



Turning a Coin over Instead of Tossing It

János Engländer¹ · Stanislav Volkov² 

Received: 13 June 2016 / Revised: 7 November 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract Given a sequence of numbers $(p_n)_{n \geq 2}$ in $[0, 1]$, consider the following experiment. First, we flip a fair coin and then, at step n , we turn the coin over to the other side with probability p_n , $n \geq 2$, independently of the sequence of the previous terms. What can we say about the distribution of the empirical frequency of heads as $n \rightarrow \infty$? We show that a number of phase transitions take place as the turning gets slower (i. e., p_n is getting smaller), leading first to the breakdown of the Central Limit Theorem and then to that of the Law of Large Numbers. It turns out that the critical regime is $p_n = \text{const}/n$. Among the scaling limits, we obtain uniform, Gaussian, semicircle, and arcsine laws.

Keywords Coin tossing · Central Limit Theorem · Laws of Large Numbers

Mathematics Subject Classification (2010) 60F05 · 60J10

The hospitality of Microsoft Research and the University of Washington is gratefully acknowledged by the first author.

Research of the second author was supported in part by the Swedish Research Council Grant VR2014–5157.

✉ Stanislav Volkov
s.volkov@maths.lth.se
<http://www.maths.lth.se/~s.volkov/>

János Engländer
janos.englander@colorado.edu
<http://euclid.colorado.edu/~englandj/MyBoulderPage.html>

¹ Department of Mathematics, University of Colorado, Boulder, CO 80309-0395, USA

² Centre for Mathematical Sciences, Lund University, 22100-118 Lund, Sweden

1 General Model

In this paper, we examine what happens if, instead of tossing a coin, we turn it over (from heads to tails and from tails to heads), with certain probabilities.

To define the model precisely, let $p_n, n = 2, 3, \dots$ be a given deterministic sequence of numbers between 0 and 1. We define a time-dependent “coin turning process” X with $X_n \in \{0, 1\}, n \geq 1$, as follows. Let $X_1 = 1$ (“heads”) or $= 0$ (“tails”) with probability $1/2$. For $n \geq 2$, set recursively

$$X_n := \begin{cases} 1 - X_{n-1}, & \text{with probability } p_n; \\ X_{n-1}, & \text{otherwise,} \end{cases}$$

that is, we turn the coin over with probability p_n and do nothing with probability $1 - p_n$, independently of the sequence of the previous terms.

Consider $\frac{1}{N} \sum_{n=1}^N X_n$, that is, the empirical frequency of 1’s (“heads”) in the sequence of X_n ’s. We are interested in the asymptotic behavior, in law, of this random variable as $N \rightarrow \infty$.

Since we are interested in limit theorems, we center the variable X_n ; for convenience, we also multiply it by two, thus focus on $Y_n := 2X_n - 1 \in \{-1, +1\}$ instead of X_n . We have

$$Y_n := \begin{cases} -Y_{n-1}, & \text{with probability } p_n; \\ Y_{n-1}, & \text{otherwise.} \end{cases}$$

Note that the sequence $\{Y_n\}$ can be defined equivalently as follows.

Let $Y_n := (-1)^{\sum_{i=1}^n W_i}$, where W_1, W_2, W_3, \dots are independent Bernoulli variables with parameters p_1, p_2, p_3, \dots , respectively, and $p_1 = 1/2$. The number of turns that occurred up to n is $\sum_{i=1}^n W_i$. (Its distribution is *Poisson binomial*).

This representation is important for the proofs, and below are some easy observations it implies. However, it is important to point out that the process Y is also a non-homogeneous Markov chain with state space $\{-1, 1\}$, initial distribution $\mathbb{P}(Y_1 = 1) = \mathbb{P}(Y_1 = -1) = 1/2$, and doubly stochastic symmetric transition matrices $\begin{pmatrix} 1 - p_n & p_n \\ p_n & 1 - p_n \end{pmatrix}, n \geq 2$. Using the symmetry in the definition (or the double stochasticity and induction), $Y_n (n \geq 2)$ has the same distribution as Y_1 , namely **Bernoulli**($1/2$). Hence, the limit theorems in this paper involve particular cases of this two-state Markov chain, but the methods of our proofs rely on the above representation for the random variables $\{Y_n\}$, and not on Markovian techniques.

The following quantity will play an important role: for $1 \leq i < j \leq N$, let

$$e_{i,j} := \prod_{k=i+1}^j (1 - 2p_k).$$

Using the representation for the random variables $\{Y_n\}$, we have

$$Y_j = Y_i \cdot (-1)^{\sum_{k=i+1}^j W_k}, \quad \text{for } 1 \leq i \leq j,$$

and hence if $i = 1$, then we get $\mathbb{E}(Y_j) = e_{1,j} \mathbb{E}(Y_1)$. In particular, since we assumed $p_1 = 1/2$, $\mathbb{E}(Y_1)$ as well as all consecutive $\mathbb{E}(Y_j)$ equal 0 for all $j \geq 2$. In fact, for arbitrary $\{p_n, n \geq 1\}$ satisfying $p_n \neq 1/2, n \geq 2$, the entire sequence $(Y_n)_{n \geq 1}$ is centered in expectation (equivalently, $\mathbb{E}(X_n) = 1/2, n \geq 1$) if and only if $p_1 = 1/2$.

Throughout the paper for $N \geq 1$, we set

$$T_N := X_1 + \dots + X_N, \quad S_N := Y_1 + \dots + Y_N.$$

Then $S_N = 2T_N - N$ and hence limit theorems we establish below for S_N/N will easily imply analogous results for $T_N/N = S_N/(2N) + 1/2$. At a more elementary level, we first observe that S_N is symmetric in distribution around zero (hence its odd moments vanish), and as a result, T_N is symmetric about $N/2$. The symmetry of the law of S_N about zero follows from the symmetry in the definition of the model. In fact, a straightforward calculation gives $\mathbb{E}e^{itS_N} = \mathbb{E} \cos(U_N t)$, where $U_N := 1 + (-1)^{W_2} + (-1)^{W_2+W_3} + \dots + (-1)^{W_2+W_3+\dots+W_N}$.

Using Corr and Cov for correlation and covariance, respectively, one also has

$$\begin{aligned} \text{Corr}(Y_i, Y_j) &= \text{Cov}(Y_i, Y_j) = \mathbb{E}(Y_i Y_j) = \mathbb{E}(-1)^{\sum_{i+1}^j W_k} \\ &= \prod_{i+1}^j \mathbb{E}(-1)^{W_k} = \prod_{k=i+1}^j (1 - 2p_k) = e_{i,j}; \end{aligned} \tag{1}$$

$$\mathbb{E}(Y_j | Y_i) = Y_i \mathbb{E}(-1)^{\sum_{i+1}^j W_k} = e_{i,j} Y_i.$$

Corollary 1 (Correlation estimate) *Assume that $\lim_{k \rightarrow \infty} p_k = 0$ and let $n_* \in \mathbb{N}$ be such that $p_k \leq 1/2$ for $k \geq n_*$. For $n_* \leq i < j$,*

$$\exp\left(-2 \sum_{i+1}^j p_k\right) \cdot \prod_{i+1}^j (1 - r_k) \leq e_{i,j} \leq \exp\left(-2 \sum_{i+1}^j p_k\right),$$

where $r_k := 2p_k^2 e^{2p_k}$, which is tending to zero rapidly.

Furthermore, for any given $C > 1$, there exists an $n_* \in \mathbb{N}$ such that for $n_* \leq i < j$,

$$\exp\left(-2 \sum_{i+1}^j C p_k\right) \leq e_{i,j} \leq \exp\left(-2 \sum_{i+1}^j p_k\right),$$

Proof Use the Remainder Theorem for Taylor series, yielding

$$0 \leq e^{-2p_k} - (1 - 2p_k) \leq 2p_k^2,$$

that is,

$$\exp(-2p_k) \cdot (1 - r_k) \leq 1 - 2p_k \leq \exp(-2p_k),$$

and multiply these inequalities, to get the first statement.

For the second statement, use that for sufficiently small positive x ,

$$e^{-Cx} \leq 1 - x \leq e^{-x}.$$

□

Similarly to (1), if $K = 2m$ is a positive even number, and $i_1 < i_2 < \dots < i_K$ then, using the fact that

$$\begin{aligned} & \sum_{k=1}^{i_1} W_k + \sum_{k=1}^{i_2} W_k + \dots + \sum_{k=1}^{i_K} W_k \\ &= \sum_{k=i_1+1}^{i_2} W_k + \sum_{k=i_3+1}^{i_4} W_k + \dots + \sum_{k=i_{K-1}+1}^{i_K} W_k \pmod{2}, \end{aligned}$$

we obtain that

$$\begin{aligned} \mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_K}) &= \mathbb{E}(-1)^{\sum_1^{i_1} W_k + \sum_1^{i_2} W_k + \dots + \sum_1^{i_K} W_k} \\ &= \mathbb{E}(-1)^{\sum_{i_1+1}^{i_2} W_k + \sum_{i_3+1}^{i_4} W_k + \dots + \sum_{i_{K-1}+1}^{i_K} W_k} \\ &= \mathbb{E}(-1)^{\sum_{i_1+1}^{i_2} W_k} \cdot \mathbb{E}(-1)^{\sum_{i_3+1}^{i_4} W_k} \cdot \dots \cdot \mathbb{E}(-1)^{\sum_{i_{K-1}+1}^{i_K} W_k} \\ &= e_{i_1, i_2} e_{i_3, i_4} \dots e_{i_{K-1}, i_K}. \end{aligned} \tag{2}$$

We close this section with introducing some frequently used notation.

Notation: In the sequel, $\text{Bessel } I_\alpha$ and $\text{Bessel } K_\alpha$ will denote the modified Bessel function of the first kind (or Bessel-I function) and the modified Bessel function of the second kind (or Bessel-K function), respectively.

Writing out these functions explicitly, one has

$$\text{Bessel } I_\alpha(x) = \sum_{m=0}^{\infty} \frac{1}{m! \Gamma(m + \alpha + 1)} \left(\frac{x}{2}\right)^{2m+\alpha},$$

and

$$\text{Bessel } K_\alpha(x) = \frac{\pi}{2} \frac{\text{Bessel } I_{-\alpha}(x) - \text{Bessel } I_\alpha(x)}{\sin(\alpha\pi)},$$

if α is not an integer (otherwise it is defined through limits), where Γ is Euler’s gamma function. See, e.g., Sections 9–10 in [1], and formula (6.8) in [2].

2 Review of Literature and Comparison with Our Results

As the Associate Editor kindly pointed out for us, the problem has a history, going back to at least the 1950s. Below we give a review of the relevant achievements in the past and compare them to the results presented in this paper.

The case $p_n = 1/n$ was already introduced in R. Dobrushin’s thesis¹ in the 1950s and it is attributed to Bernstein (see [4]). Dobrushin did not seem to explicitly identify

¹ He considered his thesis work a continuation of the work of Markov, Bernstein, Sapogov, and Linnik on time inhomogeneous Markov chains.

the limiting frequency of heads with the uniform distribution. However, in their 2007 paper [3] Dietz and Sethuraman proved that for the more general $p_n = a/n$ case ($a > 0$), the limiting frequency is **Beta**(a, a)—see their Theorems 1.3 and 1.4. (They consider a state space consisting of $m \geq 2$ points and so they treat the more general Dirichlet distributions). Therefore, in the $p_n = a/n$ case, our contribution is providing a different proof only. The proof in [3] is significantly longer and more complicated than ours; however, we only consider $m = 2$.

The case $p_n = a/n^\gamma$ is also treated in [3], albeit only the Weak Law of Large Numbers (SLLN when $0 < \gamma < 1/2$). The authors note that simulations suggest that actually a.s. convergence might hold also on the range $1/2 \leq \gamma < 1$; in our case, we prove this statement in our Corollary 2. Fluctuations about the mean are not considered in [3] though.

The situation with the Central Limit Theorem is more interesting in this case. First, Dobrushin’s Central Limit Theorem for inhomogeneous Markov chains (Theorem 1.1 in [11]) only provides the statement in the sense that, after centering and normalizing with the standard deviation, the limit is standard normal. We did not find, however, any result in the literature identifying the order of the standard deviation, which we do provide. (See the estimates on p. 411 and also Corollary 15 on p. 421 in [10]).

Secondly, and more importantly, Dobrushin’s Theorem only applies to the case when $0 < \gamma < 1/3$. Although the condition given in that result (formula (1.3) in [11]) is known to be optimal (see Section 2 in [11]), this is only true in the very general setting in which the theorem is stated. It is therefore interesting, we believe, that we also prove the CLT in the $1/3 \leq \gamma < 1$ case excluded in Dobrushin’s result. (In [10], Dobrushin’s condition was improved by Peligrad, but it is still not applicable in our case when $1/3 \leq \gamma < 1$: formula (7) in [10] is actually more stringent than the Dobrushin condition (9)).

To the best of our knowledge, Theorem 3 is completely new.

3 Supercritical Cases

First, if $\sum_n p_n < \infty$, then by the Borel–Cantelli lemma, only finitely many turns will occur a.s.; therefore, the X_j ’s will eventually become all ones or all zeros, and hence

$$\frac{T_N}{N} \rightarrow \zeta \text{ a.s.,}$$

where $\zeta \in \{0, 1\}$. By the symmetry of the definition with respect to heads and tails (or, by the bounded convergence theorem), ζ is a **Bernoulli**($1/2$) random variable.

4 The Critical Case

Fix $a > 0$, and let

$$p_n = \frac{a}{n}, \quad n \geq n_0$$

for some $n_0 \in \mathbb{N}$. Denote by $\text{Beta}(a, a)$ the symmetric (around the point $1/2$) Beta distribution with $a > 0$, with density

$$f_{\text{Beta}(a,a)}(x) = \frac{[x(1-x)]^{a-1}}{B(a, a)}$$

on the unit interval (the normalizing constant is $B(a, a) := \Gamma^2(a)/\Gamma(2a)$, using Euler’s Gamma function), and moment generating function

$$M_{\text{Beta}(a,a)}(t) = 1 + \sum_{k=1}^{\infty} \left(\prod_{l=1}^k \frac{a+l-1}{2a+l-1} \right) \frac{t^k}{k!} = e^{t/2} \left(\frac{t}{4} \right)^{\frac{1}{2}-a} \Gamma\left(a + \frac{1}{2}\right) \text{Bessel} I_{a-\frac{1}{2}}\left(\frac{t}{2}\right). \tag{3}$$

Theorem 1 *The law of $\frac{1}{N} \sum_{i=1}^N X_i$ converges to $\text{Beta}(a, a)$ as $N \rightarrow \infty$.*

Remark 1 It turns out that the convergence in distribution cannot be strengthened to convergence in probability; see, e.g., the example in Section 2 of [7].

Proof We will verify the statement by analyzing the moments of S_N . The odd moments are all zero from the symmetry of S_N around 0; on the other hand, for even K we can use the multinomial theorem:

$$S_N^K = I + K! \sum_{1 \leq i_1 < i_2 < \dots < i_K \leq N} \mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_K}),$$

where I stands for the sum of those products in which not all terms are different. Note that $|Y_i^l| \leq 1$ for any $l \geq 1$ (and $Y_i^l \equiv 1$ for l even). Therefore, $|I| \leq m(N, K)$, where $m(N, K)$ is the number of such products. But $m(N, K) \leq N \cdot N^{K-2} = N^{K-1}$, because each such product can be written (not uniquely) as $Y_{i_\ell}^2 \cdot Y_{i_1} Y_{i_2} \dots Y_{i_{K-2}} = Y_{i_1} Y_{i_2} \dots Y_{i_{K-2}}$, where the numbers $i_\ell, i_1, \dots, i_{K-2}$ are between 1 and N and are not necessarily distinct. Hence, also using (2), we get

$$\begin{aligned} \mathbb{E}S_N^K &= K! \sum_{1 \leq i_1 < i_2 < \dots < i_K \leq N} \mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_K}) + \mathcal{O}(N^{K-1}) \\ &= K! \sum_{1 \leq i_1 < i_2 < \dots < i_K \leq N} e_{i_1, i_2} e_{i_3, i_4} \dots e_{i_{K-1}, i_K} + \mathcal{O}(N^{K-1}). \end{aligned} \tag{4}$$

Let us now analyze the elements in the sum above. From (1), for $j > i > \max\{2a, n_0\}$ we have

$$\begin{aligned} e_{i,j} &= \exp \left\{ \sum_{n=i+1}^j \log \left(1 - \frac{2a}{n} \right) \right\} = \exp \left\{ \mathcal{O} \left(\frac{j-i}{i^2} \right) - 2a \sum_{n=i+1}^j \frac{1}{n} \right\} \\ &= \exp \left\{ \mathcal{O} \left(\frac{j-i}{i^2} \right) - 2a \log \left(\frac{j}{i} \right) \right\} = \frac{i^{2a}}{j^{2a}} \cdot \left(1 + \mathcal{O} \left(\frac{j-i}{i^2} \right) \right). \end{aligned} \tag{5}$$

Consequently, (4) can be approximated as

$$\begin{aligned} \mathbb{E}S_N^K &= K! \sum_{1 \leq i_1 < i_2 < \dots < i_K \leq N} \frac{i_1^{2a}}{i_2^{2a}} \cdot \frac{i_3^{2a}}{i_4^{2a}} \cdot \dots \cdot \frac{i_{K-1}^{2a}}{i_K^{2a}} + \mathcal{O}(N^{K-1}) \\ &= \frac{K! N^K (1 + \mathcal{O}(N^{-1}))}{(1 + 2a) \cdot 2 \cdot (3 + 2a) \cdot 4 \cdot \dots \cdot (K - 1 + 2a) \cdot K} + \mathcal{O}(N^{K-1}) \end{aligned} \tag{6}$$

(the contribution from the terms where $i_1 \leq \max\{2a, n_0\}$ as well as the other remainder terms in the formula for $e_{i,j}$ is of order at most N^{K-1}). Since we are working on a compact interval, we may conclude (see, e.g., Section 2, Exercise 3.27 in [5]) that $S_N/N \rightarrow \xi_a$ in distribution, where ξ_a is distributed on $[-1, 1]$ and has the following moments:

$$\mathbb{E} \left[\xi_a^K \right] = \begin{cases} 0, & K \text{ is odd;} \\ \frac{(2m)!}{2^m \cdot m! \cdot (2a + 1)(2a + 3) \dots (2a + (2m - 1))}, & K = 2m \text{ is even,} \end{cases}$$

which, for even moments, can be equivalently written as

$$\mathbb{E} \left[\xi_a^{2m} \right] = \frac{(2m)! \Gamma(a + 1/2)}{2^{2m} \Gamma(m + a + 1/2)}. \tag{7}$$

The moment generating function of ξ_a is

$$\begin{aligned} M_a(t) &= \mathbb{E}e^{t\xi_a} = 1 + \sum_{m=1}^{\infty} \frac{t^{2m}}{2^m \cdot m! \cdot \prod_{i=1}^m (2a + 2i - 1)} \\ &= \text{Bessel } I_{a-1/2}(t) \Gamma(a + 1/2) (t/2)^{1/2-a}. \end{aligned}$$

Let $\zeta_a := (\xi_a + 1)/2$. We know that $\frac{1}{N} \sum_i^N X_i \rightarrow \zeta_a$ in distribution, and using (3),

$$\begin{aligned} \mathbb{E}e^{t\zeta_a} &= e^{t/2} M_a(t/2) = e^{t/2} \text{Bessel } I_{a-1/2}(t/2) \Gamma(a + 1/2) (t/4)^{1/2-a} \\ &= M_{\text{Beta}(a,a)}(t), \end{aligned}$$

completing the proof. □

Remark 2 (Particular cases and densities) Note that in particular, for $a = 1$, $a = 1/2$ and $a = 3/2$, the limiting law $\text{Beta}(a, a)$ of the relative frequencies in Theorem 1 is $\text{Uniform}([0, 1])$, the arcsine law, and the transformed semicircle law² on $[0, 1]$, respectively.

Turning to S_N/N , the transformation $x \mapsto 2x - 1$ yields that $\lim_{N \rightarrow \infty} \text{Law}(S_N/N)$ equals to $\text{Uniform}([-1, 1])$, the transformed arcsine law on $[-1, 1]$ and Wigner’s semicircle law on $[-1, 1]$, respectively. Concerning the corresponding densities on $[-1, 1]$, we have the following explicit formulas.

² The density is a semi-ellipse.

- **Transformed Arcsine Law:** Let $a = 1/2$. Then

$$\mathbb{E}e^{t\xi} = \sum_{m=0}^{\infty} \frac{t^{2m}}{(2^m \cdot m!)^2} = \text{Bessel } I_0(t) = \frac{1}{\pi} \int_0^\pi e^{t \cos(\theta)} d\theta;$$

consequently (see, e.g., [1], formula 29.3.60) $\xi_{1/2}$ has the transformed arcsine density

$$f_{\xi_{1/2}}(x) = \begin{cases} \frac{1}{\pi\sqrt{1-x^2}}, & -1 < x < 1; \\ 0, & \text{otherwise.} \end{cases}$$

- **Wigner’s semicircle law on $[-1, 1]$:** Let $a = 3/2$. Then

$$\begin{aligned} \mathbb{E}e^{t\xi} &= \frac{2 \text{Bessel } I_1(t)}{t} = \frac{1}{\pi} \int_0^\pi e^{t \cos(\theta)} \cos(\theta) d\theta \\ &= \text{Bessel } I_0(t) - \text{Bessel } I_2(t); \end{aligned}$$

consequently (see [1], formula 9.6.19) $\xi_{3/2}$ has the Wigner semicircle density

$$f_{\xi_{3/2}}(x) = \begin{cases} \frac{2}{\pi} \sqrt{1-x^2}, & -1 < x < 1; \\ 0, & \text{otherwise.} \end{cases}$$

- **General case:** The density of ξ_a is given by

$$f_{\xi_a}(x) = \frac{\Gamma(a + 1/2)}{\Gamma(a) \sqrt{\pi}} (1 - x^2)^{a-1}$$

for $-1 < x < 1$. Indeed, for $m \in \mathbb{N}$ we have

$$\begin{aligned} \int_{-1}^1 x^{2m} (1 - x^2)^{a-1} dx &= \int_0^1 y^{m-1/2} (1 - y)^{a-1} dy \\ &= \text{Beta}(m + 1/2, a) = \frac{\Gamma(m + 1/2)\Gamma(a)}{\Gamma(m + a + 1/2)}, \end{aligned}$$

which is consistent with the moments $\mathbb{E}\xi^{2m}$ given by (7).

5 Subcritical Case

Now fix $\gamma, a > 0$, and let

$$p_n = \frac{a}{n^\gamma}, \quad n \geq n_0$$

for some n_0 . Note that $\gamma > 1$ corresponds to the supercritical case studied in Sect. 3, so from now on assume $0 < \gamma < 1$.

Theorem 2 *The law of $S_N/N^{(1+\gamma)/2}$ converges to Normal(0, σ^2) where*

$$\sigma := \frac{1}{\sqrt{a(1 + \gamma)}}. \tag{8}$$

Proof Let $\eta_{a,\gamma,N} := S_N/N^{(1+\gamma)/2}$. Let $\eta_{a,\gamma}$ be normally distributed with variance σ^2 where σ is as in (8). We will prove that $\lim_{N \rightarrow \infty} \eta_{a,\gamma,N} = \eta_{a,\gamma}$ in law.

In the proof we will use that

$$\mathbb{E}\eta_{a,\gamma}^K = \begin{cases} 0, & \text{if } K \text{ is odd;} \\ \sigma^K (K - 1)!!, & \text{if } K \text{ is even} \end{cases} \tag{9}$$

(see, e.g., [5], Section 2, Exercise 3.15).

Let A ($A > a$) be a given constant. By Corollary 1, there exists an $n_* = n_*(a, A, \gamma)$ (w.l.o.g. assume that $n_* > n_0$) such that for $j > i > n_*$

$$\exp \left\{ -2A \sum_{n=i+1}^j \frac{1}{n^\gamma} \right\} \leq e_{i,j} \leq \exp \left\{ -2a \sum_{n=i+1}^j \frac{1}{n^\gamma} \right\}.$$

Using the fact that $x^{-\gamma}$ is decreasing and bounding the sum by the integral, we have

$$\frac{(j + 1)^{1-\gamma} - (i + 1)^{1-\gamma}}{1 - \gamma} = \int_{i+1}^{j+1} \frac{dx}{x^\gamma} \leq \sum_{n=i+1}^j \frac{1}{n^\gamma} \leq \int_i^j \frac{dx}{x^\gamma} = \frac{j^{1-\gamma} - i^{1-\gamma}}{1 - \gamma},$$

yielding

$$\exp \left(-\frac{2A}{1-\gamma} [j^{1-\gamma} - i^{1-\gamma}] \right) \leq e_{i,j} \leq \exp \left(-\frac{2a}{1-\gamma} [(j + 1)^{1-\gamma} - (i + 1)^{1-\gamma}] \right), \tag{10}$$

that is, using the shorthand $c := 2a(1 - \gamma)^{-1}$ and $d := 2A(1 - \gamma)^{-1}$,

$$\exp \left(-d[j^{1-\gamma} - i^{1-\gamma}] \right) \leq e_{i,j} \leq \exp \left(-c[(j + 1)^{1-\gamma} - (i + 1)^{1-\gamma}] \right).$$

One can check that $\sup_N \mathbb{E}\eta_{a,\gamma,N}^2 < \infty$ (see the computation below with $m = 1$), and thus Chebyshev’s inequality implies that $(\eta_{a,\gamma,N})_{N \geq 1}$ is a tight sequence of random variables. Hence, it is enough to show that each subsequential limit is the same.

Assume that $(N_l)_{l \geq 1}$ is a subsequence and $\lim_{l \rightarrow \infty} \text{Law}(\eta_{a,\gamma,N_l}) = \mathcal{L}$. Since

$$\left| \frac{Y_1 + \dots + Y_{n_*-1}}{N_l^{\frac{1+\gamma}{2}}} \right| \leq \frac{n_* - 1}{N_l^{\frac{1+\gamma}{2}}},$$

one has $\mathcal{L} = \lim_{l \rightarrow \infty} \mathcal{L}_{N_l, A}$ too, where

$$\mathcal{L}_{N_l, A} := \text{Law} \left(\frac{Y_{n_*} + \dots + Y_{N_l}}{N_l^{\frac{1+\gamma}{2}}} \right),$$

and in fact, this limit must be the same for any $A > a$ (and corresponding $n_* = n_*(a, A, \gamma)$). Informally, this just means that we may throw away a finite chunk of the sequence of Y_i 's' (at the beginning) without affecting its limit.

Let us denote the even moments of \mathcal{L} by $M_{2m} \in [0, \infty]$, $m \geq 1$, while we note again that the odd moments must be zero by symmetry. Also, $M_{N_l, A, K}$ will denote the K th moment under $\mathcal{L}_{N_l, A}$.

We will show below that for a fixed $A > a$ and $K = 2m, m \geq 1$,

$$\begin{aligned} \frac{(2m - 1)!!}{[A(1 + \gamma)]^m} &\leq \liminf_{l \rightarrow \infty} M_{N_l, A, K} = \liminf_{l \rightarrow \infty} \mathbb{E} \left[\frac{Y_{n_*} + \dots + Y_{N_l}}{N_l^{\frac{1+\gamma}{2}}} \right]^K \\ &\leq \limsup_{l \rightarrow \infty} M_{N_l, A, K} = \limsup_{l \rightarrow \infty} \mathbb{E} \left[\frac{Y_{n_*} + \dots + Y_{N_l}}{N_l^{\frac{1+\gamma}{2}}} \right]^K \leq \frac{(2m - 1)!!}{[a(1 + \gamma)]^m}. \end{aligned} \tag{11}$$

Once (11) is shown, it will follow from the upper estimate and from the relation $\mathcal{L} = \lim_{l \rightarrow \infty} \mathcal{L}_{N_l, A}$ for all $A > a$ that

$$\lim_{l \rightarrow \infty} M_{N_l, A, K} = M_K \tag{12}$$

for all $K \geq 1$ and all $A > a$. Since (11) holds for any $A > a$, letting $A \downarrow a$ and using (11) and (12), one has that in fact

$$M_K = \frac{(2m - 1)!!}{[a(1 + \gamma)]^m}.$$

In summary, we obtain that for any fixed $A > a$,

$$\lim_{l \rightarrow \infty} M_{N_l, A, K} = \frac{(2m - 1)!!}{[a(1 + \gamma)]^m}. \tag{13}$$

At the same time, we recall that the normal distribution is uniquely determined by its moments, and therefore the convergence toward a normal law is implied by the convergence of all the moments (see, e.g., [5], Section 2.3.e). In our case, (13) along with (9) implies $\mathcal{L} = \lim_{l \rightarrow \infty} \mathcal{L}_{N_l, A} = \text{Normal}(0, \sigma^2)$. Therefore, it only remains to prove (11).

Let us start with the upper estimate in (11). For $K = 2m$, one has

$$\mathbb{E}[Y_{n_*} + \dots + Y_N]^K = I + K! \sum_{n_* \leq i_1 < i_2 < \dots < i_K \leq N} \mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_K})$$

where I are lower-order terms, as it will be shown below. Using (2) along with (10), we may continue with

$$\leq I + K! \sum_{n_*+1 \leq i_1 < i_2 < \dots < i_K \leq N+1} \exp(cU_{i_1, \dots, i_k}), \tag{14}$$

where

$$U_{i_1, \dots, i_k} := i_1^{1-\gamma} - i_2^{1-\gamma} + i_3^{1-\gamma} - i_4^{1-\gamma} + \dots + i_{K-1}^{1-\gamma} - i_K^{1-\gamma}.$$

By the calculation in the ‘‘Appendix,’’ the RHS of (14) is

$$I + K! \times \frac{N^{K(1+\gamma)/2}}{c^m(1-\gamma^2)^m m!} \cdot (1 + o(1)).$$

By the same token,

$$\begin{aligned} \mathbb{E}[Y_{n_*} + \dots + Y_N]^K &\geq I + K! \sum_{1 \leq i_1 < i_2 < \dots < i_K \leq N} \exp(dU_{i_1, \dots, i_k}) \\ &= I + \frac{K! N^{K(1+\gamma)/2}}{d^m(1-\gamma^2)^m m!} \cdot (1 + o(1)). \end{aligned}$$

The reason the remaining terms, collected in I , are of lower order is as follows. Apart from the already estimated term, in the expansion for $\mathbb{E}(Y_{n_*} + \dots + Y_N)^K$ for $r = 1, 2, \dots, K - 1$ we also have to sum up the terms of the type

$$\begin{aligned} &\mathbb{E}(Y_{i_1}^{p_1} Y_{i_2}^{p_2} \dots Y_{i_r}^{p_r}), \text{ where } n_* \\ &\leq i_1 < \dots < i_r \leq N, \text{ all } p_j \geq 1, \text{ and } p_1 + p_2 + \dots + p_r = K. \end{aligned}$$

Since $Y_i = \pm 1$, and thus $Y_i^p = 1$ if p is even and $Y_i^p = Y_i$ if p is odd, it suffices to estimate only the sums

$$\mathcal{R}(r; \ell_1, \dots, \ell_r; N; K; \gamma) := \sum \mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_r}),$$

where the summation is taken over all sets (i_1, \dots, i_r) such that $i_{k+1} \geq i_k + \ell_k$, $1 \leq \ell_k \leq K$, for all k , $i_1 \geq 1$ and $i_r \leq N$. However, since $r \leq K - 1$, each of the sums $\mathcal{R}(r; \ell_1, \dots, \ell_r; N; K; \gamma)$ is at most of order $N^{r(1+\gamma)/2} \leq N^{(K-1)(1+\gamma)/2}$, precisely by the same arguments which were used to estimate the sum in (14). The number of those sums can be large, as it is the number of integer partitions of K , but it depends only on K and does not increase with N .

Consequently, for $m \geq 1$ we have

$$(I) \leq \liminf_{l \rightarrow \infty} \mathbb{E} \left[\frac{Y_{n_*} + \dots + Y_{N_l}}{N_l^{\frac{1+\gamma}{2}}} \right]^K \leq \limsup_{l \rightarrow \infty} \mathbb{E} \left[\frac{Y_{n_*} + \dots + Y_{N_l}}{N_l^{\frac{1+\gamma}{2}}} \right]^K \leq (II),$$

where

$$(II) := \frac{(2m)!}{[c(1 - \gamma^2)]^m m!} = \frac{(2m)!}{[2a(1 + \gamma)]^m m!} = \frac{(2m)!}{2^m m!} \cdot [a(1 + \gamma)]^{-m} \\ = \frac{(2m - 1)!!}{[a(1 + \gamma)]^m},$$

and by similar computation,

$$(I) := \frac{(2m - 1)!!}{[A(1 + \gamma)]^m}.$$

The proof is complete. □

6 When Does the Law of Large Numbers Hold for General Sequences $\{p_n\}$?

A natural question to ask is when does S_N obey the Strong (Weak) Law of Large Numbers. The following result gives a partial answer.

For a positive even number K , introduce the shorthand

$$E(N, K) := N^{-K} \sum_{1 \leq i_1 < i_2 < \dots < i_K \leq N} e_{i_1, i_2} e_{i_3, i_4} \dots e_{i_{K-1}, i_K},$$

and note that

$$\text{var} \left(\frac{S_N}{N} \right) = \frac{1}{N} + 2 E(N, 2).$$

The first condition in the next theorem may look reminiscent of Kolmogorov’s sufficient condition for the Strong Law of Large Numbers.

Theorem 3 (a) (Strong Law) *Assume that at least one of the following two conditions holds. (C1)*

$$\sum_N \frac{E(N, 2)}{N} < \infty.$$

(C2) For some even number K ,

$$\sum_N E(N, K) < \infty. \tag{15}$$

Then SLLN holds, that is, $S_N/N \rightarrow 0$ a.s.

(b) (Weak Law) The WLLN holds if and only if for each positive even number K ,

$$\lim_{N \rightarrow \infty} E(N, K) = 0.$$

(c) (no LLN) If for each positive even number K ,

$$\exists \lim_{N \rightarrow \infty} E(N, K) =: \mu_K > 0,$$

and

$$\sum_{K \text{ even}} \frac{1}{\mu_K^{1/K}} = \infty, \tag{16}$$

then the Law of Large Numbers breaks down, and in fact, $\mathbf{Law}(S_N/N)$ converges to a law which has zero odd moments and even moments $\{\mu_K\}$.

Note that (16) is the so-called *Carleman condition*, guaranteeing that the μ_K s correspond to at most one probability law (see Theorem 3.11, Section 2, in [5]).

Proof We will use the facts about the *method of moments* for weak convergence discussed in the proof of Theorem 2, along with the fact that from (4), it follows that

$$\begin{aligned} \mathbb{E} \left(\frac{S_N}{N} \right)^K &= \frac{K!}{N^K} \sum_{1 \leq i_1 < i_2 < \dots < i_K \leq N} \mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_K}) + \mathcal{O}(1/N) \\ &= K! E(N, K) + \mathcal{O} \left(\frac{1}{N} \right). \end{aligned} \tag{17}$$

(a) Let us consider the two assumptions separately:

Under (C1), the statement follows from Theorem 1 in [9], as $e_{i,j} = \text{Cov}(Y_i, Y_j)$. Under (C2), along the lines of Theorem 6.5 in Section 1 of [5], we note that for $\varepsilon > 0$, one has

$$\mathbb{P} \left(\left| \frac{S_N}{N} \right| > \varepsilon \right) \leq \frac{\mathbb{E} S_N^K}{\varepsilon^K N^K}$$

by the Markov inequality (recall that K is even). Since, by (6), the expression on the left-hand side of (15) is the leading order term in $\mathbb{E} S_N^K$, by (15), we have $\sum_N \mathbb{P}(|S_N/N| > \varepsilon) < \infty$, and thus, by the Borel–Cantelli lemma, $\mathbb{P}(|S_N/N| > \varepsilon \text{ i.o.}) = 0$, which implies the statement.

- (b) Since $|S_N/N| \leq 1$, we know that S_N/N converges to zero in law (i.e., in probability, since the limit is deterministic) if and only if all its moments converge to zero. (One direction is to realize that the k th moment is the same as $\mathbb{E}f^K(S_N/N)$, where $f(x) = x$ on the unit interval, $f(x) := 1$ for $x > 1$ and $f(x) = -1$ for $x < -1$; then f is bounded and continuous. The other direction is also known, since the deterministically zero distribution is uniquely determined by its moments). By symmetry, it is enough to check the even moments, for which we know (17). The statement then follows from the fact that the remainder term is $\mathcal{O}(1/N)$.
- (c) Assume that the conditions in (c) hold. Since the moments of S_N/N converge (the odd moments are zero by symmetry), the corresponding laws are tight and, by the Carleman condition, all subsequential limits are the same. That is, as $N \rightarrow \infty$, $\text{Law}(S_N/N)$ converges to a law with moments given by μ_K , and since $\mu_K > 0$, the limit cannot be deterministically zero.

The following corollary proves a conjecture in [3] when $1/2 \leq \gamma < 1$.

Corollary 2 *When $p_n = a/n^\gamma$ with $0 < \gamma < 1$, $a > 0$ (subcritical case), the Strong Law of Large Numbers holds: $S_N/N \rightarrow 0$, \mathbb{P} -a.s. (Observe that in view of Theorem 2, convergence in probability is immediate).*

Proof We have seen in the proof of Theorem 2 that all moments, and in particular the second moment of the ratio $S_N/N^{\frac{1+\gamma}{2}}$, converge as $N \rightarrow \infty$. Thus $\mathbb{E}(S_N^2) \sim N^{1+\gamma}$ and

$$E(N, 2) = \frac{1}{2} \left(\mathbb{E}(S_N^2/N^2) - 1/N \right) \sim N^{\gamma-1},$$

hence condition (C1) of Theorem 3 is satisfied. (Here $f_N \sim g_N$ means that $\lim_{N \rightarrow \infty} f_N/g_N$ exists and is positive). □

Corollary 3 (Monotonicity) *If WLLN holds for the sequence $\{p_n\}$, then it also holds for the sequence $\{\hat{p}_n\}$, whenever $\hat{p}_n \geq p_n$ for all n .*

Proof This follows from Theorem 3(b) along with the definition of $E(N, K)$ in terms of the $e_{i,j}$ and the fact that $e_{i,j} = \prod_{k=i+1}^j (1 - 2p_k)$ is monotone decreasing in the p_k 's for each given $1 \leq i < j$. □

7 Giving Up Symmetry

Now we will show that, in the supercritical case as well as in the setups of Theorem 1 and of Theorem 2, the initial condition being symmetric (i.e., X_1 is equally likely to be 0 or 1) is actually not essential for the limiting distributions.

Thus, in this section we assume w.l.o.g. that $X_1 \equiv 0$ and thus $Y_1 \equiv 1$ and $Y_k = (-1)^{W_2 + \dots + W_k}$, $k \geq 2$.

In the supercritical case, we will have again $T_N/N \rightarrow \zeta \in \{0, 1\} \sim \text{Bernoulli}(q)$ a.s., but because of the lack of symmetry, we can no longer claim that $q = 1/2$. Our next statement, the proof of which may already be known, gives the exact value of q

for any sequence of $\{p_n\}$. In particular, if at least one of p_i 's is $1/2$, then $q = 1/2$, which is already clear from the symmetry.

Proposition 1 *Let $e_{1,\infty} = \prod_{i=2}^{\infty}(1 - 2p_i)$, consistently with our previous definition. Then*

$$q = \mathbb{P}(\zeta = 1) = \frac{1 + e_{1,\infty}}{2}.$$

Proof Since we are in the supercritical regime, only finitely many turns occur a.s., and hence $Y_n = Y_\infty \in \{-1, 1\}$ for all large n ; as a result, $Y_n \rightarrow Y_\infty$ a.s. Hence, using Cesàro mean, $S_N/N \rightarrow Y_\infty$ a.s. as well. By the Bounded Convergence Theorem, $\mathbb{E}Y_\infty = \lim_{n \rightarrow \infty} \mathbb{E}Y_n = e_{1,\infty}$. Since $Y_\infty = 2\zeta - 1$ and $\mathbb{E}\zeta = q$, we have $2q - 1 = e_{1,\infty}$. □

In the critical case, Eq. (1) still holds and so does (2) for even K , but for odd $K = 2m + 1$ we have

$$\mathbb{E}(Y_{i_1} Y_{i_2} \dots Y_{i_K}) = e_{0,i_1} e_{i_2,i_3} \dots e_{i_{K-1},i_K}. \tag{18}$$

The calculation (6) remains valid for even K ; however, if K is odd, one cannot claim any more that $\mathbb{E}S_N^K = 0$. At the same time, a calculation similar to (6) immediately shows that if $K = 2m + 1$, then $\mathbb{E}S_N^K = \mathbb{E}S_N^{2m+1} = \mathcal{O}(N^{2m}) = o(N^K)$, and hence the rescaled odd moments tend to zero, while the even moments are the same as in the original model. Hence the limiting distribution must be the same.

Similar argument holds for the subcritical case as well. Indeed, the even moments $\mathbb{E}S_N^K$ remain the same, while the odd moments for $K = 2m + 1$ will be $\mathbb{E}S_N^{2m+1} = \mathcal{O}(N^{m(1+\gamma)}) = o(N^{K(1+\gamma)/2})$ due to (18) and the result from the ‘‘Appendix.’’

8 Further Heuristic Arguments and a Conjecture

In this section, we omit the details of some calculations—the reader can find them in the preprint of this paper [6].

To avoid ambiguity, by the ‘‘classical CLT’’ we mean the situation where, after normalizing with the standard deviation, the limit has a standard normal distribution, and the standard deviation itself is of order \sqrt{N} ; a ‘‘nonstandard CLT’’ will mean that the standard deviation (and thus the fluctuation) is of a different order.

Consider a sum of $N \geq 1$ variables having the same law with finite variance. As is well known, the two ‘‘extreme cases’’ for a sum are the independent case, when the variance is linear and one gets the Central Limit Theorem, and the other one is when all the variables are identical and the variance grows like N^2 . By analogy then, (after recalling that in our model

$$\text{Var}(S_N) = N + 2N^2 E(N, 2)$$

holds), it seems that the first crucial question is whether

$$E(N, 2) = \mathcal{O}(1/N), \quad N \rightarrow \infty \tag{19}$$

is still the case. If (19) is true, then $\text{Var}(S_N)$ is of order N , and one can expect that the classical CLT holds. This happens when $p_n \equiv p \in (0, 1)$.

In a situation when (19) fails, one should know at least if

$$E(N, 2) = o(1) \tag{20}$$

holds. Indeed, we know from Theorem 3(b) that the exact criterion for WLLN to hold is that

$$E(N, K) = o(1), \text{ for all } K = 2m. \tag{21}$$

In light of this, we make the following conjecture.

Conjecture 1 *Let $p_n \in [0, 1]$ for $n \geq 1$, and assume that (21) holds.*

- (i) *If (19) holds for $\{p_n\}$, then the proportion of heads obeys classical CLT. (see Example 1 below).*
- (ii) *If (19) fails for $\{p_n\}$, then there is a nonstandard CLT for the proportion. (see example 2 below).*

If (21) fails for $\{p_n\}$, then WLLN is no longer valid for the proportion, that is, the proportion is not concentrated about $1/2$ at all. (see examples 3, 4 below).

Examples Supporting the Discussion and Conjecture 1

In the examples below, the deviations from the classical CLT are becoming more marked as we go from Example 2 to Example 3 to Example 4. Recall that $S_N := Y_1 + \dots + Y_N$ and $T_N := X_1 + \dots + X_N$ with $T_N = (Y_N + 1)/2$; the frequency of heads is T_N/N .

Example 1 (Markov chain CLT) Consider the case $p_n = c$ for all $n \geq 1$, where $0 < c < 1$. If $c = 1/2$, we get an i.i.d. sequence of $+1$'s and -1 's and the classical CLT applies.

Now assume $c \neq 1/2$. Then the outcomes are not independent. Indeed, denoting $\kappa := 1 - 2c \in (-1, 1)$, we have

$$N^2 E(N, 2) = \frac{\kappa(N - 1)}{1 - \kappa} - \frac{\kappa^2 (1 - \kappa^{N-1})}{(1 - \kappa)^2}.$$

Therefore, the variance is still of order N , but the constant has changed. Recall that $\text{Cov}(Y_i, Y_j) = e_{i,j} = \kappa^{j-i}$, and, following [8], define $\sigma_c^2 := 1 + 2 \sum_{i=1}^{\infty} \text{Cov}(Y_i, Y_j) = \frac{1-c}{c}$, assuming that $Y_1 \sim \text{Bernoulli}(1/2)$. In this case,

since we are dealing with a time homogeneous Markov chain, it is well known (see [8]) that

$$\text{Law} \left(\frac{S_N}{\sqrt{N}} \right) \rightarrow \text{Normal} \left(0, \sigma_c^2 \right), \text{ i.e., } \text{Law} \left(\frac{T_N - N/2}{\sqrt{N}} \right) \rightarrow \text{Normal} \left(0, \frac{\sigma_c^2}{4} \right).$$

Therefore, unless $c = 1/2$, the classical CLT is slightly changed, since $\text{Var}(S_n) \sim \sigma_c^2 n$. It is also clear that the limiting normal variance can be arbitrarily large when c is sufficiently small and thus turns occur very rarely. On the other hand, it can be arbitrarily small if c is sufficiently close to 1 and thus turns occur very frequently. (If the turns were *certain*, then the limiting variance would vanish of course).

Example 2 (Classical CLT breaks down) Consider the case $p_n := a/n^\gamma$ with $0 < \gamma < 1$. Then

$$N^2 E(N, 2) \cong \frac{N^{\gamma+1}}{2a(1 + \gamma)},$$

that is $\text{Var}(S_N)$ is of order $N^{\gamma+1}$, and the power is strictly between 1 and 2. Now (21) is true, WLLN is still in force, and S_N/N is still around zero (the proportion of heads is around 1/2). But (19) is false. The closer γ to 1, the more the situation differs from the classical CLT. We now have a nonstandard CLT, with larger than classical fluctuations.

Example 3 (LLN breaks down) Consider the case when $p_n = 1/n$. Then $e_{i,j} = \frac{(i-1)i}{(j-1)j}$. Consequently,

$$N^2 E(N, 2) = \binom{N + 1}{2}$$

is of order N^2 , that is, (19) and even (21) are false, causing The Law of Large Numbers to break down, and S_N/N is no longer around zero. This means that the correlation is as strong as in the case of identical variables, and the fluctuations are now of order N , destroying LLN. Similar is the situation when $p_k = \frac{a}{k}$ with $a > 0$.

In terms of the relative frequency of heads, instead of being around the $\delta_{1/2}$ distribution, now it tends to the **Beta**(a, a) distribution.

Example 4 (Extreme limit) Consider the case when $\sum_n p_n < \infty$. Then $\liminf_{N \rightarrow \infty} E(N, 2) > 0$ holds (hence (19) and even (21) are false).

Indeed, as we know, the limit of S_N/N is “extreme”: $\frac{1}{2}(\delta_{-1} + \delta_1)$, which is as far away from δ_0 as possible! (i.e., **Beta**(0, 0) = $\frac{1}{2}(\delta_0 + \delta_1) \equiv$ **Bernoulli**(1/2) for the frequencies of heads).

We conclude this section with an open problem.

Problem 1 (*Monotonicity for SLLN*) Is it true that if SLLN holds for the sequence $\{p_n\}$, then it also holds for the sequence $\{\hat{p}_n\}$, whenever $\hat{p}_n \geq p_n$ for all n ?

The corresponding statement for WLLN is true by Corollary 3.

Acknowledgements We would like to thank the Associate Editor and the referee for many useful suggestions and recommendations on how to improve the present paper. The Associate Editor’s remarks about the history of the problem and the suggested references were especially valuable.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

In this appendix, we will estimate the quantity

$$Q(n_*, N) := \sum_{n_* \leq i_1 < i_2 < \dots < i_K \leq N} \exp \left(c \left[i_1^{1-\gamma} - i_2^{1-\gamma} + i_3^{1-\gamma} - i_4^{1-\gamma} + \dots + i_{K-1}^{1-\gamma} - i_K^{1-\gamma} \right] \right),$$

for large N and fixed n_*, γ, c , with $K = 2m, m \geq 1$. As needed for equation (14), we will show that it asymptotically equals $\frac{N^{K(1+\gamma)/2}}{c^m(1-\gamma^2)^m m!}$ as $N \rightarrow \infty$.

The result will immediately follow from the next statement, as $Q(n_*, N)$ (asymptotically) does not depend on n_* as $N \rightarrow \infty$.

Lemma 1 *Define*

$$Z_l := \rho_l \cdot \frac{i_{2l+1}^{(1+\gamma)l}}{l! c^l (1-\gamma^2)^l} + o\left(N^{(1+\gamma)l}\right),$$

where $\rho_l \rightarrow 1$ as $N \rightarrow \infty$. Then, for $l = 0, 1, \dots, m - 1$, we have

$$Q(n_*, N) = \sum_{n_*+2l \leq i_{2l+1} < i_{2l+2} < \dots < i_{2m-1} < i_{2m} \leq N} \exp \left\{ c \left[i_{2l+1}^{1-\gamma} - i_{2l+2}^{1-\gamma} + \dots + i_{2m-1}^{1-\gamma} - i_{2m}^{1-\gamma} \right] \right\} \times Z_l \tag{22}$$

and also,

$$Q(n_*, N) = Z_m. \tag{23}$$

Proof We are going to prove the statement by induction on l .

For $l = 0$, it is true. Now assume that we have established (22) for some $l \geq 0$. Then

$$\begin{aligned}
 Q(n_*, N) = & \sum_{n_*+2l+2 \leq i_{2l+3} < i_{2l+4} < \dots < i_{2m-1} < i_{2m} \leq N} \exp \left\{ c \left[i_{2l+3}^{1-\gamma} - i_{2l+4}^{1-\gamma} \right. \right. \\
 & \left. \left. + \dots + i_{2m-1}^{1-\gamma} - i_{2m}^{1-\gamma} \right] \right\} \\
 & \times \sum_{n_*+2l \leq i_{2l+1} < i_{2l+2} < i_{2l+3}} \exp \left\{ c \left[i_{2l+1}^{1-\gamma} - i_{2l+2}^{1-\gamma} \right] \right\} \left[\frac{\rho_l i_{2l+1}^{(1+\gamma)l}}{l! c^l (1-\gamma^2)^l} \right. \\
 & \left. + o \left(N^{l(1+\gamma)} \right) \right],
 \end{aligned}$$

where the sum in the second line is taken over i_{2l+1} and i_{2l+2} only. We shall estimate the sum below.

First, note that

$$(*) = \sum_{n_* \leq i_{2l+1} < i_{2l+2} < N} \exp \left\{ c \left[i_{2l+1}^{1-\gamma} - i_{2l+2}^{1-\gamma} \right] \right\}$$

where each expression is between 0 and 1, can be very well approximated by the corresponding integral, since, whenever $y \leq x$, $|\tilde{x} - x| \leq 1$ and $|\tilde{y} - y| \leq 1$, the ratio

$$\frac{e^{c[\tilde{y}^{1-\gamma} - \tilde{x}^{1-\gamma}]}}{e^{c[y^{1-\gamma} - x^{1-\gamma}]}}$$

is bounded above by $e^{c_1[y^{-\gamma} + x^{-\gamma}]}$ where $c_1 > 0$ is some constant. Hence, outside of the area where x and y are both smaller than \sqrt{N} , the above ratio is very close to 1, while the double sum over that area can be at most N . Therefore, as $N \rightarrow \infty$,

$$\begin{aligned}
 & \sum_{n_* \leq i_{2l+1} < i_{2l+2} < N} \exp \left\{ c \left[i_{2l+1}^{1-\gamma} - i_{2l+2}^{1-\gamma} \right] \right\} \\
 & = (1 + o(1)) \int_{n_*}^N \int_y^N e^{cy^{1-\gamma} - cx^{1-\gamma}} dx dy + \mathcal{O}(N).
 \end{aligned}$$

To calculate the inner integral, observe

$$\int e^{-cx^{1-\gamma}} \cdot \left[1 - \frac{\gamma x^{\gamma-1}}{c(1-\gamma)} \right] dx = -\frac{x^\gamma e^{-cx^{1-\gamma}}}{c(1-\gamma)} + \text{const},$$

implying

$$\begin{aligned}
 R(y, N) & \leq \int_y^N e^{-cx^{1-\gamma}} dx \leq (1 + \psi(y)) \times R(N) \\
 \text{where } R(y, N) & := \frac{y^\gamma e^{-cy^{1-\gamma}}}{c(1-\gamma)} - \frac{N^\gamma e^{-cN^{1-\gamma}}}{c(1-\gamma)} \text{ and } \psi(y) = \left[1 - \frac{\gamma y^{\gamma-1}}{c(1-\gamma)} \right]^{-1}. \tag{24}
 \end{aligned}$$

Note that $\psi(y) \downarrow 1$ as $y \rightarrow \infty$, hence, since $y \geq n_*$,

$$\int_{n_*}^N \left[\int_y^N e^{-cx^{1-\gamma}} dx \right] e^{cy^{1-\gamma}} dy \leq \int_{n_*}^N (1 + \psi(n_*)) R(y, N) e^{cy^{1-\gamma}} dy$$

$$\leq \frac{1 + \psi(n_*)}{c(1 - \gamma)} \int_0^N y^\gamma dy = \frac{N^{1+\gamma}}{c(1 - \gamma^2)}.$$

Consequently, $(*) = \mathcal{O}(N^{1+\gamma})$ and

$$\sum_{n_*+2l \leq i_{2l+1} < i_{2l+2} < i_{2l+3}} \exp \left\{ c \left[i_{2l+1}^{1-\gamma} - i_{2l+2}^{1-\gamma} \right] \right\} \cdot o \left(N^{l(1+\gamma)} \right) = o \left(N^{(l+1)(1+\gamma)} \right).$$

The next step is to compute

$$\sum_{n_*+2l \leq i_{2l+1} < i_{2l+2} < i_{2l+3}} i_{2l+1}^{(1+\gamma)l} \times \exp \left\{ c \left[i_{2l+1}^{1-\gamma} - i_{2l+2}^{1-\gamma} \right] \right\}. \tag{25}$$

W.l.o.g. assume that $N^{\gamma/2} > n_* + 2l$, then (25) can be split into two terms, depending whether i_{2l+1} is smaller or larger than $N^{\gamma/2}$.

In the case $i_{2l+1} \leq N^{\gamma/2}$, we have

$$\sum_{n_*+2l \leq i_{2l+1} < i_{2l+2} < i_{2l+3}, i_{2l+1} \leq N^{\gamma/2}} i_{2l+1}^{(1+\gamma)l} \times \exp \left\{ c \left[i_{2l+1}^{1-\gamma} - i_{2l+2}^{1-\gamma} \right] \right\}$$

$$\leq N^{\gamma/2} \sum_{1 < i_{2l+2} < N} i_{2l+1}^{(1+\gamma)l} \leq N^{\gamma/2} \cdot N \cdot N^{(1+\gamma)l} = o \left(N^{(l+1)(1+\gamma)} \right)$$

For $i_{2l+1} > N^{\gamma/2}$, the sum in (25) is well approximated by the integral

$$\int_{N^{\gamma/2}}^b \int_y^b y^q e^{c[y^{1-\gamma} - x^{1-\gamma}]} dx dy = \int_{N^{\gamma/2}}^b y^q e^{cy^{1-\gamma}} \left[\int_y^b e^{-cx^{1-\gamma}} dx \right] dy$$

where $q := (1 + \gamma)l$ and $b := i_{2l+3}$. We are again allowed to do this as for $|\tilde{x} - x| \leq 1$ and $|\tilde{y} - x| \leq 1$ the ratio

$$\frac{\tilde{y}^q e^{c[\tilde{y}^{1-\gamma} - \tilde{x}^{1-\gamma}]}}{y^q e^{c[y^{1-\gamma} - x^{1-\gamma}]}}$$

is bounded above by $\left(1 + \frac{2l}{y}\right) e^{c_2(x^{-\gamma} + y^{-\gamma})}$ where $c_2 > 0$ is some constant. Observing that $x \geq y \geq N^{\gamma/2}$ yields that this ratio is very close to 1, and thus the double sum in (25) equals

$$(1 + o(1)) \int_{N^{\gamma/2}}^b y^q e^{cy^{1-\gamma}} \left[\int_y^b e^{-cx^{1-\gamma}} dx \right] dy + o \left(N^{(l+1)(1+\gamma)} \right). \tag{26}$$

From (24), since $y \geq N^{\gamma/2}$ and hence is large, we get that the inner integral equals $(1 + o(1)) R(y, b)$. Therefore, the main expression in (26), up to a factor $1 + o(1)$ and the remainder term, equals

$$\begin{aligned} & \int_{N^{\gamma/2}}^b \frac{y^{q+\gamma}}{c(1-\gamma)} - \frac{y^q b^\gamma e^{cy^{1-\gamma} - cb^{1-\gamma}}}{c(1-\gamma)} dy \\ &= \frac{b^{q+1+\gamma} - N^{(q+1+\gamma)\gamma/2}}{(q+1+\gamma)c(1-\gamma)} - \int_{N^{\gamma/2}}^b \frac{y^q b^\gamma e^{cy^{1-\gamma} - cb^{1-\gamma}}}{c(1-\gamma)} dy \\ &=: \frac{i_{2l+3}^{(l+1)(1+\gamma)} - o(N^{(l+1)(1+\gamma)})}{(l+1) \cdot c(1-\gamma)^2} - (**). \end{aligned}$$

Now the only remaining step is to show that the integral (**) is of smaller order; then the induction step is finished. To this end, fix some $\gamma < \theta < 1$, and do the following estimation.

$$\begin{aligned} (**) &\propto \int_{N^{\gamma/2}}^b y^q b^\gamma e^{cy^{1-\gamma} - cb^{1-\gamma}} dy \leq \int_0^{b-b^\theta} y^q b^\gamma e^{cy^{1-\gamma} - cb^{1-\gamma}} dy \\ &\quad + \int_{b-b^\theta}^b y^q b^\gamma e^{cy^{1-\gamma} - cb^{1-\gamma}} dy \\ &\leq \int_0^{b-b^\theta} y^q b^\gamma e^{c[(b-b^\theta)^{1-\gamma} - b^{1-\gamma}]} dy + b^\theta \cdot b^{q+\gamma} \\ &\leq b^{q+\gamma} \left[b \cdot e^{-c(1-\gamma+o(1))b^{\theta-\gamma}} + b^\theta \right] \\ &= o\left(b^{q+1+\gamma}\right) \\ &= o\left(N^{(l+1)(1+\gamma)}\right), \end{aligned}$$

since $b \leq N$.

Finally, (23) follows from repeating the above steps verbatim with the sum

$$\sum_{n_*+2m < i_{2m-1} < i_{2m} < i_{2m+1}} \exp \left\{ c \left[i_{2m-1}^{1-\gamma} - i_{2m}^{1-\gamma} \right] \right\} \times Z_{m-1},$$

by replacing i_{2m+1} by $N + 1$. □

References

1. Abramowitz, M.; Stegun, I. A. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Reprint of the 1972 edition. Dover (1992)
2. Bowman, F.: Introduction to Bessel Functions. Dover Publications Inc., New York (1958)
3. Dietz, Z., Sethuraman, S.: Occupation laws for some time-nonhomogeneous Markov chains. Electron. J. Probab. **12**(23), 661–683 (2007)
4. Dobrushin, R.L.: Central limit theorems for non-stationary Markov chains. I, II. Theory Probab. Appl. **1**, 65–80 (1956)

5. Durrett, R.: Probability: Theory and Examples, 2nd edn. Duxbury Press, Belmont (1995)
6. Engländer, J., Volkov, S.: Turning a coin over instead of tossing it, preprint. [arXiv:1606.03281](https://arxiv.org/abs/1606.03281)
7. Gantert, N.: Laws of large numbers for the annealing algorithm. *Stoch. Process. Appl.* **35**, 309–313 (1990)
8. Jones, G.L.: On the Markov chain central limit theorem. *Probab. Surv.* **1**, 299–320 (2004)
9. Lyons, R.: Strong laws of large numbers for weakly correlated random variables. *Mich. Math. J.* **35**(3), 353–359 (1988)
10. Peligrad, M.: Central limit theorem for triangular arrays of non-homogeneous Markov chains. *Probab. Theory Relat. Fields* **154**(3–4), 409–428 (2012)
11. Sethuraman, S., Varadhan, S.R.S.: A martingale proof of Dobrushin’s theorem for non-homogeneous Markov chains. *Electron. J. Probab.* **10**(36), 1221–1235 (2005)