.1. There is considerable flexibility in the choice of the breakpoint A. The basic object is to protect against rare events. Our basic argument will break down when, say, $|S_i| \geq 0.01i$. This occurs with probability exponentially small in i. However, Theorem 1.1 deals with polynomially small (in n) probabilities. Thus *a priori*, for small i, the probability of this rare event is not automatically negligible on the scale of interest. Restricting to $i \geq A$, exponentially small probabilities in i are less than polynomially small in n, and hence negligible. The split at A should be considered an artifact of the proof and it is quite possible that an argument exists that does not use this artificial split. Both the restriction to $i \geq A$ and the restriction to a random excursion will be removed later.

We shall actually find the asymptotics of the Laplace transform of L^{main}. For notational convenience, given $\alpha > 1$ we define

$$\lambda = \frac{\alpha(\alpha - 1)}{2}. \tag{4.2}$$

Theorem 4.2. *For any $\alpha > 1$,*

$$\mathbb{E}[e^{-\lambda L^{\text{main}}} \mid \text{COND}] = n^{-(\alpha-1)/2 + o(1)}.$$

As α ranges over $(1, \infty)$, $t := -\lambda$ ranges over the negative reals. Theorem 4.2 then gives the asymptotics of the Laplace transform of $L^{\text{main}} \mid \text{COND}$: letting $\hat{L}_n$ be $L^{\text{main}} \mid \text{COND}$ for a particular value of n,

$$\lim_{n \to \infty} \frac{1}{\ln n} \ln \mathbb{E} \, e^{t \hat{L}_n} = \Lambda(t) := -(\alpha - 1)/2, \quad t < 0.$$

Then (as done in more detail in Section 8.3), the Legendre transform of $\Lambda(t)$ is, by a simple calculation,

$$\Lambda^*(x) = K(x),$$

and the Gärtner–Ellis theorem [9, Theorem 2.3.6] yields the asymptotics of $\mathbb{P}[\hat{L}_n \leq c \ln n]$ of Theorem 4.1.

Remark 4.2. The main contribution to $\mathbb{E}[e^{-\lambda L^{\text{main}}} \mid \text{COND}]$ comes when $L^{\text{main}} \approx c \ln n$ with $c = (2\alpha - 1)^{-1}$, i.e. $\alpha = (c + 1)/(2c)$.

Now we study $\mathbb{P}[\text{BINGO}\,(A, A; n, n)]$ as the sum of the probabilities of all paths $\mathbf{P}$ from (A, A) to (n, n). Let $P_{2A} = (A, A), \dots, P_{2n} = (n, n)$ denote the points of path $\mathbf{P}$, P_i having sum of coordinates i, $2A \leq i \leq 2n$. Let $P_i = (x_i, y_i)$. Critically, we parametrize, as in (1.5),

$$x_i = \frac{i + \delta_i}{2} \quad \text{so that} \quad y_i = \frac{i - \delta_i}{2}.$$

Here δ_i reflects the 'distance' of the path from the main diagonal. By $\Pr(\mathbf{P})$ we mean the probability of following precisely the path $\mathbf{P}$. The $\mathbb{P}(\mathbf{P})$ vary in an interesting way. The numerators multiply out the same with factors i^α for $A \leq i < n$ and j^α for $A \leq j < n$. The denominator factor $x_i^\alpha + y_i^\alpha$ is minimal when $x_i = y_i = i/2$. We define

$$\text{FIT}_i = \frac{2(i/2)^\alpha}{x_i^\alpha + y_i^\alpha}. \tag{4.3}$$

Then $\mathrm{FIT}_i = f(\varepsilon)$, where $\varepsilon = \delta_i/i$ and

$$f(\varepsilon) = \frac{2}{(1+\varepsilon)^\alpha + (1-\varepsilon)^\alpha} \le 1.$$

We shall make critical use of the asymptotics

$$\ln(f(\varepsilon)) \sim -\lambda\varepsilon^2 \text{ as } \varepsilon \to 0. \tag{4.4}$$

Set

$$\mathrm{FIT} = \mathrm{FIT}(\mathbf{P}) = \prod_{i=2A}^{2n-1} \mathrm{FIT}_i. \tag{4.5}$$

Each $\mathrm{FIT}_i \le 1$ and hence $\mathrm{FIT} \le 1$. A low FIT tells us that the path $\mathbf{P}$ is relatively unlikely. Roughly, paths $\mathbf{P}$ which stay close to the main diagonal will have a high FIT, meaning they will be more likely than those that stray far from the main diagonal. We now split

$$\mathbb{P}[\mathbf{P}] = \mathrm{BASE} \cdot \mathrm{FIT}. \tag{4.6}$$

Here BASE is what $\mathbf{P}$ would be if the terms $x_i^\alpha + y_i^\alpha$ were replaced with $2(i/2)^\alpha$ and FIT is the additional factor with the actual x_i, y_i, $2A \le i < 2n$. Then the denominator would be precisely the product of $2(i/2)^\alpha$ over $2A \le i < 2n$. That is,

$$\mathrm{BASE} = \frac{\prod_{i=A}^{n-1} i^\alpha \prod_{j=A}^{n-1} j^\alpha}{\prod_{i=2A}^{2n-1} 2(i/2)^\alpha}. \tag{4.7}$$

BINGO $(A, A; n, n)$ is the sum of $\mathrm{BASE} \cdot \mathrm{FIT}(\mathbf{P})$ over all $\binom{2(n-A)}{n-A}$ paths from (A, A) to (n, n). Summing over all paths $\mathbf{P}$, we rewrite (4.6) with the *exact* formula

$$\Pr[\mathrm{BINGO}(A, A; n, n)] = \mathrm{BASE} \cdot \binom{2(n-A)}{n-A} \cdot \mathbb{E}[\mathrm{FIT}(\mathbf{P})], \tag{4.8}$$

where expectation is over a uniformly chosen path from (A, A) to (n, n). Equation (4.8) is of fundamental importance in our analysis; indeed, this relation connects the study of the reinforced urn model with the random walk. Stirling's formula asymptotics give

$$\mathrm{BASE} = n^{-\alpha/2 + \mathrm{o}(1)} 2^{-2(n-A)}$$

and

$$\Pr[\mathrm{BINGO}(A, A; n, n)] = n^{-(1+\alpha)/2 + \mathrm{o}(1)} \cdot \mathbb{E}[\mathrm{FIT}(\mathbf{P})]. \tag{4.9}$$

Applying (3.2) we deduce

$$\mathbb{E}[\mathrm{FIT}(\mathbf{P})] = n^{(1-\alpha)/2 + \mathrm{o}(1)}. \tag{4.10}$$

Remark 4.3. We had originally hoped to apply (4.9) in reverse. That is, a combinatorial (or other) argument for the asymptotics of $\mathbb{E}[\mathrm{FIT}(\mathbf{P})]$ would yield an alternative proof, a non-Book Proof, for $\mathbb{P}[\mathrm{BINGO}]$. It was surprising that the continuous-time approach led to (4.10), which is quite difficult to prove directly.

Now we try to estimate FIT using (4.4). The technical difficulty is that we do not have $\varepsilon = \delta_i/i = \mathrm{o}(1)$ tautologically. Call a walk **P** *weird* if $|S_i| > i^{0.99}$ for some $A \leq i \leq n$. Otherwise call **P** *normal*. Large deviation results give that the probability **P** is weird for a particular i is at most $\exp[-i^{0.98}/2]$. We only look at $i \geq A$. We have selected A so that this probability is subpolynomial. As $\mathrm{FIT}(\mathbf{P}) \leq 1$ tautologically, the effect on $\mathbb{E}[\mathrm{FIT}(\mathbf{P})]$ of weird **P** is negligible. Hence in calculating $\mathbb{E}[\mathrm{FIT}(\mathbf{P})]$ we can restrict ourselves to normal **P**. Normal **P** have $\varepsilon = |S_i|/i < i^{-0.01} = \mathrm{o}(1)$ uniformly. We apply (4.4), each $\ln(\mathrm{FIT}_i) \sim -\lambda S_i^2 i^{-2}$ so that $\ln(\mathrm{FIT}) \sim -\lambda L^{\mathrm{main}}$. Therefore

$$\mathbb{E}[\mathrm{e}^{-\lambda L^{\mathrm{main}}} \mid \mathrm{COND}] = n^{(1-\alpha)/2 + \mathrm{o}(1)}$$

as desired, giving Theorem 4.2 and hence Theorem 4.1.

We now extend Theorem 4.1 to $L = L^{\mathrm{init}} + L^{\mathrm{main}}$. As $L^{\mathrm{init}} \geq 0$,

$$\mathbb{P}[L \leq c \ln n \mid \mathrm{COND}] \leq \Pr[L^{\mathrm{main}} \leq c \ln n \mid \mathrm{COND}] \leq \mathrm{e}^{-(K(c) + \mathrm{o}(1)) \ln n}.$$

Now we show L^{init} under COND is appropriately negligible. We have $\xi_1, \ldots, \xi_{2A-2} = \pm 1$ conditioned on their sum being zero. $S_i = \xi_1 + \cdots + \xi_i$. A standard second moment calculation gives the precise value $E[S_i^2] = i - i(i-1)(2A-3)^{-1}$ but we shall only use $E[S_i^2] \leq i$. (That is, the conditioning lowers the variance.) Then, using (4.1),

$$\mathbb{E}[L^{\mathrm{init}} \mid \mathrm{COND}] \leq \sum_{i=1}^{2A-2} \frac{i}{i^2} \leq (10 + \mathrm{o}(1)) \ln \ln n.$$

By Markov's inequality, with n sufficiently large, $L^{\mathrm{init}} \leq 21 \ln \ln n$ with probability at least 0.5. We have created $L^{\mathrm{init}}, L^{\mathrm{main}}$ to be independent, so with probability at least $0.5 \cdot \mathrm{e}^{-(K(c) + \mathrm{o}(1)) \ln n}$ both $L^{\mathrm{init}} \leq 21 \ln \ln n$ and $L^{\mathrm{main}} \leq c \ln n$. Hence

$$\mathbb{P}[L \leq c \ln n + 21 \ln \ln n \mid \mathrm{COND}] \geq \frac{1}{2} \mathrm{e}^{-(K(c) + \mathrm{o}(1)) \ln n}.$$

The multiplicative factor of $1/2$ and the additive factor of $21 \ln \ln n$ get absorbed in the asymptotics, giving

$$\mathbb{P}[L \leq c \ln n \mid \mathrm{COND}] \geq \mathrm{e}^{-(K(c) + \mathrm{o}(1)) \ln n}.$$

We have shown the following.

Theorem 4.3. *Under* COND,

$$\mathbb{P}[L \leq c \ln n \mid \mathrm{COND}] = \mathrm{e}^{-(K(c) + \mathrm{o}(1)) \ln n}.$$

5. The lower bound

Let $|\Lambda| \leq \sqrt{A}$, $|\Gamma| \leq \sqrt{n}$. We generalize the Basic Case. Initially $\xi_1, \ldots, \xi_{2n-2} = \pm 1$ are uniform and independent. Here we *condition* on $S_{2A-2} = 2\Lambda$ *and* $S_{2n-2} = 2\Gamma$. We may and shall consider the S_i in two regimes. For $0 \leq i \leq 2A - 2$, the S_i form a random excursion, beginning ($i = 0$) at 0 and ending ($i = 2A - 2$) at 2Λ. For $2A - 2 \leq i \leq 2n - 2$, the S_i form a random excursion beginning ($i = 2A - 2$) at 2Λ and ending ($i = 2n - 2$) at 2Γ. As in Section 4, the two excursions are independent. Let $\mathrm{COND}(\Lambda, \Gamma)$ denote this condition.

Theorem 5.1. *Uniformly in* $|\Lambda| \leq \sqrt{A}$, $|\Gamma| \leq \sqrt{n}$, *under* $\mathrm{COND}(\Lambda, \Gamma)$,

$$\mathbb{P}[L \leq c \ln n \mid \mathrm{COND}(\Lambda, \Gamma)] = \mathrm{e}^{-(K(c) + \mathrm{o}(1)) \ln n}.$$

Proof. This follows the same lines as Theorem 4.3. The critical preferential attachment Theorem 3.2 and (3.2) are replaced by Corollary 3.1 and (3.4). ☐

From Theorem 5.1 we derive the lower bound of Theorem 1.1.

Theorem 5.2. *For $c \in (0, 1)$,*

$$\mathbb{P}[L \leq c \ln n] \geq e^{-K(c)+o(1)) \ln n}$$

with

$$K(c) = \frac{(1-c)^2}{8c}.$$

Proof. We split (sum over $|\Lambda| \leq \sqrt{A}$, $|\Gamma| \leq \sqrt{n}$)

$$\mathbb{P}[L \leq c \ln n] \geq \sum_{\Lambda, \Gamma} \mathbb{P}[L \leq c \ln n \mid \mathrm{COND}\,(\Lambda, \Gamma)]\, \mathrm{Pr}\,[\mathrm{COND}\,(\Lambda, \Gamma)].$$

These conditionings are disjoint. An unrestricted random walk has probability $\Omega(1)$ of having these 'reasonable' values at $2A - 2$ and $2n - 2$, so the sum of the probabilities of COND is $\Omega(1)$. From Theorem 5.1 the conditional probabilities of $L \leq c \ln n$ are all bounded from below. ☐

6. The upper bound

We employ coupling arguments to give upper bounds on the large deviation of L.

Theorem 6.1. *For any Λ with $|\Lambda| \leq A$ and any z,*

$$\mathbb{P}[L^{\mathrm{main}} \leq z \mid \mathrm{COND}(\Lambda, 0)] \leq \mathbb{P}[L^{\mathrm{main}} \leq z \mid \mathrm{COND}(0, 0)]. \tag{6.1}$$

Proof. We couple paths

$$P_{2A} = (A + \Lambda, A - \Lambda), \ldots, P_{2n} = (n, n)$$

with paths

$$P_{2A}^* = (A, A), \ldots, P_{2n}^* = (n, n).$$

Determine the random paths P, P^* sequentially, starting at $2A$. Let t be the first value (if any) for which, setting $P_t = (a, b)$, either $P_t^* = (a, b)$ or $P_t^* = (b, a)$. In the first case couple $P_s = P_s^*$ for all $t \leq s \leq 2n$. In the second case couple P_s^* to be P_s with coordinates reversed (i.e. flip the path on the diagonal) for all $t \leq s \leq 2n$. For any paired P, P^*, $|S_i| > |S_i^*|$ for $2A \leq i < t$ and $|S_i| = |S_i^*|$ for $t \leq i \leq 2n$. Thus $L^{\mathrm{main}}(P) \geq L^{\mathrm{main}}(P^*)$ and (6.1) follows. ☐

Corollary 6.1. *For any Λ with $|\Lambda| \leq A$ and any z,*

$$\mathbb{P}[L \leq z \mid \mathrm{COND}(\Lambda, 0)] \leq \mathbb{P}[L^{\mathrm{main}} \leq z \mid \mathrm{COND}(0, 0)]. \tag{6.2}$$

Proof. $L^{\mathrm{main}} \leq L$ so $\mathbb{P}[L \leq z] \leq \mathbb{P}[L^{\mathrm{main}} \leq z]$. ☐

Corollary 6.2. *For any z,*

$$\mathbb{P}[L \leq z \mid P_{2n} = (n, n)] \leq \mathbb{P}[L^{\mathrm{main}} \leq z \mid \mathrm{COND}(0, 0)].$$

Proof. The event $P_{2n} = (n, n)$ is the disjoint union of the events $\mathrm{COND}(\Lambda, 0)$. Since, from (6.2), $\mathbb{P}[L \leq z]$ is uniformly bounded conditional under each of the events $\mathrm{COND}(\Lambda, 0)$, it has the same bound conditional on their union. $\square$

Theorem 6.2. *For any Γ with $|\Gamma| \leq n$,*

$$\mathbb{P}[L \leq z \mid P_{2n} = (n + \Gamma, n - \Gamma)] \leq \mathbb{P}[L \leq z \mid P_{2n} = (n, n)]. \tag{6.3}$$

Proof. We reverse time, and consider the random walk starting at $P_{2n}^* = (n + \Gamma, n - \Gamma)$ and ending at $P_2^* = (1, 1)$. That is, at a state (a, b) one moves to either $(a - 1, b)$ or $(a, b - 1)$ with the probabilities that the random walk from $(1, 1)$ to (a, b) goes through those states. We couple walks $P_{2n}^*, \ldots, P_2^*$ with walks $P_{2n}, \ldots, P_2$. Let t be the first value (here, highest index value) so that, with $P_t^* = (a, b)$, either $P_t = (a, b)$ or $P_t = (b, a)$. In the first case we couple $P_s^* = P_s$ for $2 \leq s \leq t$ and in the second case P_s is P_s^* with coordinates reversed for $2 \leq s \leq t$. For any paired paths P^*, P, $L(P^*) \geq L(P)$ and so the lower-tail inequality (6.3) follows. $\square$

Corollary 6.3. *For any Γ with $|\Gamma| \leq n$,*

$$\mathbb{P}[L \leq z \mid P_{2n} = (n + \Gamma, n - \Gamma)] \leq \mathbb{P}[L^{\mathrm{main}} \leq z \mid \mathrm{COND}(0, 0)].$$

Proof. Combine Corollary 6.2 and Theorem 6.2. $\square$

Theorem 6.3. *Let $\xi_3, \ldots, \xi_{2n} = \pm 1$ independently and uniformly. Let S_t be the walk with initial value $S_2 = 0$ and step $S_t = S_{t-1} + \xi_t$. Set $L = \sum_{i=2}^{2n} S_i^2/i^2$. Then*

$$\Pr[L \leq z] \leq \mathbb{P}[L^{\mathrm{main}} \leq z \mid \mathrm{COND}(0, 0)].$$

Proof. The unrestricted walk is the disjoint union of the excursions ending at $P_{2n} = (n + \Gamma, n - \Gamma)$. Corollary 6.3 gives the upper bound under any of these conditions, so the upper bound holds under their union. $\square$

We set $z = c \ln n$. Theorems 6.3 and 4.1 yield the upper bound to Theorem 1.1 and hence, together with Theorem 5.2, prove Theorem 1.1.

7. Brownian approximations

In this section we introduce a Brownian analog for L, and establish that for the purposes of establishing Theorem 1.1, it is enough to establish the corresponding statement for the Brownian analog. To this end, let $\{B_t : t \geq 0\}$ be a standard Brownian motion (with $B_0 = 0$). Thus B_n is a natural approximation of S_n.

Recall $L = L_n$ from (1.3) and define the two natural approximations

$$\widetilde{L} = \widetilde{L}_n = \sum_{i=1}^{n} \frac{B_i^2}{i^2}, \tag{7.1}$$

$$\widehat{L} = \widehat{L}_n = \int_1^n \frac{B_t^2}{t^2} \, \mathrm{d}t. \tag{7.2}$$

We introduce a cut-off A; $A := \lfloor \ln^{10} n \rfloor$ as in (4.1) works in this case as well, except that we assume that A is an even integer (this is convenient and simplifies the argument in Lemma 7.2 below, but is not essential). Define

$$L' = L'_n = \sum_{i=A+1}^{n} \frac{S_i^2}{i^2},$$

$$\widetilde{L}' = \widetilde{L}'_n = \sum_{i=A+1}^{n} \frac{B_i^2}{i^2},$$

$$\widehat{L}' = \widehat{L}'_n = \int_A^n \frac{B_t^2}{t^2}\, dt$$

and

$$L'' = L''_n = \sum_{i=A+1}^{n} \frac{(S_i - S_A)^2}{i^2}, \tag{7.3}$$

$$\widetilde{L}'' = \widetilde{L}''_n = \sum_{i=A+1}^{n} \frac{(B_i - B_A)^2}{i^2}, \tag{7.4}$$

$$\widehat{L}'' = \widehat{L}''_n = \int_A^n \frac{(B_t - B_A)^2}{t^2}\, dt. \tag{7.5}$$

Note that

$$\mathbb{E}L_n = \mathbb{E}L'_n = \sum_{i=1}^{n} \frac{1}{i} = \ln n + O(1) \tag{7.6}$$

and

$$\mathbb{E}L''_n = \int_1^n \frac{1}{t}\, dt = \ln n.$$

Throughout this discussion, C denotes some unspecified finite constants, changing from one occurrence to the next (in contrast to c, which is our main parameter). We implicitly assume that n is large. At least, assume $n \geq 8$ throughout, so $\ln \ln n \geq 1$.

Lemmas 7.1–7.4 establish that the random variable L_n and those defined in (7.1)–(7.5) are equivalent for our purposes.

Lemma 7.1. *For any $c > 0$ and $\varepsilon > 0$, for n large enough,*

$$\mathbb{P}(L_n \leq c \ln n) \geq \frac{1}{2}\mathbb{P}(L''_n \leq (c - \varepsilon) \ln n), \tag{7.7}$$

$$\mathbb{P}(\widetilde{L}_n \leq c \ln n) \geq \frac{1}{2}\mathbb{P}(\widetilde{L}''_n \leq (c - \varepsilon) \ln n), \tag{7.8}$$

$$\mathbb{P}(\widehat{L}_n \leq c \ln n) \geq \frac{1}{2}\mathbb{P}(\widehat{L}''_n \leq (c - \varepsilon) \ln n). \tag{7.9}$$

Proof. The proofs of all three parts are identical, up to obvious (notational) changes. Hence we consider only (7.7).

By Minkowski's inequality (the triangle inequality in ℓ^2),

$$\sqrt{L'_n} \leq \sqrt{L''_n} + \left(\sum_{i=A+1}^{n} \frac{S_A^2}{i^2} \right)^{1/2} \leq \sqrt{L''_n} + \frac{|S_A|}{\sqrt{A}}$$

and thus

$$\sqrt{L_n} = \sqrt{L_n' + L_A} \le \sqrt{L_n'} + \sqrt{L_A} \le \sqrt{L_n''} + \frac{|S_A|}{\sqrt{A}} + \sqrt{L_A}. \tag{7.10}$$

Furthermore, $\mathbb{E}S_A^2 = A$ and $\mathbb{E}L_A = O(\ln A) = O(\ln \ln n)$ by (7.6); hence, by Chebyshev's and Markov's inequalities, for a suitable C,

$$\mathbb{P}\left(\frac{|S_A|}{\sqrt{A}} > C\right) \le \frac{1}{4},$$

$$\mathbb{P}(L_A > C \ln \ln n) \le \frac{1}{4},$$

and thus

$$\mathbb{P}\left(\frac{|S_A|}{\sqrt{A}} + \sqrt{L_A} > C\sqrt{\ln \ln n}\right) \le \frac{1}{2}. \tag{7.11}$$

Since L_n'' is independent of S_A and L_A, it follows from (7.10) and (7.11) that

$$\mathbb{P}(L_n \le c \ln n) \ge \mathbb{P}(L_n'' \le (c - \varepsilon) \ln n)\mathbb{P}\left(\frac{|S_A|}{\sqrt{A}} + \sqrt{L_A} \le C\sqrt{\ln \ln n}\right)$$

$$\ge \frac{1}{2}\mathbb{P}\{L_n'' \le (c - \varepsilon) \log n\}. \qquad \square$$

Obviously, $L_n \ge L_n'$, $\widetilde{L}_n \ge \widetilde{L}_n'$ and $\widehat{L}_n \ge \widehat{L}_n'$. The next lemma says that L_n' is stochastically larger than L_n'', and so on.

Lemma 7.2. *For any $y \ge 0$,*

$$\mathbb{P}(L_n \le y) \le \mathbb{P}(L_n' \le y) \le \mathbb{P}(L_n'' \le y), \tag{7.12}$$

$$\mathbb{P}(\widetilde{L}_n \le y) \le \mathbb{P}(\widetilde{L}_n' \le y) \le \mathbb{P}(\widetilde{L}_n'' \le y), \tag{7.13}$$

$$\mathbb{P}(\widehat{L}_n \le y) \le \mathbb{P}(\widehat{L}_n' \le y) \le \mathbb{P}(\widehat{L}_n'' \le y). \tag{7.14}$$

Proof. Consider first (7.12). Define $\overline{S}_i := S_i - S_A$ for $i \ge A$. Then $\overline{S}_i$, $i \ge A$, is a simple random walk starting at $\overline{S}_A = 0$.

If we condition $(S_i)_{i \ge A}$ on $S_A = x$, we obtain a simple random walk starting at x. This has the same distribution as $x + \overline{S}_i$, but we shall use a different coupling defined as follows. Recall that A is chosen to be even, and thus S_A is an even integer.

For a given even integer x, define the stopping time $\tau := \inf\{k \ge A : \overline{S}_k = x/2\}$, and

$$\overline{S}_i^{(x)} := \begin{cases} x - \overline{S}_i & A \le i \le \tau, \\ \overline{S}_i & i > \tau. \end{cases}$$

Then $\overline{S}_i^{(x)}$ is a simple random walk, started at $\overline{S}_A^{(x)} = x$, and thus $(\overline{S}_i^{(x)})_A^\infty$ has the same distribution as $(x + \overline{S}_i)_A^\infty$. Furthermore, it is easily seen that, for all $i \ge A$,

$$|\overline{S}_i^{(x)}| \ge |\overline{S}_i|.$$

(To see this, we may by symmetry assume $x \ge 0$. It suffices to consider $A \le i \le \tau$, and then $\overline{S}_i \le x/2$, and thus either $\overline{S}_i \le 0$ and $\overline{S}_i^{(x)} = x + |\overline{S}_i|$, or $0 < \overline{S}_i \le x/2 \le \overline{S}_i^{(x)}$.)

Consequently, for every even integer x and every $y \geq 0$,

$$
\begin{aligned}
\mathbb{P}(L_n'' \leq y) &= \mathbb{P}\left(\sum_{i=A+1}^n \frac{\bar{S}_i^2}{i^2} \leq y \right) \\
&\geq \mathbb{P}\left(\sum_{i=A+1}^n \frac{(\bar{S}_i^{(x)})^2}{i^2} \leq y \right) \\
&= \mathbb{P}\left(\sum_{i=A+1}^n \frac{S_i^2}{i^2} \leq y \;\Big|\; S_A = x \right) \\
&= \mathbb{P}(L_n' \leq y \mid S_A = x).
\end{aligned}
$$

Thus, $\mathbb{P}(L_n'' \leq y) \geq \mathbb{P}(L_n' \leq y \mid S_A)$, and we obtain (7.12) by taking the expectation.

The proofs of (7.13) and (7.14) are the same, with S_n replaced by B_t. $\square$

Lemma 7.3. *For every $\varepsilon > 0$, $c > 0$, and $a < \infty$,*

$$
\mathbb{P}(L_n'' \leq c \ln n) \leq \mathbb{P}(\widetilde{L}_n'' \leq (c+\varepsilon) \ln n) + \mathrm{O}(n^{-a}), \tag{7.15}
$$

$$
\mathbb{P}(\widetilde{L}_n'' \leq c \ln n) \leq \mathbb{P}(L_n'' \leq (c+\varepsilon) \ln n) + \mathrm{O}(n^{-a}), \tag{7.16}
$$

$$
\mathbb{P}(\widetilde{L}_n'' \leq c \ln n) \leq \mathbb{P}(\widehat{L}_n'' \leq (c+\varepsilon) \ln n) + \mathrm{O}(n^{-a}), \tag{7.17}
$$

$$
\mathbb{P}(\widehat{L}_n'' \leq c \ln n) \leq \mathbb{P}(\widetilde{L}_n'' \leq (c+\varepsilon) \ln n) + \mathrm{O}(n^{-a}). \tag{7.18}
$$

Proof. By [18], there exists a coupling (the 'dyadic coupling') of the simple random walk $(S_i)_{i \geq 0}$ and the Brownian motion $(B_t)_{t \geq 0}$ such that, with probability $1 - \mathrm{O}(n^{-a})$, for some constant C_a,

$$
\max_{i \leq n} |S_i - B_i| \leq C_a \ln n \tag{7.19}
$$

(see also [19, Chapter 7]). If (7.19) holds, then $|(S_i - S_A) - (B_i - B_A)| \leq 2C_a \ln n$ for $A \leq i \leq n$, and thus, by Minkowski's inequality,

$$
\begin{aligned}
|\sqrt{L_n''} - \sqrt{\widetilde{L}_n''}| &\leq \left(\sum_{i=A+1}^n \frac{((S_i - S_A) - (B_i - B_A))^2}{i^2} \right)^{1/2} \\
&\leq 2C_a \frac{\ln n}{\sqrt{A}} \\
&= \mathrm{o}(1).
\end{aligned} \tag{7.20}
$$

Hence, (7.15) and (7.16) follow.

In order to prove (7.17)–(7.18), we introduce yet another version of L_n:

$$
\check{L}'' = \check{L}_n'' := \int_A^n \frac{(B_{\lceil t \rceil} - B_A)^2}{t^2} \, dt = \sum_{i=A+1}^n \frac{(B_i - B_A)^2}{i(i-1)}.
$$

Then (see (7.4)),

$$
\widetilde{L}_n'' \leq \check{L}_n'' \leq \frac{A+1}{A} \widetilde{L}_n'' = (1 + \mathrm{o}(1))\widetilde{L}_n''. \tag{7.21}
$$

Moreover, by simple standard properties of Brownian motion,

$$\mathbb{P}\Big(\sup_{t \le n} |B_{\lceil t \rceil} - B_t| > \ln n\Big) \le n\mathbb{P}\Big(\sup_{0 \le t \le 1} |B_1 - B_t| > \ln n\Big)$$

$$\le n\mathbb{P}\Big(\sup_{0 \le t \le 1} |B_t| > \frac{1}{2} \ln n\Big)$$

$$\le 4n\mathbb{P}\Big(B_1 > \frac{1}{2} \ln n\Big)$$

$$\le Cn\,e^{-\ln^2 n/8}$$

$$= \mathrm{O}(n^{-a}).$$

Hence, similarly to (7.20), with probability $1 - \mathrm{O}(n^{-a})$,

$$|\sqrt{\check{L}''_n} - \sqrt{\widehat{L}''_n}| \le \frac{\ln n}{\sqrt{A}} = \mathrm{o}(1). \tag{7.22}$$

We obtain (7.17) and (7.18) from (7.21), and (7.22). $\qquad\square$

Lemma 7.4. *For every $\varepsilon > 0$, $c > 0$, and $a < \infty$,*

$$\mathbb{P}(L_n \le c \ln n) \le 2\mathbb{P}(\widehat{L}_n \le (c + \varepsilon) \ln n) + \mathrm{O}(n^{-a}), \tag{7.23}$$

$$\mathbb{P}(\widehat{L}_n \le c \ln n) \le 2\mathbb{P}(L_n \le (c + \varepsilon) \ln n) + \mathrm{O}(n^{-a}). \tag{7.24}$$

Proof. By (7.12), (7.15), (7.17), and (7.9),

$$\mathbb{P}(L_n \le c \ln n) \le \mathbb{P}(L''_n \le c \ln n)$$

$$\le \mathbb{P}(\widetilde{L}''_n \le (c + \varepsilon) \ln n) + \mathrm{O}(n^{-a})$$

$$\le \mathbb{P}(\widehat{L}''_n \le (c + 2\varepsilon) \ln n) + \mathrm{O}(n^{-a})$$

$$\le 2\mathbb{P}(\widehat{L}_n \le (c + 3\varepsilon) \ln n) + \mathrm{O}(n^{-a}),$$

which yields (7.23) after replacing ε with $\varepsilon/3$.

Similarly, (7.14), (7.18), (7.16), and (7.7) yield

$$\mathbb{P}(\widehat{L}_n \le c \ln n) \le \mathbb{P}(\widehat{L}''_n \le c \ln n)$$

$$\le \mathbb{P}(\widetilde{L}''_n \le (c + \varepsilon) \ln n) + \mathrm{O}(n^{-a})$$

$$\le \mathbb{P}(L''_n \le (c + 2\varepsilon) \ln n) + \mathrm{O}(n^{-a})$$

$$\le 2\mathbb{P}(L_n \le (c + 3\varepsilon) \ln n) + \mathrm{O}(n^{-a}). \qquad\square$$

Consequently, it does not matter whether we use L_n or $\widehat{L}_n$ (or $\widetilde{L}_n$) in Theorem 1.1: the different versions are equivalent.

8. Analysis of the Brownian versions

Note from (7.1)–(7.2) that both $\widetilde{L}_n$ and $\widehat{L}_n$ are quadratic functionals of Gaussian variables. There is a general theory for such studying large deviation for such variables. This facilitates a direct analysis of the moment generating function of (7.2).

8.1. Moment generating function of $\widehat{L}$

We utilize the general theory of Gaussian Hilbert spaces to compute the moment generating function of $\widehat{L}_n$. For the convenience of the reader, we include a brief summary, relevant for this application, in Appendix A, and refer the interested reader to [16, Chapters VII and VI] for further details.

By Theorem A.2 and Lemma A.1, for every $t < (2 \max \lambda_j)^{-1}$,

$$\mathbb{E}\, e^{t\widehat{L}_n} = \prod_j (1 - 2\lambda_j t)^{-1/2}, \tag{8.1}$$

where (λ_j) are the non-zero eigenvalues of the integral operator

$$Tf(x) := \int_0^n \left(\frac{1}{1 \vee x \vee y} - \frac{1}{n} \right) f(y)\, \mathrm{d}y, \tag{8.2}$$

acting in $L^2(0, n)$. As shown in Appendix A (see Remark A.1), $T = T_n$ is a positive compact operator, and thus $\lambda_j > 0$; furthermore, $\sum_j \lambda_j = \mathbb{E}\widehat{L}_n = \ln n < \infty$.

Suppose that f is an eigenfunction with a non-zero eigenvalue λ. Thus $f \in L^2(0, n)$ is not identically 0, and $Tf = \lambda f$. It follows from (8.2) by dominated convergence that $Tf(x)$ is continuous in $x \in [0, n]$; thus $f = \lambda^{-1} Tf$ is continuous on $[0, n]$. Similarly, $f = \lambda^{-1} Tf$ is constant on $[0, 1]$, and $f(n) = 0$. By (8.2), we have

$$\lambda f(x) = Tf(x) = \int_0^{1 \vee x} \left(\frac{1}{1 \vee x} - \frac{1}{n} \right) f(y)\, \mathrm{d}y + \int_{1 \vee x}^n \left(\frac{1}{y} - \frac{1}{n} \right) f(y)\, \mathrm{d}y,$$

and it follows that f is continuously differentiable on $(1, n)$, with

$$\lambda f'(x) = (Tf)'(x) = -\frac{1}{x^2} \int_0^x f(y)\, \mathrm{d}y, \quad 1 < x < n. \tag{8.3}$$

Conversely, if f is continuous on $[0, n]$, constant on $[0, 1]$ and satisfies (8.3) on $(1, n)$ with the boundary condition $f(n) = 0$, then $Tf = \lambda f$.

Letting $F(x) := \int_0^x f(y)\, \mathrm{d}y$, we have $F'(x) = f(x)$, and thus (8.3) yields the differential equation

$$F''(x) = -\lambda^{-1} x^{-2} F(x), \quad 1 < x < n. \tag{8.4}$$

Furthermore, $F(1) = \int_0^1 f(x)\, \mathrm{d}x = f(1) = F'(1)$ and $F'(n) = f(n) = 0$. Hence, we have the boundary conditions (with derivatives at the endpoints 1 and n interpreted by continuity)

$$F'(1) = F(1), \tag{8.5}$$
$$F'(n) = 0. \tag{8.6}$$

Conversely, if F solves (8.4) on $(1, n)$ with the boundary conditions (8.5)–(8.6), then $f(x) := F'(x \vee 1)$ solves (8.3) and λ is an eigenvalue of T.

For a given $\lambda > 0$, the differential equation (8.4) has the solutions

$$F(x) = Ax^{\alpha_+} + Bx^{\alpha_-}, \tag{8.7}$$

where $\alpha_{\pm}$ are the solutions of $\alpha(\alpha - 1) = -\lambda^{-1}$, and thus

$$\alpha_{\pm} = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{\lambda}}. \tag{8.8}$$

If $\lambda = 4$, so we have a double root $\alpha_+ = \alpha_- = 1/2$, we instead have the solutions

$$F(x) = Ax^{1/2} + Bx^{1/2} \ln x. \tag{8.9}$$

Suppose that $\lambda > 0$ with $\lambda \neq 4$. It is easily verified that the solutions (8.7) that satisfy (8.5) are multiples of $F(x) := \alpha_+ x^{\alpha_+} - \alpha_- x^{\alpha_-}$. Hence, λ is an eigenvalue of T if and only if this function satisfies (8.6), that is, if and only if

$$\alpha_+^2 n^{\alpha_+ - 1} = \alpha_-^2 n^{\alpha_- - 1}. \tag{8.10}$$

Furthermore, this eigenvalue is simple.

Consider first the case $0 < \lambda < 4$. Then (8.8) yields the complex roots $\alpha_\pm = \frac{1}{2} \pm \omega\,\mathrm{i}$, with $\omega = \sqrt{1/\lambda - 1/4}$ and thus

$$\lambda = \frac{1}{\omega^2 + 1/4} = \frac{4}{1 + 4\omega^2}. \tag{8.11}$$

We rewrite (8.10) as

$$\left(\frac{1/2 + \omega\,\mathrm{i}}{1/2 - \omega\,\mathrm{i}}\right)^2 \mathrm{e}^{2\omega \ln n\,\mathrm{i}} = 1, \tag{8.12}$$

or, taking logarithms,

$$4\,\mathrm{Im}\ln(1 + 2\omega\,\mathrm{i}) + 2\omega \ln n \in 2\pi\,\mathbb{Z}. \tag{8.13}$$

The left-hand side of (8.13) is a continuous increasing function of $\omega \in [0, \infty)$, with the value 0 for $\omega = 0$. Hence, for a given $n \geq 2$, there is for each integer $k \geq 1$ exactly one solution $\omega_k > 0$ with

$$4\,\mathrm{Im}\ln(1 + 2\omega_k\,\mathrm{i}) + 2\omega_k \ln n = 2\pi k, \tag{8.14}$$

and it follows, by (8.11), that the eigenvalues of T in (0, 4) are

$$\lambda_k := \frac{4}{4\omega_k^2 + 1}, \quad k = 1, 2, \dots. \tag{8.15}$$

In fact, these are all the non-zero eigenvalues, since if $\lambda > 4$, so $\alpha_\pm$ are real with $\alpha_+ > \alpha_-$, then (8.10) cannot hold, and a similar argument shows that no non-zero F of the form (8.9) satisfies (8.5)–(8.6). (This also follows from Remark 8.1 below.) Hence, (8.1) shows that, for every $t > -1/8$, at least,

$$\mathbb{E}\,\mathrm{e}^{-t\widehat{L}_n} = \prod_{k=1}^{\infty}(1 + 2\lambda_k t)^{-1/2} = \prod_{k=1}^{\infty}\left(1 + \frac{8t}{1 + 4\omega_k^2}\right)^{-1/2}. \tag{8.16}$$

Note that $\mathrm{Im}\ln(1 + 2\omega_k\,\mathrm{i}) \in (0, \pi/2)$, and thus (8.14) yields

$$\frac{\pi}{\ln n}(k - 1) < \omega_k < \frac{\pi}{\ln n}k. \tag{8.17}$$

Remark 8.1. The norm of $T = T_n$ is $\lambda_1 = 4/(1 + 4\omega_1^2) = 4 - \mathrm{O}(1/\ln^2 n)$ (see (8.15) and (8.17)). If we replace the lower cut-off 1 in (8.2) with a, which by homogeneity and a change of variables is equivalent to considering $T_{n/a}$, and then let $a \to 0$ and $n \to \infty$, we obtain as a weak limit of T the integral operator on L^2 with kernel $1/(x \vee y)$. This limiting operator T_∞ is bounded on $L^2[0, \infty)$ with norm 4, but it is not compact and has no eigenvectors. That the norm is 4 follows from the result for T_n above; that it is at most 4 follows also from [14, Theorem 319]; that there are no eigenvectors in $L^2[0, \infty)$ is seen by a direct calculation similar to the one above; that T_∞ is bounded but not compact follows also from [2, Theorems 3.1 and 3.2], where a class of integral operators (including both T_∞ and T_n) is studied.

8.2. Asymptotics of the moment generating function

So far we have kept n fixed. Now consider asymptotics as $n \to \infty$. Taking logarithms in (8.16), and using (8.17), we obtain for $t > 0$

$$\frac{1}{2}\sum_{k=1}^{\infty} \ln\left(1 + \frac{8t}{1 + (4\pi^2/\ln^2 n)k^2}\right) < -\ln \mathbb{E}\, e^{-t\widehat{L}_n} < \frac{1}{2}\sum_{k=0}^{\infty} \ln\left(1 + \frac{8t}{1 + (4\pi^2/\ln^2 n)k^2}\right). \tag{8.18}$$

For $-1/8 < t < 0$, (8.18) holds with the inequalities reversed. Hence, for a fixed $t > -1/8$, uniformly in n,

$$\begin{aligned}
\ln \mathbb{E}\, e^{-t\widehat{L}_n} &= -\frac{1}{2}\sum_{k=1}^{\infty} \ln\left(1 + \frac{8t}{1 + (4\pi^2/\ln^2 n)k^2}\right) + O(1) \\
&= -\frac{1}{2}\int_0^{\infty} \ln\left(1 + \frac{8t}{1 + (4\pi^2/\ln^2 n)x^2}\right) dx + O(1). \tag{8.19} \\
&= -\frac{\ln n}{4\pi}\int_0^{\infty} \ln\left(1 + \frac{8t}{1 + y^2}\right) dy + O(1).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&\int_0^{\infty} \ln\left(1 + \frac{8t}{1 + y^2}\right) dy \\
&= \int_0^{\infty} \left(\ln(1 + 8t + y^2) - \ln(1 + y^2)\right) dy \\
&= [y(\ln(1 + 8t + y^2) - \ln(1 + y^2)) + 2\sqrt{1 + 8t}\arctan(y/\sqrt{1 + 8t}) - 2\arctan(y)]_0^{\infty} \\
&= \pi(\sqrt{1 + 8t} - 1). \tag{8.20}
\end{aligned}$$

Consequently, by (8.19) and (8.20), we have shown the following.

Theorem 8.1. *For any fixed $t > -1/8$, and all $n \geq 2$,*

$$\ln \mathbb{E}\, e^{-t\widehat{L}_n} = \frac{1 - \sqrt{1 + 8t}}{4}\ln n + O(1).$$

For $t = -1/8$, a little extra work shows that (8.19) holds with the error term $O(\ln \ln n)$. If $t < -1/8$, then $-2t\lambda_1 > 1$ for large n, and thus $\mathbb{E}\, e^{-t\widehat{L}_n} = \infty$.

8.3. A second proof of Theorem 1.1

By Theorem 8.1 (and the comments after it, for completeness),

$$\lim_{n\to\infty} \frac{1}{\ln n}\ln \mathbb{E}\, e^{t\widehat{L}_n} = \Lambda(t) := \begin{cases} \dfrac{1 - \sqrt{1 - 8t}}{4} & t \leq 1/8, \\ +\infty & t > 1/8. \end{cases} \tag{8.21}$$

The Legendre transform of $\Lambda(t)$ is, by a simple calculation,

$$\Lambda^*(x) := \sup_{t\in\mathbb{R}}(tx - \Lambda(t)) = \begin{cases} \dfrac{1}{8x}(x - 1)^2 = \dfrac{x}{8} + \dfrac{1}{8x} - \dfrac{1}{4} & x > 0, \\ +\infty & x \leq 0. \end{cases} \tag{8.22}$$

By (8.21) and the Gärtner–Ellis theorem (see e.g. [9, Theorem 2.3.6], and Remark (a) after it), the large deviation principle holds for the variables $\widehat{L}_n / \ln n$ with rate function $\Lambda^*(x)$ in (8.22), in the sense that, for example,

$$\lim_{n \to \infty} \frac{\ln \mathbb{P}(\widehat{L}_n \leq c \ln n)}{\ln n} = -\Lambda^*(c), \quad 0 < c \leq 1.$$

Note that $\Lambda^*(c) = K(c)$ given by (1.4). Consequently, we have shown the following Brownian analog of Theorem 1.1.

Theorem 8.2. *For every $c \in (0, 1]$,*

$$\mathbb{P}(\widehat{L}_n \leq c \ln n) = n^{-K(c)+o(1)}.$$

Second proof of Theorem 1.1. We use Theorem 8.2 and Lemma 7.4. □

Moreover, (8.21) and the Gärtner–Ellis theorem also give a corresponding result for the upper tail.

Theorem 8.3. *For every $c \in [1, \infty)$,*

$$\mathbb{P}(\widehat{L}_n \geq c \ln n) = n^{-K(c)+o(1)}.$$

This result too transfers from the Brownian version to the random walk.

Theorem 8.4. *For every $c \in [1, \infty)$,*

$$\mathbb{P}(L_n \geq c \ln n) = n^{-K(c)+o(1)}.$$

Proof. This follows by Theorem 8.3 and an upper-tail version of Lemma 7.4 with $\mathbb{P}(L_n \leq c \ln n)$ replaced by $\mathbb{P}(L_n \geq c \ln n)$, and so on; this version is proved in the same way as above, so we omit the details. □

Remark 8.2. In their original paper, Chatterjee and Dembo [7] considered $f : [0, 1]^n \to \mathbb{R}$, and $\mathbf{Y} = (Y_1, \dots, Y_n)$, with $Y_i \sim^{\text{i.i.d.}} \text{Ber}(p)$ random variables. For each $t > 0$, they derived explicit non-asymptotic bounds for the upper-tail probability $\mathbb{P}[f(Y) > tn]$. For functions f satisfying a 'complexity' condition on its gradient, these bounds are asymptotically tight, and establish an LDP for this non-linear functional.

Our setting differs from that of [7] in certain crucial ways. First, the relevant deviations for us occur on the $\log n$ scale, rather than the n scale. Second, the approach of [7] is inherently based on Taylor approximation, and works for a wide number of examples, for example subgraph counts, random arithmetic progressions etc. On the other hand, our approach is tailored to analyzing the specific quadratic functional under consideration.

9. Conditional functional limit laws

In this section we study the preferential attachment process $\{(X_k, Y_k) : k \geq 2\}$ defined in Section 1, and establish functional limit theorems for the trajectories, conditional on the event $\text{BINGO}(n, n)$. We define $\Delta_k := X_k - Y_k$, so that the process is given by (1.5), and state the results in terms of the stochastic process $\{\Delta_k : 2 \leq k \leq 2n\}$, conditional on $\text{BINGO}(n, n)$; recall that $\text{BINGO}(n, n)$ in this notation is the event $\Delta_{2n} = 0$.

In particular, we prove Theorem 1.2 stated in Section 1. We also state and prove related functional limit results for the process at times $o(n)$. We establish the results using the usual

two-step approach: first we establish finite-dimensional convergence and then we establish tightness (see e.g. [5]). The proofs proceed using the local CLT estimates in Section 3, in particular Theorem 3.3. Finite-dimensional convergence follows by straightforward calculations, but our proof of tightness is rather complicated, and uses several lemmas. We base the proof of tightness on a theorem by Aldous [1] (see Section 9.1 below), but for technical reasons discussed there, we do not use Aldous's result directly. Instead, in Section 9.1 we state and prove a variant that is convenient in our situation. We then prove Theorem 1.2 in Section 9.2, and give corresponding results for small times in Section 9.3.

Note that the processes (X_k, Y_k), Δ_k, and $n^{-1/2}\Delta_{\lfloor 2nt \rfloor}$ are Markov processes, and so they are (by a simple, general, calculation) also conditioned on BINGO (n, n).

9.1. A general criterion for tightness

Our proof uses a tightness criterion by Aldous [1] (and, in a slightly different formulation, Mackevičius [21]); see also [24, Lemma 3.12]. Recall that a sequence of $\mathcal{D}[0, \infty)$-valued stochastic processes $\{Z_n(t): n \geq 1\}$ is *stochastically bounded* if, for every $T > 0$,

$$\lim_{M \to \infty} \sup_n \mathbb{P}\left[\max_{0 \leq t \leq T} |X_n(t)| > M\right] = 0. \tag{9.1}$$

It is well known, and easy to see, that it suffices to show (9.1) with $\sup_n$ replaced by $\limsup_{n \to \infty}$.

Lemma 9.1. ([1, 21, 24].) *Suppose that $Z_n(t)$ is a sequence of stochastic processes in $\mathcal{D}[0, \infty)$ satisfying the following conditions.*

(i) *$\{Z_n(t): n \geq 1\}$ is stochastically bounded.*

(ii) *For each $n \geq 1$, $T > 0$, $\varepsilon > 0$, $\lambda < \infty$, and $\delta > 0$, there exists a number $\alpha_n(\lambda, \varepsilon, \delta, T)$ such that*

$$\mathbb{P}[|Z_n(u) - Z_n(t_m)| > \varepsilon \mid Z_n(t_1), \ldots, Z_n(t_m)] \leq \alpha_n(\lambda, \varepsilon, \delta, T) \tag{9.2}$$

a.s. on the event $\{\max_i |Z_n(t_i)| \leq \lambda\}$, for every finite sequence $\{t_i: 1 \leq i \leq m\}$ and u with $0 \leq t_1 \leq t_2 \leq \cdots \leq t_m \leq u \leq T$ and $u - t_m \leq \delta$. Furthermore, these numbers $\alpha_n(\lambda, \varepsilon, \delta, T)$ satisfy

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \alpha_n(\lambda, \varepsilon, \delta, T) = 0, \tag{9.3}$$

for every λ, T, ε.

Then the sequence $Z_n(t)$ is tight in $\mathcal{D}[0, \infty)$.

For Markov processes (as in our case), the condition (9.2) simplifies: by the Markov property, it suffices to consider the case $m = 1$.

A technical problem that prevents us from a direct application of Lemma 9.1 to our processes, using Theorem 3.3 to verify the condition, is that in (9.2), $u - t_m$ may be arbitrarily small, while in Theorem 3.3, B/A is supposed to be bounded below by some $\theta > 1$. We thus first prove the following variant of Lemma 9.1, where we have a lower bound on $u - t_m$. For simplicity, we state the lemma only in the Markov case. We assume also, again for simplicity, that the processes are strong Markov; recall that this means, informally, that the Markov property holds not only at fixed times but also at stopping times. A discrete-time Markov process, or a process such as our $n^{-1/2}\Delta_{\lfloor 2nt \rfloor}$ that essentially has discrete time, is automatically strong Markov.

The main difference from Lemma 9.1 is that the condition $0 \leq u - t_m \leq \delta$ is replaced by $\delta \leq u - t \leq 2\delta$. We also add a condition that the jumps are uniformly bounded (which trivially holds in our case); we do not know whether this condition really is needed. (The condition can presumably be weakened to stochastic boundedness of the jumps, as in [6], but we have not pursued this.)

Lemma 9.2. *Suppose that $Z_n(t)$ is a sequence of strong Markov processes in $\mathcal{D}[0, \infty)$ satisfying the following conditions.*

(i) *$\{Z_n(t) \colon n \geq 1\}$ is stochastically bounded.*

(ii) *For each $n \geq 1$, $T > 0$, $\varepsilon > 0$, $\lambda < \infty$, and $\delta > 0$, there exists a number $\alpha_n(\lambda, \varepsilon, \delta, T)$ such that*

$$\mathbb{P}[|Z_n(u) - Z_n(t)| > \varepsilon \mid Z_n(t)] \leq \alpha_n(\lambda, \varepsilon, \delta, T) \tag{9.4}$$

a.s. on the event $\{|Z_n(t)| \leq \lambda\}$, for every t and u with $0 \leq t \leq u \leq T$ and $t + \delta \leq u \leq t + 2\delta$. Furthermore, these numbers α_n satisfy

$$\lim_{\delta \downarrow 0} \limsup_{n \to \infty} \alpha_n(\lambda, \varepsilon, \delta, T) = 0, \tag{9.5}$$

for every λ, T, ε.

(iii) *The jumps are bounded by 1:*

$$|Z_n(t) - Z_n(t-)| \leq 1 \tag{9.6}$$

for all n and t.

Then the sequence $Z_n(t)$ is tight in $\mathcal{D}[0, \infty)$.

We reduce to Lemma 9.1 using the following lemma.

Lemma 9.3. *Suppose that $Z(t)$ is a strong Markov process in $\mathcal{D}[0, \infty)$, such that, for some given numbers $\lambda, T, \varepsilon, \delta, \alpha > 0$,*

$$\mathbb{P}[|Z(u) - Z(t)| \geq \varepsilon \mid Z(t)] \leq \alpha \tag{9.7}$$

a.s. on the event $\{|Z(t)| \leq \lambda + 2\varepsilon + 1\}$, for each t and u with $0 \leq t \leq u \leq T + 2\delta$ with $t + \delta \leq u \leq t + 2\delta$. Suppose further that the jumps in $Z(t)$ are bounded by 1, that is,

$$|Z(t) - Z(t-)| \leq 1 \tag{9.8}$$

for all $t \geq 0$. Then, for each $t \leq T$,

$$\Pr\left[\sup_{u \in [t, t+\delta]} |Z(u) - Z(t)| > 2\varepsilon \mid Z(t)\right] \leq 2\alpha$$

a.s. on the event $\{|Z(t)| \leq \lambda\}$.

Proof. Let $\mathcal{F}_t$ be the σ-field generated by $\{Z(s) \colon s \leq t\}$, and define a stopping time by

$$\tau := \inf\{u \in [t, t + \delta] \colon |Z(u) - Z(t)| \geq 2\varepsilon\}, \tag{9.9}$$

using the definition $\inf \varnothing := \infty$ if there is no such u.

Let $v := t + 2\delta$. If $\tau < \infty$, then $\tau \in [t, t + \delta]$, and thus $\tau + \delta \leq v \leq \tau + 2\delta$; hence, by (9.7) and the strong Markov property,

$$\mathbb{P}[|Z(v) - Z(\tau)| > \varepsilon \mid \mathcal{F}_\tau] \leq \alpha \tag{9.10}$$

a.s. on the event $\{\tau < \infty\} \cap \{|Z(\tau)| \leq \lambda + 2\varepsilon + 1\}$. Furthermore, $\tau < \infty$ implies, by the definition (9.9) and right-continuity,

$$|Z(\tau) - Z(t)| \geq 2\varepsilon, \tag{9.11}$$

and thus also $\tau > t$ and, by (9.7) again,

$$|Z(\tau-) - Z(t)| \leq 2\varepsilon. \tag{9.12}$$

Let $\mathcal{E}$ be any event with $\mathcal{E} \in \mathcal{F}_t$ and $\mathcal{E} \subseteq \{|Z(t)| \leq \lambda\}$. Then, on the event $\mathcal{E} \cap \{\tau < \infty\}$, by (9.8) and (9.12),

$$|Z(\tau)| \leq |Z(t)| + |Z(\tau-) - Z(t)| + |Z(\tau) - Z(\tau-)| \leq \lambda + 2\varepsilon + 1.$$

Hence, (9.10) applies, and thus, since $\mathcal{F}_t \subseteq \mathcal{F}_\tau$,

$$\mathbb{P}[\{|Z(v) - Z(\tau)| \leq \varepsilon\} \cap \mathcal{E} \cap \{\tau < \infty\}] \geq (1 - \alpha)\mathbb{P}[\mathcal{E} \cap \{\tau < \infty\}].$$

In other words, recalling that $\mathcal{E}$ can be any event in $\mathcal{F}_t$ with $\mathcal{E} \subseteq \{|Z(t)| \leq \lambda\}$,

$$\mathbb{P}[\{|Z(v) - Z(\tau)| \leq \varepsilon\} \cap \{\tau < \infty\} \mid \mathcal{F}_t] \geq (1 - \alpha)\mathbb{P}[\{\tau < \infty\} \mid \mathcal{F}_t] \tag{9.13}$$

a.s. on the event $\{|Z(t)| \leq \lambda\}$.

Furthermore, $\tau < \infty$ and $|Z(v) - Z(\tau)| \leq \varepsilon$ imply, using (9.11), $|Z(v) - Z(t)| \geq \varepsilon$. Consequently, (9.7) implies (using the Markov property)

$$\mathbb{P}[\{|Z(v) - Z(\tau)| \leq \varepsilon\} \cap \{\tau < \infty\} \mid \mathcal{F}_t] \leq \mathbb{P}[\{|Z(v) - Z(t)| \geq \varepsilon\} \mid \mathcal{F}_t] \leq \alpha \tag{9.14}$$

a.s. on the event $\{|Z(t)| \leq \lambda\}$.

Assume first $\alpha \leq 1/2$. Combining (9.13) and (9.14), we obtain

$$\mathbb{P}[\{\tau < \infty\} \mid \mathcal{F}_t] \leq \frac{\alpha}{1 - \alpha} \leq 2\alpha,$$

a.s. on the event $\{|Z(t)| \leq \lambda\}$. Since the event satisfies

$$\left\{ \sup_{u \in [t, t+\delta]} |Z(u) - Z(t)| > 2\varepsilon \right\} \subseteq \{\tau < \infty\},$$

the result follows. The case $\alpha > 1/2$ is trivial. □

Proof of Lemma 9.2. Lemma 9.3 applies to each Z_n with ε replaced by $\varepsilon/2$ and $\alpha := \alpha_n(\lambda + \varepsilon + 1, \varepsilon/2, \delta, T + 2\delta)$. This shows that, for each $t \leq T$,

$$\mathbb{P}\left[\sup_{u \in [t, t+\delta]} |Z_n(u) - Z_n(t)| > \varepsilon \mid Z_n(t) \right] \leq \alpha'_n(\lambda, \varepsilon, \delta, T) := 2\alpha_n(\lambda + \varepsilon + 1, \varepsilon/2, \delta, T + 2\delta) \tag{9.15}$$

a.s. on the event $\{|Z_n(t)| \leq \lambda\}$. Hence, assumption (ii) holds with α_n replaced by α'_n; note that (9.3) holds for α'_n by (9.5) and (9.15) (since it suffices to consider $\delta \leq 1$).

Hence, Lemma 9.1 applies and the result follows. □

9.2. Proof of Theorem 1.2

Note first that if we define $G_\alpha(t)$ by (1.8), then its covariance function agrees with (1.6); this shows that the Gaussian process $G_\alpha(t)$ in Theorem 1.2 really exists and is continuous on [0,1] (also at $t = 0$, since $\mathrm{Br}(t)$ is Hölder($\frac{1}{2} - \varepsilon$) for every $\varepsilon > 0$). Equivalently, we can define $G_\alpha(t)$ from a Brownian motion $B(t)$ by either

$$G_\alpha(t) := (\alpha - 1/2)^{-1}(t^{1-\alpha} - t^\alpha)B_{t^{2\alpha-1}/(1-t^{2\alpha-1})}, \quad 0 < t < 1, \tag{9.16}$$

or (reversing the flow of time)

$$G_\alpha(t) := (\alpha - 1/2)^{-1}t^\alpha B_{t^{1-2\alpha}-1}, \quad 0 < t \le 1. \tag{9.17}$$

Again, these are verified by calculating the covariances. Note that $G_\alpha(0) = G_\alpha(1) = 0$, for example by (1.8).

Lemma 9.4. *Finite-dimensional convergence holds in* (1.7), *that is, if* $0 \le t_1 < \cdots < t_m \le 1$ *are fixed, then, conditioned on* BINGO (n, n), *as* $n \to \infty$,

$$n^{-1/2}(\Delta_{\lfloor 2nt_1\rfloor}, \ldots, \Delta_{\lfloor 2nt_m\rfloor}) \overset{d}{\longrightarrow} (G_\alpha(t_1), \ldots, G_\alpha(t_m)). \tag{9.18}$$

Proof. Since $\Delta_0 = \Delta_{2n} = 0$ and $G_\alpha(0) = G_\alpha(1) = 0$ by definition, we may assume $0 < t_1 < \cdots < t_m < 1$. We also fix some $M > 0$. Let $n_i := \lfloor 2nt_i\rfloor$, and let $k_1, \ldots, k_m \in \mathbb{Z}$ with $n_i + k_i \in 2\mathbb{Z}$ and $|k_i| \le M\sqrt{n}$. (Assume n so large that each $n_i \ge 2$.) Then BINGO (n, n) holds together with $\Delta(n_i) = k_i$ for $i = 1, \ldots, m$, if and only if the events $\mathcal{A}_1, \ldots \mathcal{A}_{m+1}$ occur, where we set

$$\mathcal{A}_1 = \mathrm{BINGO}\left(\frac{n_1 + k_1}{2}, \frac{n_1 - k_1}{2}\right), \tag{9.19}$$

and for $2 \le i \le m + 1$, with $n_{m+1} := 2n$ and $k_{m+1} := 0$,

$$\mathcal{A}_i = \mathrm{BINGO}\left(\frac{n_{i-1} + k_{i-1}}{2}, \frac{n_{i-1} - k_{i-1}}{2}; \frac{n_i + k_i}{2}, \frac{n_i - k_i}{2}\right). \tag{9.20}$$

The events $\mathcal{A}_1, \ldots, \mathcal{A}_{m+1}$ are independent, and thus

$$\mathbb{P}[\Delta(n_1) = k_1, \ldots, \Delta(n_m) = k_m \mid \mathrm{BINGO}\,(n, n)]$$

$$= \frac{\prod_{i=1}^{m+1} \mathbb{P}[\mathcal{A}_i]}{\mathbb{P}[\mathrm{BINGO}\,(n, n)]}$$

$$= \frac{\mathbb{P}[\mathcal{A}_1]}{\mathbb{P}[\mathrm{BINGO}\,(n, n)]} \prod_{i=2}^{m+1} \mathbb{P}[\mathcal{A}_i]. \tag{9.21}$$

By Theorem 2.3, and $|k_1| \le M\sqrt{n} \le M'\sqrt{n_1}$,

$$\frac{\mathbb{P}[\mathcal{A}_1]}{\mathbb{P}[\mathrm{BINGO}\,(n, n)]} = \left(\mathbb{P}\left[\mathrm{BINGO}\left(\frac{n_1 + k_1}{2}, \frac{n_1 - k_1}{2}\right)\right]\right) \Big/ \mathbb{P}[\mathrm{BINGO}\,(n, n)]$$

$$\sim \left(\left(\frac{n_1 + k_1}{2}\right)^{-\alpha} + \left(\frac{n_1 - k_1}{2}\right)^{-\alpha}\right) \Big/ 2n^{-\alpha}$$

$$\sim \frac{2(n_1/2)^{-\alpha}}{2n^{-\alpha}}$$

$$\sim t_1^{-\alpha}, \tag{9.22}$$

where, as in the estimates below, the implicit factors $1 + o(1)$ tend to 1 as $n \to \infty$, uniformly for all $k_1, \ldots, k_m$ as above, for fixed $t_1, \ldots, t_m$ and M.

Denote the probability density function of the normal distribution $N(0, t)$ by

$$\phi_t(x) := (2\pi t)^{-1/2} \, e^{-x^2/2t}. \tag{9.23}$$

Furthermore, let

$$\kappa := \sqrt{\alpha - 1/2}, \tag{9.24}$$

$$T_i := t_i^{1-2\alpha} - 1, \tag{9.25}$$

$$y_i := \kappa t_i^{-\alpha} k_i / \sqrt{n}. \tag{9.26}$$

Note that $n_i/2 \sim t_i n$, where we define $t_{m+1} := 1$. Thus (9.20) and Corollary 3.2 yield, for $2 \leq i \leq m+1$,

$$\begin{aligned}
\mathbb{P}[\mathcal{A}_i] &\sim \sqrt{\frac{2\alpha - 1}{\pi}} \frac{1}{n^{1/2} t_i^\alpha \sqrt{t_{i-1}^{1-2\alpha} - t_i^{1-2\alpha}}} \exp\left(-\frac{(y_{i-1} - y_i)^2}{2(t_{i-1}^{1-2\alpha} - t_i^{1-2\alpha})} \right) \\
&= \frac{2\kappa}{t_i^\alpha \sqrt{n}} \phi_{T_{i-1} - T_i}(y_{i-1} - y_i).
\end{aligned} \tag{9.27}$$

Consequently, (9.21), (9.22), and (9.27) yield, uniformly for $|k_i| \leq M\sqrt{n}$,

$$\begin{aligned}
\mathbb{P}[\Delta(n_1) &= k_1, \ldots, \Delta(n_m) = k_m \mid \text{BINGO}(n, n)] \\
&\sim t_1^{-\alpha} \prod_{j=1}^m \frac{2\kappa}{t_{j+1}^\alpha \sqrt{n}} \phi_{T_j - T_{j+1}}(y_j - y_{j+1}) \\
&= \prod_{j=1}^m \frac{2\kappa}{t_j^\alpha \sqrt{n}} \phi_{T_j - T_{j+1}}(y_j - y_{j+1}).
\end{aligned} \tag{9.28}$$

Note that $T_1 > \cdots > T_m > T_{m+1} = 0$, and thus (recalling $y_{m+1} = 0$) $\prod_{i=1}^m \phi_{T_j - T_{j+1}}(y_j - y_{j+1})$ is the joint density function of $(B(T_1), \ldots, B(T_m))$ for a Brownian motion $B(t)$. Since M is arbitrary, it follows easily, recalling the scaling (9.26) and noting that $k_i + n_i \in 2\mathbb{Z}$, so k_i takes values spaced by 2, and thus y_i takes values spaced by $2\kappa t_i^{-\alpha} n^{-1/2}$, that

$$n^{-1/2}(\kappa t_1^{-\alpha} \Delta(n_1), \ldots, \kappa t_m^{-\alpha} \Delta(n_m)) \xrightarrow{\text{d}} (B(T_1), \ldots, B(T_m)).$$

Hence,

$$n^{-1/2}(\Delta(n_1), \ldots, \Delta(n_m)) \xrightarrow{\text{d}} (\kappa^{-1} t_1^\alpha B(T_1), \ldots, \kappa^{-1} t_m^\alpha B(T_m)), \tag{9.29}$$

which, using (9.25) and the construction (9.17), is the same as (9.18). □

Remark 9.1. We assumed in the proof above that $t_1, \ldots, t_m$ are fixed. In fact, the proof shows, using the uniformity assertion in Theorem 3.3, that the estimates, in particular (9.28), hold uniformly for all $0 < t_1 < \cdots < t_m < 1$ with $\min\{t_1, t_{i+1} - t_i, 1 - t_m\} \geq \delta$, for any fixed $M < \infty$ and $\delta > 0$.

We let Z_n denote the processes $Z_n(t) := n^{-1/2}\Delta_{\lfloor 2nt \rfloor}$, always conditioned on BINGO (n, n). We also let $\widehat{\mathbb{P}}$ denote probabilities conditional on BINGO (n, n).

We proceed to tightness and functional convergence. Our proof requires special arguments for t close to the endpoints 0 and 1, mainly because of the lack of uniformity in Theorem 3.3 when A is close to 0 or B. We first prove that the sequence $Z_n(t)$ is stochastically bounded on a sub-interval $[a, b] \subset (0, 1)$.

Lemma 9.5. *Let* $0 < a < b < 1$. *Then, conditioned on* BINGO (n, n)*, the sequence of stochastic processes* $Z_n(t) := n^{-1/2}\Delta_{\lfloor 2nt \rfloor}$ *is stochastically bounded on* $[a, b]$.

Proof. Let $A_0 := \lfloor 2na \rfloor$ and $B_0 := \lfloor 2nb \rfloor$. Further, let $K > 0$ be a large number. Define (for each n) the stopping time

$$\tau_K := \inf\{k \geq A_0 \colon |\Delta_k| \geq Kn^{1/2}\}, \tag{9.30}$$

as always with $\inf \varnothing := \infty$. Then

$$\Pr\left[\{A_0 < \tau_K \leq B_0\} \wedge \mathrm{BINGO}\,(n, n)\right] = \sum_{k=A_0+1}^{B_0} \mathbb{P}[\tau_K = k]\Pr\left[\mathrm{BINGO}\,(n, n) \mid \tau_K = k\right]. \tag{9.31}$$

If $\tau_K = k > A_0$, then $|\Delta_k| = \lceil Kn^{1/2} \rceil$ or $\lceil Kn^{1/2} \rceil + 1$ (depending on the parity of k). Denoting this number by $\hat{\Delta}_k$, and assuming $A_0 < k \leq B_0$, we have by Corollary 3.2, for all $n \geq n_0$ for some n_0 not depending on k, and some C not depending on n, k, or K (but perhaps on a, b, α),

$$\begin{aligned}
\mathbb{P}[\mathrm{BINGO}\,(n, n) \mid \tau_K = k] &= \mathbb{P}\left[\mathrm{BINGO}\left(\frac{k \pm \hat{\Delta}_k}{2}, \frac{k \mp \hat{\Delta}_k}{2}; n, n\right)\right] \\
&\leq Cn^{-1/2}\,\mathrm{e}^{-\hat{\Delta}_k^2/2k} \\
&\leq Cn^{-1/2}\,\mathrm{e}^{-K^2 n/2k} \\
&\leq Cn^{-1/2}\,\mathrm{e}^{-K^2/4}. \tag{9.32}
\end{aligned}$$

Note also that $\tau_K = k > A_0$ implies $|\Delta_{A_0}| < Kn^{1/2}$ by (9.30). Hence, (9.31) and (9.32) yield, for $n \geq n_0$,

$$\begin{aligned}
\mathbb{P}[\{A_0 < \tau_K \leq B_0\} \wedge \mathrm{BINGO}\,(n, n)] &\leq \sum_{k=A_0+1}^{B_0} \mathbb{P}[\tau_K = k]Cn^{-1/2}\,\mathrm{e}^{-K^2/4} \\
&= Cn^{-1/2}\,\mathrm{e}^{-K^2/4}\mathbb{P}[A_0 < \tau_K \leq B_0] \\
&\leq Cn^{-1/2}\,\mathrm{e}^{-K^2/4}\mathbb{P}[|\Delta_{A_0}| < Kn^{1/2}]. \tag{9.33}
\end{aligned}$$

Furthermore, by Theorem 2.3, as $n \to \infty$,

$$\begin{aligned}
\mathbb{P}[|\Delta_{A_0}| < Kn^{1/2}] &= \sum_{|j| < Kn^{1/2}} \Pr\left[\mathrm{BINGO}\left(\frac{A_0 + j}{2}, \frac{A_0 - j}{2}\right)\right] \\
&\sim 2Kn^{1/2} \cdot 2\beta(A_0/2)^{-\alpha} \\
&\sim 4\beta Ka^{-\alpha}n^{1/2-\alpha}. \tag{9.34}
\end{aligned}$$

By the same theorem, $\mathbb{P}[\text{BINGO}(n, n)] \sim 2\beta n^{-\alpha}$. Consequently, (9.33) implies

$$
\begin{aligned}
\limsup_{n\to\infty} \widehat{\mathbb{P}}[\{A_0 < \tau_K \le B_0\}] &= \limsup_{n\to\infty} \frac{\mathbb{P}[\{A_0 < \tau_K \le B_0\} \wedge \text{BINGO}(n, n)]}{\mathbb{P}[\text{BINGO}(n, n)]} \\
&\le \limsup_{n\to\infty} \frac{Cn^{-1/2} e^{-K^2/4} \cdot 4\beta K a^{-\alpha} n^{1/2-\alpha}}{2\beta n^{-\alpha}} \\
&= C_1 K e^{-K^2/4}.
\end{aligned}
$$

We have also, by Lemma 9.4, as $n \to \infty$,

$$
\widehat{\mathbb{P}}[\tau_K = A_0] = \widehat{\mathbb{P}}[|\Delta_{A_0}| \ge Kn^{1/2}] \to \mathbb{P}[G_\alpha(a) \ge K].
$$

Consequently, conditioned on BINGO (n, n),

$$
\begin{aligned}
\limsup_{n\to\infty} \widehat{\mathbb{P}}\Big[\sup_{a \le t \le b} |Z_n(t)| \ge K\Big] &= \limsup_{n\to\infty} \widehat{\mathbb{P}}[A_0 \le \tau_k \le B_0] \\
&\le C_1 K e^{-K^2/4} + \mathbb{P}[G_\alpha(a) \ge K].
\end{aligned}
$$

The right-hand side tends to 0 as $K \to \infty$, and the result follows. $\qquad\square$

We next prove that $Z_n(t)$ is (with large probability) uniformly small for small t.

Lemma 9.6. *For every $\varepsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ such that*

$$
\sup_{n \ge 1} \widehat{\mathbb{P}}\Big[\sup_{0 \le t \le \delta} |Z_n(t)| > \varepsilon\Big] \le \eta. \tag{9.35}
$$

Proof. The argument is similar to the proof of Lemma 9.5, but with some differences. Define (for each n) the stopping time

$$
\tau_\varepsilon := \inf\{k \ge 0 \colon |\Delta_k| \ge \varepsilon n^{1/2}\}.
$$

Let $A_0 := \lceil \varepsilon n^{1/2} \rceil$ and fix also some $M \ge \varepsilon$. Note that if $k \le A_0$, then $|\Delta_k| \le k - 2 < \varepsilon n^{1/2}$. Thus, $\tau_\varepsilon > A_0$.

Suppose that

$$
A_0 \le A \le \frac{\varepsilon}{16M} n. \tag{9.36}
$$

Then, with $\hat{\Delta}_k$ as in the proof of Lemma 9.5 (with ε instead of K),

$$
\begin{aligned}
&\mathbb{P}[\{A < \tau_\varepsilon \le 2A\} \wedge \{|\Delta_n| \le Mn^{1/2}\} \wedge \text{BINGO}(n, n)] \\
&= \sum_{k=A+1}^{2A} \sum_{|j| \le Mn^{1/2}} \mathbb{P}[\tau_\varepsilon = k] \mathbb{P}\Big[\text{BINGO}\Big(\frac{k \pm \hat{\Delta}_k}{2}, \frac{k \mp \hat{\Delta}_k}{2}; \frac{n+j}{2}, \frac{n-j}{2}\Big)\Big] \\
&\quad \times \mathbb{P}\Big[\text{BINGO}\Big(\frac{n+j}{2}, \frac{n-j}{2}; n, n\Big)\Big].
\end{aligned} \tag{9.37}
$$

Let c and C denote positive constants, not depending on n, A, ε, or M, but possibly varying from one occurrence to another. Lemma 3.1 applies by (9.36), provided $n \ge n_1(M)$ for some $n_1(M)$ depending on M only, and yields, for every k in the sum in (9.37),

$$
\sum_{|j| \le Mn^{1/2}} \mathbb{P}\Big[\text{BINGO}\Big(\frac{k + \hat{\Delta}_k}{2}, \frac{k - \hat{\Delta}_k}{2}; \frac{n+j}{2}, \frac{n-j}{2}\Big)\Big] \le e^{-c\varepsilon^2 n/2k} \le e^{-c\varepsilon^2 n/A}.
$$

By symmetry the same holds with $\hat{\Delta}_k$ replaced by $-\hat{\Delta}_k$. Similarly, Corollary 3.2 yields, provided $n \geq n_2(M)$,

$$\mathbb{P}\left[\text{BINGO}\left(\frac{n+j}{2}, \frac{n-j}{2}; n, n\right)\right] \leq Cn^{-1/2}.$$

Hence, (9.37) implies, assuming from now on that $n \geq n_1(M) \vee n_2(M)$,

$$\mathbb{P}[\{A < \tau_\varepsilon \leq 2A\} \wedge \{|\Delta_n| \leq Mn^{1/2}\} \wedge \text{BINGO}(n, n)] \leq C\mathbb{P}[A < \tau_\varepsilon \leq 2A]\, e^{-c\varepsilon^2 n/A} n^{-1/2}. \tag{9.38}$$

If $A \geq 2A_0$, we use Theorem 2.3 similarly to (9.34) and obtain

$$\mathbb{P}[A < \tau_\varepsilon \leq 2A] \leq \mathbb{P}[|\Delta_A| < \varepsilon n^{1/2}] \leq C\varepsilon n^{1/2} A^{-\alpha}. \tag{9.39}$$

Combining (9.38) and (9.39), and recalling (2.2), we then obtain

$$\widehat{\mathbb{P}}[\{A < \tau_\varepsilon \leq 2A\} \wedge \{|\Delta_n| \leq Mn^{1/2}\}] \leq C\varepsilon\, e^{-c\varepsilon^2 n/A} A^{-\alpha} n^\alpha \leq C\varepsilon^{1-2(\alpha+1)} A/n. \tag{9.40}$$

If $A_0 \leq A < 2A_0$, we instead use $\mathbb{P}[A < \tau_\varepsilon \leq 2A] \leq 1$ and obtain from (9.38) similarly

$$\begin{aligned}
\widehat{\mathbb{P}}[\{A < \tau_\varepsilon \leq 2A\} \wedge \{|\Delta_n| \leq Mn^{1/2}\}] &\leq C\, e^{-c\varepsilon^2 n/A} n^{-1/2} n^\alpha \\
&\leq C\, e^{-c\varepsilon n^{1/2}} n^{\alpha-1/2} \\
&\leq C\varepsilon^{-2\alpha-1} A_0/n, \tag{9.41}
\end{aligned}$$

yielding the same conclusion as (9.40).

Assume $0 < \delta \leq \varepsilon/32M$ and sum (9.40) or (9.41) with $A = A_j := 2^j A_0$, for $0 \leq j \leq j_0 := \lfloor \log_2(2\delta n/A_0) \rfloor$; note that $\delta n < A_{j_0} \leq 2\delta n$, and that (9.36) holds for each A_j. This yields, with $C(\varepsilon)$ denoting constants that may depend on ε,

$$\widehat{\mathbb{P}}[\{A_0 < \tau_\varepsilon \leq 2\delta n\} \wedge \{|\Delta_n| \leq Mn^{1/2}\}] \leq C(\varepsilon) \sum_{j=0}^{j_0} \frac{A_j}{n} \leq C(\varepsilon)\delta$$

and thus, recalling that $\tau_\varepsilon > A_0$,

$$\widehat{\mathbb{P}}[\max\{|\Delta_k|: k \leq 2\delta n\} \geq \varepsilon n^{1/2}] = \widehat{\mathbb{P}}[A_0 < \tau_\varepsilon \leq 2\delta n] \leq C(\varepsilon)\delta + \widehat{\mathbb{P}}[|\Delta_n| > Mn^{1/2}].$$

The right-hand side can be made smaller than η, uniformly in n, by choosing M large and δ small. We assumed above that $n \geq n_1(M) \vee n_2(M)$. The result extends trivially to all n as stated in (9.35) by decreasing δ. □

Lemma 9.7. *For every $\varepsilon > 0$ and $\eta > 0$, there exists $\delta > 0$ such that*

$$\limsup_{n\to\infty} \widehat{\mathbb{P}}\left[\sup_{1-\delta \leq t \leq 1} |Z_n(t)| > \varepsilon\right] \leq \eta. \tag{9.42}$$

Unlike in Lemma 9.6, we cannot replace $\limsup_n$ with $\sup_n$, since trivially

$$\sup_{t \geq 1-\delta} |Z_n(t)| \geq n^{-1/2}$$

for every n and δ.

Proof. Assume $0 < \delta < 1/2$, let $A_0 := \lfloor (1-\delta)2n \rfloor$ and define (for each n) the stopping time

$$\tau_\varepsilon := \inf\{k \geq A_0 : |\Delta_k| \geq \varepsilon n^{1/2}\}.$$

Let $\hat{\Delta}_k$, C, and $C(\varepsilon)$ be as in the proof of Lemma 9.6. Then, using Lemma 3.2,

$$\mathbb{P}[\{A_0 < \tau_\varepsilon < 2n\} \wedge \mathrm{BINGO}\,(n, n)]$$

$$= \sum_{k=A_0+1}^{2n-1} \mathbb{P}[\tau_\varepsilon = k] \mathbb{P}\left[\mathrm{BINGO}\left(\frac{k \pm \hat{\Delta}_k}{2}, \frac{k \mp \hat{\Delta}_k}{2}; n, n\right)\right]$$

$$\leq \sum_{k=A_0+1}^{2n-1} \mathbb{P}[\tau_\varepsilon = k] \frac{C}{\hat{\Delta}_k} \, \mathrm{e}^{-\varepsilon^2 n/4(2n-k)}$$

$$\leq C(\varepsilon) n^{-1/2} \, \mathrm{e}^{-\varepsilon^2/8\delta} \mathbb{P}[\tau_\varepsilon > A_0]. \tag{9.43}$$

Furthermore, similarly to (9.34), Theorem 2.3 yields

$$\mathbb{P}[\tau_\varepsilon > A_0] = \mathbb{P}[|\Delta_{A_0}| < \varepsilon n^{1/2}] \leq C\varepsilon n^{1/2} A_0^{-\alpha} \leq C\varepsilon n^{1/2-\alpha}. \tag{9.44}$$

Combining (9.43) and (9.44) and recalling (2.2), we obtain

$$\widehat{\mathbb{P}}[A_0 < \tau_\varepsilon < 2n] \leq C(\varepsilon) \, \mathrm{e}^{-\varepsilon^2/8\delta}. \tag{9.45}$$

Moreover, as $n \to \infty$, by Lemma 9.4,

$$\widehat{\mathbb{P}}[\tau_\varepsilon = A_0] = \widehat{\mathbb{P}}[|\Delta_{A_0}| \geq \varepsilon n^{1/2}] \to \mathbb{P}[|G_{1-\delta}| \geq \varepsilon]. \tag{9.46}$$

Hence, combining (9.45) and (9.46),

$$\limsup_{n\to\infty} \widehat{\mathbb{P}}\left[\sup_{t \geq 1-\delta} |Z_n(t)| \geq \varepsilon\right] = \limsup_{n\to\infty} \widehat{\mathbb{P}}[\{A_0 \leq \tau_\varepsilon < 2n\}]$$

$$\leq C(\varepsilon) \, \mathrm{e}^{-\varepsilon^2/8\delta} + \mathrm{Pr}\,[|G_{1-\delta}| \geq \varepsilon].$$

The right-hand side can be made less than η by choosing δ small, which yields (9.42). $\square$

Next, consider the processes only on an interval $[a, 1)$, for some fixed $a \in (0, 1)$.

Lemma 9.8. *Let $0 < a < 1$. Then $Z_n(t) \overset{\mathrm{d}}{\longrightarrow} G_\alpha(t)$ in $\mathcal{D}[a, 1)$ as $n \to \infty$.*

Proof. The space $\mathcal{D}[a, 1)$ of functions, equipped with the Skorokhod topology, is homeomorphic to the space $\mathcal{D}[0, \infty)$ by a change of variable. (Any continuous increasing bijection $[a, 1) \to [0, \infty)$ will do; we pick one, for example $t \mapsto (t-a)/(1-t)$.) It follows that Lemma 9.2 applies to $\mathcal{D}[a, 1)$ as well, considering only $T < 1$ in (i), (9.1) and (ii), and $t \geq a$ in (ii).

The stochastic boundedness on $[a, 1)$ is given by Lemma 9.5. The condition (9.6) is satisfied, because trivially each jump in $Z_n(t)$ is $n^{-1/2}$.

It remains to verify (ii) in Lemma 9.2. Let $\alpha_n(\lambda, \varepsilon, \delta, T)$ be the smallest number such that (9.4) holds, i.e. the supremum of the left-hand side over all t, u and $Z_n(t)$ satisfying the conditions.

Let $n_1 := \lfloor 2nt \rfloor$, $n_2 := \lfloor 2nu \rfloor$ and let k be an integer with $|k| \le \lambda n^{1/2}$. Then

$$\widehat{\mathbb{P}}[|Z_n(u) - Z_n(t)| \le \varepsilon \mid Z_n(t) = kn^{-1/2}] = \widehat{\mathbb{P}}[|\Delta(n_2) - \Delta(n_1)| \le \varepsilon n^{1/2} \mid \Delta(n_1) = k]$$

$$= \sum_{|j-k| \le \varepsilon n^{1/2}} \frac{\mathbb{P}[\mathcal{A}_1(j)]\mathbb{P}[\mathcal{A}_2(j)]}{\mathbb{P}[\mathcal{A}_0]}, \tag{9.47}$$

where we define the events

$$\mathcal{A}_1(j) := \text{BINGO}\left(\frac{n_1+k}{2}, \frac{n_1-k}{2}; \frac{n_2+j}{2}, \frac{n_2-j}{2}\right),$$

$$\mathcal{A}_2(j) := \text{BINGO}\left(\frac{n_2+j}{2}, \frac{n_2-j}{2}; n, n\right),$$

$$\mathcal{A}_0 := \text{BINGO}\left(\frac{n_1+k}{2}, \frac{n_1-k}{2}; n, n\right).$$

Let $\phi_t(x)$ and κ be as in (9.23) and (9.24). Further (see (9.25)–(9.26)), let $T_1 := t^{1-2\alpha} - 1$, $T_2 := u^{1-2\alpha} - 1$, $x := \kappa t^{-\alpha} k/\sqrt{n}$, and $y := \kappa u^{-\alpha} j/\sqrt{n}$. By calculations as in the proof of Lemma 9.4 (see (9.27)), we obtain

$$\frac{\mathbb{P}[\mathcal{A}_1(j)]\mathbb{P}[\mathcal{A}_2(j)]}{\mathbb{P}[\mathcal{A}_0]} \sim \frac{2\kappa}{u^\alpha \sqrt{n}} \frac{\phi_{T_1-T_2}(x-y)\phi_{T_2}(y)}{\phi_{T_1}(x)}$$

$$= \frac{2\kappa}{u^\alpha \sqrt{n}} \phi_{(T_1-T_2)T_2/T_1}\left(y - \frac{T_2}{T_1}x\right), \tag{9.48}$$

uniformly in all t, u, k, j, satisfying the conditions above (see Remark 9.1). Using (9.47) and summing (9.48) over all j with $|j-k| \le \varepsilon n^{1/2}$ and $j \equiv n_2 \pmod 2$, y takes values with step $2\kappa u^{-\alpha}/\sqrt{n}$, we obtain, with $Z \sim N(0,1)$ a standard normal variable,

$$\widehat{\mathbb{P}}[|Z_n(u) - Z_n(t)| \le \varepsilon \mid Z_n(t) = kn^{-1/2}]$$

$$\sim \int_{|u^\alpha y - t^\alpha x| \le \kappa\varepsilon} \phi_{(T_1-T_2)T_2/T_1}\left(y - \frac{T_2}{T_1}x\right) \mathrm{d}y$$

$$= \mathbb{P}\left[\left|\left(\frac{(T_1-T_2)T_2}{T_1}\right)^{1/2} Z + \frac{T_2}{T_1}x - \frac{t^\alpha}{u^\alpha}x\right| \le \kappa\varepsilon u^{-\alpha}\right],$$

uniformly in t and u satisfying the conditions. Hence, taking complements and recalling the definition of α_n,

$$\limsup_{n\to\infty} \alpha_n(\lambda, \varepsilon, \delta, T) = \sup \mathbb{P}\left[\left|\left(\frac{(T_1-T_2)T_2}{T_1}\right)^{1/2} Z + \frac{T_2}{T_1}x - \frac{t^\alpha}{u^\alpha}x\right| > \kappa\varepsilon u^{-\alpha}\right], \tag{9.49}$$

taking the supremum over t, u, x with $a \le t \le u \le T$, $t + \delta \le u \le t + 2\delta$ and $|x| \le \kappa t^{-\alpha}\lambda \le \kappa a^{-\alpha}\lambda$. For fixed $\varepsilon, \lambda > 0$ and $T < 1$, if $\delta \to 0$, then $t/u \to 1$, $T_2/T_1 \to 1$ and $T_1 - T_2 \to 0$, uniformly for all t and u satisfying these conditions. It follows from (9.49) that

$$\limsup_{n\to\infty} \alpha_n(\lambda, \varepsilon, \delta, T) \to 0 \quad \text{as } \delta \to 0,$$

which verifies (ii) in Lemma 9.2.

We have verified the conditions in Lemma 9.2, and the lemma thus shows that the sequence $Z_n(t)$ is tight in $\mathcal{D}[a, 1)$. Combined with the finite-dimensional convergence in Lemma 9.4, this shows convergence to $G_\alpha(t)$ in $\mathcal{D}[a, 1)$. □

Proof of Theorem 1.2. Lemma 9.8 shows $Z_n(t) \xrightarrow{d} G_\alpha(t)$ in $\mathcal{D}[a, 1)$ for every $a \in (0, 1)$. This can equivalently be expressed as convergence in $\mathcal{D}(0, 1)$, considering the open interval $(0, 1)$, and this can be improved to convergence in $D[0, 1]$ using Lemmas 9.6 and 9.7 (see e.g. [15, Proposition 2.4]).

A direct proof can be made as follows. Let $\varepsilon, \eta > 0$ be given. Find $\delta > 0$ such that (9.35) and (9.42) hold, and furthermore

$$\mathbb{P}\left[\sup_{0 \le t \le \delta} |G_\alpha(t)| > \varepsilon\right] \le \eta, \quad \mathbb{P}\left[\sup_{1-\delta \le t \le 1} |G_\alpha(t)| > \varepsilon\right] \le \eta.$$

By Lemma 9.8 with $a = \delta$, $Z_n(t) \xrightarrow{d} G_\alpha(t)$ in $\mathcal{D}[\delta, 1)$, and thus in $\mathcal{D}[\delta, 1 - \delta]$. By the Skorokhod coupling theorem [17, Theorem 4.30], we may assume that $Z_n(t) \to G_\alpha(t)$ uniformly on $[\delta, 1 - \delta]$. Then, now writing Pr instead of $\widehat{\mathbb{P}}$,

$$\limsup_{n \to \infty} \mathbb{P}\left[\sup_{0 \le t \le 1} |Z_n(t) - G_\alpha(t)| > 2\varepsilon\right]$$

$$\le 4\eta + \limsup_{n \to \infty} \mathbb{P}\left[\sup_{\delta \le t \le 1-\delta} |Z_n(t) - G_\alpha(t)| > 2\varepsilon\right]$$

$$= 4\eta.$$

Since η and ε are arbitrary, this shows $Z_n(t) \to G_\alpha(t)$ in $\mathcal{D}[0, 1]$ in probability, and thus in distribution. $\square$

9.3. The initial part of the process

Proof of Theorem 1.3. The proof is very similar to the proof of Theorem 1.2, and we only point out the differences. For convenience, we change the notation by replacing n with N and m_n with $2n$. Note that the normalization is then by $(2n)^{-1/2}$, differing by a factor $2^{-1/2}$ from the one in Theorem 1.2.

Finite-dimensional convergence is proved as in Lemma 9.4. The main difference is that the probability of the last event $\mathcal{A}_{m+1}$ is estimated using Corollary 3.1 instead of Corollary 3.2, and that we define $T_i := t_i^{1-2\alpha}$, where now $t_i \in (0, \infty)$. Then the same calculations as before show that (9.29) holds, which yields finite-dimensional convergence in (1.9).

Stochastic boundedness on any interval $[a, b]$ with $0 < a < b < \infty$ follows as in Lemma 9.5, again using Corollary 3.1.

The convergence (1.9) in the space $\mathcal{D}[a, \infty)$ for any $a > 0$ now follows as in Lemma 9.8, again using Lemma 9.2. This is equivalent to convergence in $\mathcal{D}(0, \infty)$.

Finally, the analog of Lemma 9.6 holds, by the same proof with trivial modifications, which yields convergence also in $\mathcal{D}[0, \infty)$. $\square$

Proof of Theorem 1.4. Apply Theorem 1.3 with $m_n := e^{t_n}$. Then (1.9) holds in $\mathcal{D}[0, \infty)$, and thus in $\mathcal{D}(0, \infty)$. The change of variables $t = e^s$ yields

$$e^{-t_n/2} \Delta(\lfloor e^{s+t_n} \rfloor) \xrightarrow{d} (2\alpha - 1)^{-1/2} e^{(1-\alpha)s} B(e^{(2\alpha-1)s}), \tag{9.50}$$

in $\mathcal{D}(-\infty, \infty)$. Multiplying (9.50) by the continuous function $e^{-s/2}$ yields

$$e^{-(s+t_n)/2} \Delta(\lfloor e^{s+t_n} \rfloor) \xrightarrow{d} Z(s) := (2\alpha - 1)^{-1/2} e^{(1/2-\alpha)s} B(e^{(2\alpha-1)s}), \tag{9.51}$$

in $\mathcal{D}(-\infty, \infty)$, which is (1.10). The covariance (1.11) follows from the definition of $Z(s)$ in (9.51). $\square$

Appendix A. Quadratic functionals of Gaussian variables

A *Gaussian Hilbert space* is a closed subspace H of $L^2(\Omega, \mathcal{F}, P)$, for some probability space $(\Omega, \mathcal{F}, P)$, such that every element f of H is a random variable with a centered Gaussian distribution $N(0, \sigma^2)$ (where $\sigma^2 = \|f\|_2^2$). (In this appendix, we consider real vector spaces and real-valued functions; thus $L^2 = L^2_{\mathbb{R}}$ is the space of real-valued square-integrable functions.) Here, for the readers' and our own convenience, we review some basic and more or less well-known facts; further details can be found in [16], for example. In our application above, H is the closed linear span of $\{B_t : 0 \le t \le n\}$, as usual defined on some anonymous probability space $(\Omega, \mathcal{F}, P)$, but we state the results generally.

If $\xi, \eta \in H$, define their *Wick product* by $:\xi\eta: := \xi\eta - \mathbb{E}(\xi\eta)$, i.e. the centered product. Let $H^{:2:}$ be the closed linear span of all Wick products $:\xi\eta:, \xi, \eta \in H$. Hence, if $X = Q(\xi_1, \ldots, \xi_N)$ is a quadratic form in Gaussian random variables $\xi_1, \ldots, \xi_N \in H$, then $X - \mathbb{E}X \in H^{:2:}$, and conversely, every element of $H^{:2:}$ is a limit of such centered quadratic forms, and can be written as a quadratic form in (in general) infinitely many variables. Moreover, this form can be diagonalized by a suitable choice of orthonormal basis in H, leading to the following representation theorem. (Note that every $X \in H^{:2:}$ has $\mathbb{E}X = 0$ as a consequence of the definition.)

Theorem A.1. ([16, Theorem 6.1].) *If $X \in H^{:2:}$, then there exists a finite or infinite sequence $(\lambda_j)_{j=1}^N$, $0 \le N \le \infty$, of non-zero real numbers such that $\sum_j \lambda_j^2 < \infty$ and*

$$X \overset{\mathrm{d}}{=} \sum_{j=1}^N \lambda_j(\xi_j^2 - 1), \tag{A.1}$$

where ξ_j are independent, identically distributed (i.i.d.) $N(0, 1)$ random variables. The numbers λ_j are the non-zero eigenvalues (counted with multiplicities) of the compact symmetric bilinear form

$$Q_X(\xi, \eta) := \frac{1}{2}\mathbb{E}(X\xi\eta), \quad \xi, \eta \in H,$$

or, equivalently, of the corresponding compact symmetric operator T_X on H defined by

$$\langle T_X(\xi), \eta \rangle = Q_X(\xi, \eta) = \frac{1}{2}\mathbb{E}(X\xi\eta), \quad \xi, \eta \in H.$$

In particular, (A.1) yields the moment generating function, for all real t such that $2\lambda_j t < 1$ for every j,

$$\mathbb{E}\,\mathrm{e}^{tX} = \prod_j (1 - 2\lambda_j t)^{-1/2}\,\mathrm{e}^{-\lambda_j t}. \tag{A.2}$$

In our application, we deal with non-centered quadratic functionals, and then the following version is more directly applicable. (See the more general but less specific [16, Theorem 6.2]. We do not know a reference for the precise statements in Theorem A.2, so we give a complete proof.)

Theorem A.2. *Suppose the following.*

(i) *X is a random variable such that $X - \mathbb{E}X \in H^{:2:}$.*

(ii) *$X \ge 0$ a.s., and $\mathbb{P}(X < \varepsilon) > 0$ for every $\varepsilon > 0$, that is, the lower bound of the support of X is 0.*

(iii) *The bilinear form $Q = Q_{X-\mathbb{E}X}$ is positive, that is,*

$$Q_{X-\mathbb{E}X}(\xi, \xi) = \frac{1}{2}\mathbb{E}((X - \mathbb{E}X)\xi^2) \geq 0$$

for every $\xi \in H$.

Then

$$X \overset{\mathrm{d}}{=} \sum_{j=1}^{N} \lambda_j \xi_j^2, \tag{A.3}$$

where ξ_j are i.i.d. $N(0, 1)$ random variables and the coefficients $\lambda_j > 0$ are the non-zero eigenvalues (counted with multiplicities) of Q. Furthermore,

$$\mathbb{E}X = \sum_{j=1}^{N} \lambda_j < \infty \tag{A.4}$$

and, for $-\infty < t < (2\max_j \lambda_j)^{-1}$,

$$\mathbb{E}\,e^{tX} = \prod_j (1 - 2\lambda_j t)^{-1/2}. \tag{A.5}$$

Proof. Theorem A.1 yields the representation

$$X - \mathbb{E}X = \sum_j \lambda_j(\xi_j^2 - 1), \tag{A.6}$$

where $\lambda_j > 0$ since the positive form Q has only non-negative eigenvalues. Hence, (A.2) applies for any $t < 0$, and thus, replacing t with $-t$, for every $t > 0$,

$$\frac{1}{t}\ln\mathbb{E}\,e^{-t(X-\mathbb{E}X)} = \sum_j \frac{t\lambda_j - \frac{1}{2}\ln(1 + 2\lambda_j t)}{t}. \tag{A.7}$$

Now let $t \to \infty$. By (ii), $\mathbb{E}\,e^{-tX} \leq 1$, but

$$\liminf_{t\to\infty} \frac{1}{t}\ln\mathbb{E}\,e^{-tX} \geq -\varepsilon \quad \text{for every } \varepsilon > 0,$$

and thus

$$\frac{1}{t}\ln\mathbb{E}\,e^{-tX} \to 0 \quad \text{and} \quad \frac{1}{t}\ln\mathbb{E}\,e^{-t(X-\mathbb{E}X)} \to \mathbb{E}X.$$

In the sum on the right-hand side of (A.7), each term is positive, and increases to λ_j as $t \to \infty$ (because ln is concave); hence the sum tends to $\sum_j \lambda_j$ by monotone convergence. Consequently, $\mathbb{E}X = \sum_j \lambda_j$, and since $\mathbb{E}X < \infty$, (A.4) holds.

The representation (A.3) follows from (A.6) and (A.4), and (A.5) is an immediate consequence. □

Remark A.1. The operator T in Theorem A.1 is a Hilbert–Schmidt operator, since $\sum_j \lambda_j^2 < \infty$. Similarly, in Theorem A.2, T is a trace class operator, with trace and trace norm $\sum_j \lambda_j = \mathbb{E}X$.

If $Y \in H$, then $X := Y^2$ satisfies (i) and (ii) in Theorem A.2. Furthermore, for $\xi, \eta \in H$, since $X - \mathbb{E}X = Y^2 - \mathbb{E}Y^2 =: Y^2:$, by [16, Theorem 3.9],

$$Q_{X-\mathbb{E}X}(\xi, \eta) = \frac{1}{2}\mathbb{E}(: Y^2 :\xi\eta) = \frac{1}{2}\mathbb{E}(: Y^2 :: \xi\eta :) = \mathbb{E}(Y\xi)\mathbb{E}(Y\eta). \tag{A.8}$$

Taking $\eta = \xi$, we see that Q is a positive form. Hence the conditions (i)–(iii) hold for $X = Y^2$, and it follows that they also hold for any finite sum of squares $\sum_i Y_i^2$ with $Y_i \in H$, for example $\widetilde{L}_n$ in (7.1). Moreover, we can take limits, and conclude that the conditions also hold for the integral $\widehat{L}_n$ in (7.2), for example. (Note that (ii) is obvious for $\widehat{L}_n$.) Hence, Theorem A.2 applies to $\widehat{L}_n$.

A.1. Stochastic integration

In order to find the eigenvalues λ_i in Theorem A.2 in our application to $\widehat{L}_n$, it is convenient to transfer from the Gaussian Hilbert space H to the function space $L^2(0, n)$ by means of stochastic integrals.

The stochastic integral $\int_0^n f(t)\, dB_t$ can be defined for every (deterministic) function $f \in L^2(0, n)$ as follows. (This is a simple form of stochastic integrals; we have no need for the general theory of random integrands here.)

First, $\int_0^n \mathbf{1}_{(0,a)}(t)\, dB_t = B_a$ for every $a \in [0, n]$. This and linearity define $\int_0^n f(t)\, dB_t$ for every step function f (in the obvious, naive way). A simple calculation shows that

$$\mathbb{E}\left(\int_0^n f(t)\, dB_t\right)^2 = \int_0^n f(t)^2\, dt.$$

Hence, the mapping $I: f \mapsto \int f\, dB_t$ is an isometry from the subspace of step functions in $L^2(0, n)$ to the linear span of the random variables B_t, $t \in [0, n]$. We let H be the closure of the latter space, regarded as a subspace of $L^2(\Omega, \mathcal{F}, P)$; then I extends by continuity to an isometry $I: L^2(0, n) \to H$. We may write $I(f) = \int_0^n f(t)\, dB_t$. This isometry enables us to regard the bilinear form Q and operator T as defined on $L^2(0, n)$.

If $Y = I(f)$, $\xi = I(g)$ and $\eta = I(h)$, for some $f, g, h \in L^2(0, n)$, and $X = Y^2$, then (A.8) yields

$$Q(\xi, \eta) = \mathbb{E}(Y\xi)\mathbb{E}(Y\eta) = \langle f, g\rangle\langle f, h\rangle = \int_0^n \int_0^n f(x)f(y)g(x)h(y)\, dx\, dy,$$

which shows that T, regarded as an operator on $L^2(0, n)$, is the integral operator with kernel $f(x)f(y)$.

Example A.1. For any $t \in [0, n]$, $B_t = I(\mathbf{1}_{(0,t)})$, and thus $X = B_t^2$ corresponds to the integral operator T with kernel $\mathbf{1}_{(0,t)}(x)\mathbf{1}_{(0,t)}(y)$. It follows easily that $\widehat{L}_n = \int_1^n (B_t^2/t^2)\, dt$ corresponds to the integral operator with kernel

$$K_n(x, y) := \int_1^n \frac{1}{t^2} \mathbf{1}_{(0,t)}(x)\mathbf{1}_{(0,t)}(y)\, dt = \int_{1 \vee x \vee y}^n \frac{dt}{t^2} = \frac{1}{1 \vee x \vee y} - \frac{1}{n}. \tag{A.9}$$

We summarize as follows.

Lemma A.1. *Theorem A.2 applies to $X = \widehat{L}_n$, with λ_j the non-zero eigenvalues of the symmetric integral operator on $L^2(0, n)$ with kernel K_n given by* (A.9).

Acknowledgement

References

[1] ALDOUS, D. (1978). Stopping times and tightness. *Ann. Prob.* **6** (2), 335–340.

[2] ALEKSANDROV, A. B., JANSON, S., PELLER, V. V. AND ROCHBERG, R. (2002). An interesting class of operators with unusual Schatten–von Neumann behavior. In *Function Spaces, Interpolation Theory and Related Topics (Lund, 2000)*, pp. 61–149. De Gruyter, Berlin.

[3] BHATTACHARYA, B. B., GANGULY, S., LUBETZKY, E. AND ZHAO, Y. (2017). Upper tails and independence polynomials in random graphs. *Adv. Math.* **319**, 313–347.

[4] BHATTACHARYA, B. B., GANGULY, S., SHAO, X. AND ZHAO, Y. (2019). Upper tails for arithmetic progressions in a random set. To appear in *Int. Math. Res. Not.*

[5] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[6] BILLINGSLEY, P. (1974). Conditional distributions and tightness. *Ann. Prob.* **2**, 480–485.

[7] CHATTERJEE, S. AND DEMBO, A. (2016). Nonlinear large deviations. *Adv. Math.* **299**, 396-450.

[8] DAVIS, B. (1990). Reinforced random walks. *Prob. Theory Relat. Fields* **84**, 203–229.

[9] DEMBO, A. AND ZEITOUNI, O. (1998). *Large Deviation Techniques and Applications*, 2nd edn. Springer, New York.

[10] EGGENBERGER, F. AND PÓLYA, G. (1923). Über die Statistik verketteter Vorgänge. *Z. Angew. Math. Mech.* **3** (4), 279–289.

[11] ELDAN, R. (2018). Gaussian-width gradient complexity, reverse log-Sobolev inequalities and non-linear large deviations. *Geom. Funct. Anal.* **28** (6), 1548–1596.

[12] ELDAN, R. AND GROSS, R. (2018). Decomposition of mean-field Gibbs distributions into product measures. *Electron. J. Probab.* **23**, paper no. 35, 24.

[13] ELDAN, R. AND GROSS, R. (2018). Exponential random graphs behave like mixtures of stochastic block models. *Ann. Appl. Probab.* **28** (6), 3698–3735.

[14] HARDY, G. H., LITTLEWOOD, J. E. AND PÓLYA, G. (1952). *Inequalities*, 2nd edn. Cambridge University Press, Cambridge.

[15] JANSON, S. (1994). *Orthogonal Decompositions and Functional Limit Theorems for Random Graph Statistics* (Mem. Amer. Math. Soc. **534**). American Mathematical Society, Providence, RI.

[16] JANSON, S. (1997). *Gaussian Hilbert Spaces*. Cambridge University Press, Cambridge.

[17] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd edn. Springer, New York.

[18] KOMLÓS, J., MAJOR, P. AND TUSNÁDY, G. (1975). An approximation of partial sums of independent RV's and the sample DF, I. *Z. Wahrscheinlichkeitsth.* **32**, 111–131.

[19] LAWLER, G. F. AND LIMIC, V. (2010). *Random Walk: A Modern Introduction*. Cambridge University Press, Cambridge.

[20] LUBETZKY, E. AND ZHAO, Y. (2017). On the variational problem for upper tails in sparse random graphs. *Random Structures Algorithms* **50** (3), 420-436.

[21] MACKEVIČIUS, V. (1974). On the question of the weak convergence of random processes in the spaces $D[0, \infty)(X)$ (in Russian). *Litovsk. Mat. Sb.* **14** (4), 117–121, 237. English translation: *Lithuanian Math. Trans.* (*Lithuanian Math. J.*) **14** (4), 620–623 (1975).

[22] OLIVEIRA, R. AND SPENCER, J. (2005). Connectivity transitions in networks with super-linear preferential attachment. *Internet Math.* **2**, 121–163

[23] PÓLYA, G. (1930). Sur quelques points de la théorie des probabilités. *Ann. Inst. H. Poincaré Prob. Statist.* **1**, 117–161.

[24] WHITT, W. (2007). Proofs of the martingale FCLT. *Probab. Surv.* **4**, 268-302.