

# ALICE AND BOB ON $\mathbb{X}$ : REVERSAL, COUPLING, RENEWAL

GEOFFREY R. GRIMMETT

ABSTRACT. A neat question involving coin flips surfaced on  $\mathbb{X}$ , and generated an intensive ‘storm’ of ‘social mathematics’. In a sequence of flips of a fair coin, Alice wins a point at each appearance of two consecutive heads, and Bob wins a point whenever a head is followed immediately by a tail. Who is more likely to win the game? The subsequent discussion illustrated conflicting intuitions, and concluded with the correct answer (it is a close thing). It is explained here why the context of the question is interesting and how it may be answered in a quantitative manner using the probabilistic techniques of reversal, coupling, and renewal.

## 1. THE PROBLEM

**Question 1.** *A fair coin is tossed  $n$  times. Alice scores one point at each appearance of two consecutive heads, and Bob scores one point each time a head is followed immediately by a tail. The winner is the player who accumulates more points. Who is the more likely to win?*

Here are some intuitive arguments in order of decreasing naiveté.

- (a) In any pair of consecutive coin tosses, Alice wins a point with probability  $\frac{1}{4}$ , and Bob wins a point with the same probability. Therefore, each has a mean total score of  $\frac{1}{4}(n - 1)$ . Since these means are equal, the game is fair and each player has the same probability of winning.
- (b) Alice’s points tend to arrive in clusters, whereas Bob’s are isolated. That favours Alice, so she is more likely to win.
- (c) When Alice wins a point, she has an increased chance of winning again. Therefore, when she wins, she tends to do so by a wider margin than does Bob. However, her *mean total* is the same as Bob’s. It follows that Bob has a greater *probability* of winning.

This question (with  $n = 100$ ) was posed by Daniel Litt on his  $\mathbb{X}$  feed [9] on 16 March 2024. At the current time of writing, his post has attracted 1.2M views. First responders tended to favour Alice above Bob on the grounds of argument (b), and a vote was reported as placing Alice (26.3%) over Bob (10.2%), with 42.8% of the 51,588 voters supporting equality. Later simulations appeared to show that Bob has a slight advantage.

ChatGPT has changed its position on the question over the intervening months. Initially it favoured Alice on the grounds given in (b) above. At the time of writing it has veered

---

*Date:* submitted August 30, 2024, revised September 1, 2025.

*2010 Mathematics Subject Classification.* 60C05, 91A15.

*Key words and phrases.* Coin tosses, random game, central limit theorem.

towards (a), with the conclusion that “The game is fair to both players in terms of expected outcomes.” Fair enough, but not very helpful. There remains hope for mathematicians.\*

More complete answers began to surface on Litt’s  $\mathbb{X}$  feed fairly soon after the original post, and suggestions were made for rigorous proof. It is not easy to convey the details and to check the correctness of a mathematical argument within the confined format of  $\mathbb{X}$ , and hence this note. We present three conclusions.

**Theorem 2.** *Consider  $n$  flips of a fair coin.*

- (a) *Bob is (strictly) more likely than Alice to win when  $n \geq 3$ .*
- (b)  $\mathbb{P}(\text{Bob wins}) - \mathbb{P}(\text{Alice wins}) \sim c/\sqrt{n}$  as  $n \rightarrow \infty$ , where

$$c = \frac{1}{2\sqrt{\pi}} \approx 0.282.$$

- (c) *The probability of a tie is asymptotically  $1/\sqrt{\pi n}$ .*

Explicit representations for the probabilities in (b) and (c) are given in equations (14)–(15). Parts (b) and (c) imply that

$$\frac{1}{2} - \mathbb{P}(\text{B wins}) \sim \frac{1}{4\sqrt{\pi n}}, \quad \frac{1}{2} - \mathbb{P}(\text{A wins}) \sim \frac{3}{4\sqrt{\pi n}}. \quad (1)$$

It is immediate that Alice and Bob are equally likely to win when  $n = 1, 2$ . By considering the eight possible outcomes when  $n = 3$ , we have  $\mathbb{P}(\text{Bob wins}) - \mathbb{P}(\text{Alice wins}) = \frac{1}{8}$  in this case.

The methods of proof may be summarized as reversal, coupling, and renewal, and the proof of Theorem 2 illustrates these standard techniques. Reversing a random sequence is a long-established activity which has contributed enormously to probability and especially to the theory of random walks (see, for example, [7, Sect. 3.10]). The coupling approach enables us to study the ‘pathwise’ relationship between winning sequences for Alice and for Bob, rather than by simply calculating probabilities (see, for example, [7, Sect. 4.12]). Renewal theory is the study of random processes that renew themselves at random times (see for example, [7, Chap. 10]).

We make two intuitive remarks about the problem (see also the next section). Let  $S_n$  denote Bob’s score minus Alice’s score after  $n$  flips.

- (a) Certainly,  $S_n$  has mean 0, but that does not imply that its distribution is symmetric, and it is precisely such asymmetry that skews the game towards Bob. One measure of asymmetry is termed ‘skewness’ (denoted ‘skw’), and is defined in terms of the so-called third central moment (see [7, p. 163]). The sign of the skewness (positive or negative) is an indicator of the more likely winner.
- (b) The term  $\sqrt{n}$  of Theorem 2 originates in the fact that  $S_n$  is the sum of independent summands. The number  $S_n$  is random but it has order  $n$ . Intuition based on the central limit theorem suggests that  $S_n$  has an approximately Gaussian distribution with mean 0 and variance of order  $n$ . In particular, one may guess that  $\mathbb{P}(S_n = 0)$  has order  $1/\sqrt{n}$ , in line with the so-called local central limit theorem (see [7, p.

---

\*Note on revision: This seems to be the view of ChatGPT also, which has now discovered the reference [13].

219]). As further evidence, the skewness of the sum of  $n$  independent copies of a random variable  $X$  is  $\text{skw}(X)/\sqrt{n}$ .

## 2. REMARKS ON RELATED PROBLEMS

One may phrase the problem addressed here as ‘HT vs HH’, and we shall write ‘HT > HH’ in paraphrase of Theorem 2(a). This implies that TH > HH for the following reason. Take a sequence  $\omega$  of heads/tails and reverse it in time to obtain  $\rho(\omega)$  — so that, for example, the sequence  $\omega = \text{HTHHHT}$  (sometimes abbreviated in the natural way to  $\text{HTH}^3\text{T}$ ) becomes the reversed sequence  $\rho(\omega) = \text{THHHTH}$  (abbreviated to  $\text{TH}^3\text{TH}$ ). Then  $\omega$  and  $\rho(\omega)$  have the same probability distribution, and the same count of consecutive head pairs HH. However, every HT in  $\omega$  becomes TH in  $\rho(\omega)$ . Such reversal is a technique used heavily in the current paper. Another simple technique is the interchange of heads and tails, whereby one deduces immediately that TH > TT.

Further questions spring easily to mind, involving sequences of length greater than 2, and indeed *sets* of sequences. There is a general method for determining the likely winner for sufficiently large  $n$ , and it can be applied to more general problems than HH vs HT. This is not the subject of the current article, but we outline it here for completeness. The aggregate score process  $S$  evolves as the number of flips increases, and it does so in the manner of a stochastic process with finite-range dependence (that is, there exists  $r < \infty$  such that the increments in  $S$  are independent when they are separated in time by  $r$  or more). Such processes have many of the properties of random walk, and in particular they satisfy a central limit theorem (CLT), in which the rate of convergence to the Gaussian distribution is determined (in the main) by the so-called third central moment. The last can be calculated explicitly in specific cases, and this yields the identity of the likely winner when  $n$  is large. Some technicalities must be overcome to complete this argument, and the nice arguments of Feller [6, Sect. XVI.4] may be useful in this regard. Mention is made of the preprint of Basdevant et al. [1]. A full account of an approach of this type is expected in [8].

This CLT argument yields a general solution to such problems, but only for sufficiently large values of  $n$ . Since it relies on asymptotics and error estimates, it cannot be expected to answer such questions for all values of  $n$ . In addition, it does not reveal detailed aspects of structure such as those described in the current work for the case HT vs HH. In a sense, Litt’s question was a lucky shot (though fortune favours the well prepared) — it was simple, captivating, and has a beautiful answer. There is a limited number of other cases that are at least partially susceptible to the methods of the current article.

As mathematicians we insist on unambiguous definitions of the objects of our study. One may capture pretty well all of probability theory by defining it as the theory of a countably infinite sequence of tosses of a fair coin. Perhaps not everything worth knowing is yet known about this primeval experiment.

The target of the current note is to obtain Theorem 2 using exact probabilistic methods to illuminate structure. The associated literature is slightly complex, owing to the nature and speed of communication of  $\mathbb{X}$ . The idea of reversal has arisen independently in certain contributions on  $\mathbb{X}$  to Litt’s post (see [12], for example, for some nice ideas). Mention is

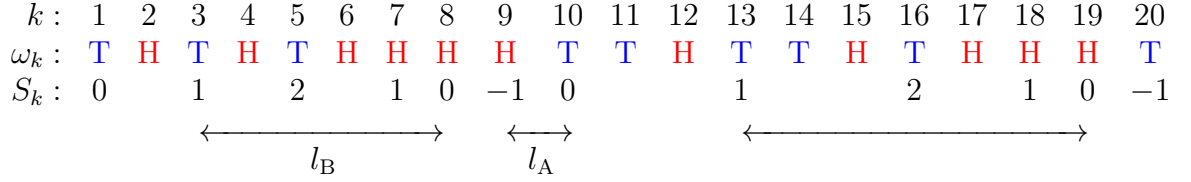


FIGURE 1. A sequence  $\omega = \text{THTHTH}^4\text{T}^2\text{HT}^2\text{HTH}^3\text{T}$  of 20 coin flips. The scores  $S_k$  are stated only when they change. Bob enters a winning period from epoch 3 to epoch 8, then Alice from 9 to 10, and then Bob from 13 to 19.

made of [5], which uses computational tools to explore the quantities in Theorem 2, and [13], which obtains the theorem by applying analytic tools to generating functions.

### 3. NOTATION AND BASIC OBSERVATIONS

Here is some notation. Abbreviate Alice to A, and Bob to B; write H for a head and T for a tail. Let  $\mathbb{N}$  denote the natural numbers, and let  $\Omega_\infty = \{\text{H}, \text{T}\}^\mathbb{N}$  be the set of all sequences  $\omega = (\omega_1, \omega_2, \dots)$  each element of which is either H or T. We shall normally express such sequences as  $\omega = \omega_1\omega_2\cdots$ . Similarly, let  $\Omega_n := \{\text{H}, \text{T}\}^n$ , the set of all sequences of  $n$  coin tosses.

Let  $\omega \in \Omega_\infty$ . Player A scores  $-1$  at each appearance of HH; player B scores 1 at each appearance of HT. The score  $S_k(\omega)$  is the aggregate score after  $k$  steps of  $\omega$ , that is,  $S_k(\omega)$  is the number of appearances of HT minus the number of appearances of HH in the subsequence  $\omega_1\omega_2\cdots\omega_k$ .

There follows an outline of the method to be followed here, with an illustration in Figure 1. First, one defines epochs of renewal, at which the scoring process restarts. Neither player scores until the first appearance of H.

- (i) If the following flip is T then Bob enters a winning period (of some length  $l_B$ ), which lasts until the next time that the aggregate score is 0; this must happen at an appearance of HH.
- (ii) If the following flip is H then Alice enters a winning period (of some length  $l_A$ ), which lasts until the next time that the aggregate score is 0; this must happen at an appearance of HT.

The process then restarts (subject to certain details to be made specific). After  $n$  coin flips, Alice is winning if she is then in a winning period, and similarly for Bob. Ties occur between winning periods. We will see that the representative periods  $l_A, l_B$  are such that  $l_A$  is (stochastically) smaller than  $l_B$  (that is to say,  $\mathbb{P}(l_A > l) \leq \mathbb{P}(l_B > l)$  for all  $l$ ), and Theorem 2(a) will follow.

The stochastic domination is proved by displaying a concrete coupling of  $l_A$  and  $l_B$ . The above explanation is given in more detail in Section 5.

The asymptotic part (b) of Theorem 2 is of course connected to the local central limit theorem (see, for example, [7, p. 219]). It suffices to use the earliest such theorem ever proved, namely the 1733 theorem of de Moivre [3, 11] (see also [4, p. 243–254] and [14,

Thm 1.1]), although one may equally use the (de Moivre–)Stirling formula alone (see the historical note [10]).

We have assumed implicitly in the above that the lengths  $l_A, l_B$  are finite (almost surely), and this requires proof. It follows from the next lemma.

Let  $\mathbb{P}$  be the probability measure on  $\Omega_\infty$  under which the coin tosses are independent random variables, each being **H** (respectively, **T**) with probability  $\frac{1}{2}$ . An event  $E$  is said to occur *almost surely* if  $\mathbb{P}(E) = 1$ .

**Lemma 3.** *The number of times  $r$  at which the aggregate score satisfies  $S_r(\omega) = 0$  is infinite almost surely.*

**Remark 4.** *Many of the arguments of this note are valid in the more general setting where heads occurs with some probability  $p \in (0, 1)$ . However, the conclusion of Lemma 3 is false when  $p \neq \frac{1}{2}$ , and indeed  $\mathbb{P}(\text{A wins})$  approaches 1 if  $p > \frac{1}{2}$ , and approaches 0 if  $p < \frac{1}{2}$ , in the limit as the number of coin flips grows to  $\infty$ .*

*Proof of Lemma 3.* There are several ways to prove this, of which we feature the neat argument of [12]. Let  $Z_1, Z_2, \dots$  be independent random variables taking the values  $\pm 1$  each with probability  $\frac{1}{2}$ . The partial sums  $T_n := Z_1 + Z_2 + \dots + Z_n$  are said to form a *simple random walk* (SRW). It is a famous theorem that

$$\mathbb{P}(T_n = 0 \text{ for some } n \geq 1) = 1$$

(see [7, Cor. 5.3.4]), and this is usually expressed by saying that the value 0 is ‘recurrent’. Having returned once, the walk will (with probability 1) return again, and so on. Therefore,

$$\mathbb{P}(T_n = 0 \text{ for infinitely many } n) = 1. \tag{2}$$

The score process  $(S_r)$  is not a SRW, but instead a ‘time-changed’ SRW. We view it as follows. First wait until the first **H**. Exactly one of the two following events then occurs, each having probability  $\frac{1}{2}$ ,

- (tail): the next flip is **T** and  $S$  increases by 1,
- (head): the next flip is **H** and  $S$  decreases by 1.

In case (tail), we wait for the next **H**; in either case, the process is then iterated. Let  $Z_i$  be the  $i$ th change in value of the score process. By the above, the  $Z_i$  are the steps of a SRW, albeit delayed in time. Every return to 0 of this SRW marks a return to 0 of the score, and the claim follows by (2).  $\square$

## 4. PROOF OF THEOREM 2(a)

A finite sequence  $\omega = \omega_1 \cdots \omega_k$  is called

a B-excursion if it starts HT and ends HH, and satisfies:

$$S_l(\omega) > 0 \text{ for } 1 < l < k, \text{ and } S_k(\omega) = 0,$$

an A-excursion if it starts HH and ends HT, and satisfies:

$$S_l(\omega) < 0 \text{ for } 1 < l < k, \text{ and } S_k(\omega) = 0,$$

an  $\hat{A}$ -excursion if it comprises an A-excursion followed by a (possibly empty) run of tails and then a single head.

We write  $\mathcal{B}$  for the set of B-excursions, and  $\mathcal{B}_k$  for those of length  $k$ , with similar notation for A-excursions and  $\hat{A}$ -excursions (using the notation  $\mathcal{A}$  and  $\hat{\mathcal{A}}$ , respectively).

$\hat{A}$ -excursions are introduced for the following reason. Players A and B can score only at a coin flip that follows an appearance of H. It will be convenient that excursions finish with a head. Since an A-excursion ends with a tail, we simply extend it to include any following run of tails, followed by the subsequent head; this extension contains neither HH nor HT and hence the score of the excursion is unchanged.

**Remark 5.** A headrun (respectively, tailrun) is a maximal consecutive sequence of heads (respectively, tails). Any finite sequence  $\omega$  comprises alternating headruns and tailruns. Alice's score equals the number of heads minus the number of headruns. Bob's score equals the number of tailruns having a preceding head. Let  $h$  be the total number of heads, and  $r$  the total number of runs (each comprising either heads or tails). The aggregate score  $S(\omega)$  equals  $r - h$  if  $\omega_1 = H$  (respectively,  $r - 1 - h$  if  $\omega_1 = T$ ).

We define sequence-reversal next. Let  $\Phi = \bigcup_{k=1}^{\infty} \Omega_k$  be the set of all non-empty finite sequences, and let  $\Phi^H$  be the subset of  $\Phi$  containing all finite sequences that begin H. For  $\omega \in \Phi$ , we write  $S(\omega)$  for the aggregate score of  $\omega$ . The score function is additive in the sense that

$$S(\omega_1 \cdots \omega_{m+n}) = S(\omega_1 \cdots \omega_m) + S(\omega_{m+1} \cdots \omega_{m+n}), \quad m, n \geq 1. \quad (3)$$

For a finite sequence  $\omega = \omega_1 \omega_2 \cdots \omega_k \in \Phi$ , define its reversal  $\rho(\omega)$  by  $\rho(\omega) = \omega_k \omega_{k-1} \cdots \omega_1$ .

**Lemma 6.**

- (a) If  $\omega = \omega_1 \cdots \omega_k \in \Phi^H$  ends in H (so that  $\omega_1 = \omega_k = H$ ), then  $S(\omega) = S(\rho(\omega))$ .
- (b) Let  $k \geq 2$ . The mapping  $\rho$  is a bijection between the sets  $\mathcal{B}_k$  and  $\hat{\mathcal{A}}_k$ .

*Proof.* (a) Let  $\omega \in \Phi^H$  end in H, so that  $\rho(\omega) \in \Phi^H$ . By Remark 5,  $S(\omega)$  equals the number of runs minus the number of heads. These counts are invariant under  $\rho$ .

(b) A B-excursion  $b = b_1 b_2 \cdots b_k \in \mathcal{B}_k$  has the form  $b = HT^r H \cdots HH$  for some  $r \geq 1$ , whence  $\rho(b) = b_k b_{k-1} \cdots b_1$  has the form  $HH \cdots HT^r H$ . By part (a),  $S(\rho(b)) = S(b) = 0$ , which implies, after removal of the final  $T^{r-1}H$ , that

$$S(b_k b_{k-1} \cdots b_{r+1}) = 0. \quad (4)$$

We claim that

$$S(b_k b_{k-1} \cdots b_j) < 0 \quad \text{for } j \in [r+2, k-1]. \quad (5)$$

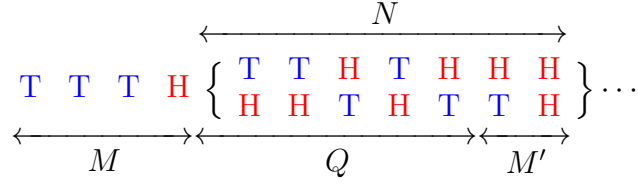


FIGURE 2. An illustration of the notation  $M$ ,  $N$ ,  $Q$ . Within the braces, the upper sequence is a B-excursion, and the lower is an  $\hat{A}$ -excursion.

Suppose, on the contrary, that  $j \in [r+2, k-1]$  is such that

$$S(b_k b_{k-1} \cdots b_j) = 0, \quad (6)$$

and pick  $j$  largest with this property. Since  $b_k b_{k-1} \cdots b_j$  starts with  $\text{HH}$ , it must end with  $\text{HT}$ . We now extend it to the next appearance in  $\rho(b)$  of  $\text{H}$ , thus obtaining  $b_k b_{k-1} \cdots b_j \text{T}^a \text{H}$  ( $= b_k b_{k-1} \cdots b_{j-a-1}$ ) for some  $a \geq 0$ . By (6),

$$S(b_k b_{k-1} \cdots b_{j-a-1}) = S(b_k b_{k-1} \cdots b_j \text{T}^a \text{H}) = 0. \quad (7)$$

By part (a), we have

$$S(b_{j-a-1} \cdots b_{k-1} b_k) = S(b_k b_{k-1} \cdots b_{j-a-1}) = 0.$$

By the additivity of  $S$ , (3),

$$S(b_1 b_2 \cdots b_{j-a-1}) = S(b_1 b_2 \cdots b_k) - S(b_{j-a-1} \cdots b_{k-1} b_k) = 0,$$

in contradiction of the assumption that  $b$  is a B-excursion, and (5) follows.

We illustrate the above argument with an example. Suppose  $k = 8$  and  $b = \text{HT}^2 \text{H} \text{TH}^3$ , so that  $\rho(b) = \text{H}^3 \text{THT}^2 \text{H}$ . Note that  $b_k b_{k-1} \cdots b_{r+1} = \text{H}^3 \text{THT}$  in agreement with (4). Equation (5) amounts to the assertion that  $S(\text{H}^3 \text{TH}), S(\text{H}^3 \text{T}), S(\text{H}^3), S(\text{H}^2) < 0$ .

We return to the proof of the lemma. By (4)–(5),  $b_k b_{k-1} \cdots b_{r+1}$  is an A-excursion, and hence  $\rho(b)$  is an  $\hat{A}$ -excursion. Therefore,  $\rho$  is an injection. By the same argument applied to  $\rho^{-1}$ , we have that  $\rho$  is a surjection.  $\square$

We make three observations, in amplification of the remarks of the last section. Let  $\omega \in \Omega_\infty$ , and observe the initial evolution of scores, as illustrated in Figure 2.

- (a) There is no score until the first  $\text{H}$  appears. Let  $M$  be the position of the first  $\text{H}$ , and note that

$$\mathbb{P}(M = m) = \left(\frac{1}{2}\right)^m, \quad m \geq 1. \quad (8)$$

That is, there is an initial sequence of tails of some random length  $M - 1$  ( $\geq 0$ ), followed by a head (so that  $\omega = \text{T}^{M-1} \text{H} \cdots$ ). Following this head, there is equal probability of  $\text{H}$  or  $\text{T}$ .

- (b) If the next toss is  $\text{T}$  (which is to say that  $\omega_{M+1} = \text{T}$ ), then Bob scores one point. He remains in the lead until the next epoch,  $M + N$  say, at which the aggregate score equals 0. It must be the case that  $\omega_{M+N-1} \omega_{M+N} = \text{HH}$ , and thus the sequence  $\omega_M \cdots \omega_{M+N}$  is a B-excursion. The process restarts from  $\text{H}$  at time  $M + N$ .
- (c) Suppose the next toss is  $\text{H}$  (that is,  $\omega_{M+1} = \text{H}$ ). The situation is slightly more complicated in this case. At time  $M + 1$ , Alice leads by 1, and she continues in the lead until the next epoch,  $M + Q$  say, when the aggregate score equals 0, and



moreover  $\omega_{M+Q-1}\omega_{M+Q} = \text{HT}$ . Thus the sequence  $\omega_1 \cdots \omega_{M+Q}$  is an A-excursion. Whereas in (b) the process restarts from the state  $\omega_{M+N} = \text{H}$ , this time we have  $\omega_{M+Q} = \text{T}$ , and we wait a random period of time for the next H. As in (a), there is a run of tails before the next head, which is to say that the process restarts from the head at epoch  $M + Q + M'$  where  $M'$  is an independent copy of  $M$ . The sequence  $\omega_M \cdots \omega_{M+N+M'}$  is an  $\hat{\text{A}}$ -excursion.

We shall adjoin certain sequences by placing them in tandem, and we write either  $\psi_1\psi_2$  or  $\psi_1 \cdot \psi_2$  for  $\psi_1$  followed by  $\psi_2$ . When scoring  $\psi_2$ , viewed as a subsequence of  $\psi_1\psi_2$ , one takes account of the final character of  $\psi_1$ , which will typically be H in the cases studied here. In order to do the necessary book-keeping we introduce the following notation: for a sequence  $\psi$  beginning with H, we write  $\text{H}^{-1}\psi$  for the sequence obtained from  $\psi$  by removing its initial H.

We construct next a random sequence of excursions drawn from the set  $\mathcal{A}$  of A-excursions. Consider a random sequence of flips beginning HH, and let  $\tau$  be the initial A-excursion of this sequence; we term  $\tau$  a *random A-excursion*; a similar definition holds for a random B-excursion. Let  $(\tau_i : i \geq 1)$  be a sequence of independent copies of  $\tau$  (that is, they are independent and each has the same distribution as  $\tau$ ). Let  $Q_i + 1$  be the length of  $\tau_i$  (so that  $\text{H}^{-1}\tau_i$  has length  $Q_i$ ). The  $\tau_i$  are almost surely finite, by Lemma 3. Let  $(M_i : i \geq 0)$  be independent copies of  $M$  (see (8)), also independent of the  $Q_i$ . Let  $N_i = Q_i + M_i$ , and let  $L_i$  be the sequence  $\text{T}^{M_i-1}\text{H}$ .

We have that  $\tau_i L_i$  is a random element of  $\hat{\mathcal{A}}$  of length  $Q_i + M_i + 1$ . By Lemma 6,  $\rho(\tau_i L_i)$  is a random element of  $\mathcal{B}$  of length  $Q_i + M_i + 1$ . We define

$$\begin{aligned} \text{the } \hat{\text{A}}\text{-excursion } \alpha_i &:= \tau_i L_i, \\ \text{the B-excursion } \beta_i &:= \rho(\tau_i L_i). \end{aligned} \tag{9}$$

Thus  $\beta_i$  is simply a reversal of  $\alpha_i$ , and this fact will provide a coupling of  $\hat{\text{A}}$ -excursions and B-excursions which is length-conserving (in that  $\alpha_i$  and  $\beta_i$  have the same length).

Both  $\alpha_i$  and  $\beta_i$  start and end with H; when placing them in tandem we shall strip the initial H. Another way of expressing (a)–(c) is as follows.

- (i) A random sequence of coin tosses begins with  $L_0$ .
- (ii) This is followed by

$$\gamma_1 := \begin{cases} \text{H}^{-1}\alpha_1 & \text{with probability } \frac{1}{2}, \\ \text{H}^{-1}\beta_1 & \text{with probability } \frac{1}{2}. \end{cases}$$

- (iii) Let  $k \geq 2$ , and suppose  $\alpha_i, \beta_i$  have been constructed for  $i = 1, 2, \dots, k$ . We then let

$$\gamma_{k+1} := \begin{cases} \text{H}^{-1}\alpha_{k+1} & \text{with probability } \frac{1}{2}, \\ \text{H}^{-1}\beta_{k+1} & \text{with probability } \frac{1}{2}. \end{cases} \tag{10}$$

Note that the length of  $\gamma_i$  is  $Q_i + M_i$ .

**Lemma 7.** *The sequence  $X = L_0\gamma_1\gamma_2 \cdots$  is an independent sequence of fair coin tosses.*

This does not require proof beyond the above remarks. The lemma provides a representation of a random sequence in terms of an initial tailrun, followed by independent copies



of  $\gamma_1$  in tandem. Each  $\gamma_i$  occupies a time-slot, and within this slot there appears a sequence of coin-tosses; according to the flip of another fair coin, we either retain this sequence or we reverse it (see (10)). (Some minor details concerning initial and final heads are omitted from this overview).

The sequence  $X$  may be expressed as follows:

$$X = L_0 \cdot \left\{ \begin{array}{l} \mathbf{H}^{-1}\alpha_1 = \mathbf{H}^{-1}\tau_1 L_1 \\ \mathbf{H}^{-1}\beta_1 = \mathbf{H}^{-1}\rho(\tau_1 L_1) \end{array} \right\} \cdot \left\{ \begin{array}{l} \mathbf{H}^{-1}\alpha_2 = \mathbf{H}^{-1}\tau_2 L_2 \\ \mathbf{H}^{-1}\beta_2 = \mathbf{H}^{-1}\rho(\tau_2 L_2) \end{array} \right\} \cdots \quad (11)$$

There are two possibilities within each matched pair of braces, and one chooses between them at random. More precisely, let  $Z_1, Z_2, \dots$  be independent random variables taking values in  $\{a, b\}$ , with each value chosen with probability  $\frac{1}{2}$ . We pick the upper element of the  $i$ th matched pair if  $Z_i = a$  and the lower element if  $Z_i = b$ .

Let  $X = (X_k : k \geq 1)$  be as in Lemma 7, so that  $X$  may be expressed in the form of (11). We shall explore the differences between the upper and lower elements of any given pair of braces there. Consider the  $i$ th matched pair of braces in (11). Let  $f_i$  be the first index in this subsequence of  $X$ , and let  $s_i$  be the last. Finally, let  $m_i = f_i + |\mathbf{H}^{-1}\tau_i| - 1$  where  $|\omega|$  denotes the length of a sequence  $\omega$ ; that is,  $m_i$  is the index of the last element of  $\mathbf{H}^{-1}\tau_i$  when  $Z_i = a$ . In illustration of these three indices, we assume that  $Z_i = a$  and express the  $i$ th matched pair in the form:

$$\left\{ \mathbf{H}^{-1}\tau_i L_i \right\} = \underset{\downarrow f_i}{(t_2 t_3 \dots t_c)} \cdot \underset{\downarrow m_i}{(l_1 l_2 \dots l_d)} \underset{\downarrow s_i}{} \quad (12)$$

where  $\tau_i = \mathbf{H}t_2 \dots t_c$  and  $L_i = l_1 l_2 \dots l_d$ . Then  $f_i$  is the index of  $t_2$  in  $X$ ,  $m_i$  is the index of  $t_c$ , and  $s_i$  is the index of  $l_d$  (' $f$ ' for *first* point, ' $m$ ' for *middle*, and ' $s$ ' for *second*). Evidently,  $m_i < s_i$ .

Here is a concrete example. If  $\tau_i = \mathbf{H}^3 \mathbf{T} \mathbf{H}^2 \mathbf{T} \mathbf{H} \mathbf{T}$  and  $L_i = \mathbf{T}^3 \mathbf{H}$ , then the  $i$ th pair of braces contains  $\mathbf{H}^{-1}\tau_i L_i = \mathbf{H}^2 \mathbf{T} \mathbf{H}^2 \mathbf{T} \mathbf{H} \mathbf{T} \cdot \mathbf{T}^3 \mathbf{H}$ , so that  $m_i = f_i + 7$  and  $s_i = m_i + 4$ , where  $f_i$  is the index in  $X$  of the first element of this brace.

*Proof of Theorem 2(a).* Let  $n \geq 3$ , and find the unique  $I = I_n$  such that  $n$  lies in the  $I$ th brace pair of (11) (write  $I = 0$  if no such pair exists). Note that  $f_I \leq n \leq s_I$  for  $I \geq 1$ . If  $n$  lies inside a B-excursion (respectively, A-excursion), then Bob (respectively, Alice) is currently winning. The third possibility is that the  $I$ th brace contains an  $\hat{\mathbf{A}}$ -excursion and  $n$  falls at or after the end of the A-excursion therein. More precisely, the score at epoch  $n$  satisfies

$$\begin{aligned} S_n &= 0 & \text{if} & & I = 0, \\ S_n &> 0 & \text{if} & & I \geq 1, Z_I = b, \text{ and } n < s_I \\ S_n &< 0 & \text{if} & & I \geq 1, Z_I = a, \text{ and } n < m_I, \\ S_n &= 0 & \text{if} & & \text{none of the above conditions hold.} \end{aligned} \quad (13)$$

Recalling (11), it follows that

$$\begin{aligned} \mathbb{P}(S_n > 0) - \mathbb{P}(S_n < 0) &= \frac{1}{2} [\mathbb{P}(I \geq 1, n < s_I) - \mathbb{P}(I \geq 1, n < m_I)] \\ &= \frac{1}{2} \mathbb{P}(I \geq 1, m_I \leq n < s_I) \geq 0. \end{aligned} \quad (14)$$

The *strict* positivity of the last probability (with  $n \geq 3$ ) follows by consideration of sequences beginning  $\mathbf{H}^2 \mathbf{T}^m \mathbf{H} \cdots$  with  $m \geq n$ , for which  $M_0 = 1$ ,  $M_1 = m$ , and  $\gamma_1 = \mathbf{H} \mathbf{T}^m \mathbf{H}$ .

Thus,  $\tau_1 = \mathbf{H}^2\mathbf{T}$ , and the interval  $\{m_1, m_1 + 1, \dots, s_1 - 1\} = \{3, 4, \dots, m + 2\}$  contains the value  $n$ .  $\square$

The probability of a tie is derived similarly to (14) as

$$\mathbb{P}(S_n = 0) = \mathbb{P}(I = 0) + \mathbb{P}(I \geq 1, n = s_I) + \frac{1}{2}\mathbb{P}(I \geq 1, m_I \leq n < s_I). \quad (15)$$

From (14)–(15) (as in (1)) one may obtain representations for the probabilities  $\mathbb{P}(S_n > 0)$  and  $\mathbb{P}(S_n < 0)$ .

## 5. PROOF OF THEOREM 2(b, c)

Returning to the representation (11) for  $X$ , we take the lower element in each matched pair of braces, and reorganize the terms slightly to obtain the sequence

$$Y := \{L_0\mathbf{H}^{-1}\tau_1\} \cdot \{L_1\mathbf{H}^{-1}\tau_2\} \cdots. \quad (16)$$

We emphasize that  $Y$  and  $X$  are very different, but nevertheless the sequence  $Y$  enables us to locate a set of renewal points that will aid the analysis of  $X$ , namely the following. Let  $R = \{m_0, m_1, m_2, \dots\}$  where  $m_0 = 0$  and the  $m_i$  for  $i \geq 1$  are as marked in (12); in other words,  $R$  is the set of indices of the rightmost elements of the braces in (16). Since  $(m_{i+1} - m_i : i \geq 0)$  are independent and identically distributed,  $R$  is indeed a renewal process. Such a process has the *renewal property* that, conditional on  $\{m \in R\}$  the process subsequent to time  $m$  has the same distribution as the process starting at time 0. Let  $\pi_m := \mathbb{P}(m \in R)$ .

Let  $m \leq n$ , and recall the notation  $Z_i, I = I_n$  from the last section. By the renewal property, conditional on the event  $\{m \in R, Z_I = a\}$ , the sequence  $\psi_{m,n} := X_{m+1}X_{m+2} \cdots X_n$  is a sequence of independent coin flips. For fixed  $n$ , the events

$$E_{m,n} = \{m \in R, Z_I = a, \psi_{m,n} = \mathbf{T}^{n-m}\}, \quad m = 1, 2, \dots, n,$$

are disjoint with union

$$\bigcup_{m \leq n} E_{m,n} = \{I \geq 1, Z_I = a, m_I \leq n < s_I\}.$$

By (14) and the above,

$$\begin{aligned} \mathbb{P}(S_n > 0) - \mathbb{P}(S_n < 0) &= \mathbb{P}(I \geq 1, Z_I = a, m_I \leq n < s_I) \\ &= \sum_{m=1}^n \mathbb{P}(m \in R, \psi_{m,n} = \mathbf{T}^{n-m} \mid Z_I = a) \mathbb{P}(Z_I = a) \\ &= \frac{1}{2} \sum_{m=1}^n \mathbb{P}(m \in R, \psi_{m,n} = \mathbf{T}^{n-m}) = \sum_{k=0}^{n-1} \left(\frac{1}{2}\right)^{k+1} \pi_{n-k}. \end{aligned} \quad (17)$$

Theorem 2(b) follows once the following has been proved.

**Proposition 8.** *We have that  $\pi_m \sim c/\sqrt{m}$  as  $m \rightarrow \infty$ , where  $c = 1/(2\sqrt{\pi}) \approx 0.282$ .*

The proof uses the de Moivre local central limit theorem (as in [14]), though this amounts only to sustained use of the (de Moivre–)Stirling formula, [10].

An outline of the proof of Theorem 2(c) is included after that of the proposition.

*Proof.* This proof is based on a combinatorial calculation associated with a certain subset of the renewal process  $R = (m_i : i \geq 0)$ , namely the subset  $R_X = \{m_i : i \geq 1, Z_i = a\}$ . It is immediate that  $R_X \subseteq R$ . Moreover, by the independence of  $(m_i)$  and  $(Z_j)$ ,

$$\mathbb{P}(m \in R_X) = \frac{1}{2}\mathbb{P}(m \in R), \quad m \geq 1. \quad (18)$$

By Lemma 7, the event  $\{m \in R_X\}$  is the set of all  $\omega$  with length  $m$  that satisfy  $\omega_{m-1}\omega_m = \text{HT}$  and  $S(\omega) = 0$ . Let  $h$  be the number of heads in such  $\omega$ , and  $r$  the number of runs (of either heads or tails). If  $r$  is even, (respectively, odd), then such sequences begin **H** (respectively, **T**), and they invariably finish with a tailrun of exactly one tail. By Remark 5, since the aggregate score is 0, we have  $h = r$  if  $r$  is even, and  $h = r - 1$  if  $r$  is odd.

We use the fact that the number of  $p$ -partitions of the integer  $q$  is  $\binom{q-1}{p-1}$ . Let  $m \geq 5$ , say. By partitioning the heads and tails into  $r$  runs subject to the above, and interleaving headruns and tailruns, we obtain that

$$|\{m \in R_X\}| = \Sigma_1 + \Sigma_2$$

where

$$\begin{aligned} \Sigma_1 &= \sum_{r \text{ odd}} \binom{m-h-2}{\frac{1}{2}(r+1)-2} \binom{h-1}{\frac{1}{2}(r-1)-1} = \sum_{r \text{ odd}} \binom{m-r-1}{\frac{1}{2}(r-3)} \binom{r-2}{\frac{1}{2}(r-3)}, \\ \Sigma_2 &= \sum_{r \text{ even}} \binom{m-h-2}{\frac{1}{2}r-2} \binom{h-1}{\frac{1}{2}r-1} = \sum_{r \text{ even}} \binom{m-r-2}{\frac{1}{2}r-2} \binom{r-1}{\frac{1}{2}r-1}. \end{aligned}$$

Set  $r = 2s + 1$  (respectively,  $r = 2s$ ) in  $\Sigma_1$  (respectively,  $\Sigma_2$ ) and add to obtain

$$\begin{aligned} |\{m \in R_X\}| &= \sum_{s=1}^{\infty} \left[ \binom{m-2s-2}{s-1} + \binom{m-2s-2}{s-2} \right] \binom{2s-1}{s-1} \\ &= \sum_{s=1}^{\infty} \binom{m-2s-1}{s-1} \binom{2s-1}{s-1}. \end{aligned} \quad (19)$$

We could now expand the binomial coefficients and use Stirling's formula directly. These calculations have been done already by de Moivre [3, 14], and we choose to refer to his results, as follows.

By (19),

$$\begin{aligned} \mathbb{P}(m \in R_X) &= \left(\frac{1}{2}\right)^m |\{m \in R_X\}| \\ &= \frac{1}{4} \sum_{s=1}^{\infty} \mathbb{P}(T_{m-2s-1} = s-1) \mathbb{P}(T_{2s-1} = s-1), \end{aligned} \quad (20)$$

where  $T_k$  has the binomial distribution with parameters  $k$  and  $\frac{1}{2}$ . By Stirling's formula (or [14, Thm 1.1]),

$$\mathbb{P}(T_{2s-1} = s-1) \sim \frac{1}{\sqrt{\pi s}} \quad \text{as } s \rightarrow \infty. \quad (21)$$

We may occasionally use real numbers in the following where integers are expected. The term  $\mathbb{P}(T_{m-2s-1} = s-1)$  in (20) is a maximum when  $m-2s-1 = 2(s-1)$ , which is to

say that  $s = \frac{1}{4}(m+1)$ . Let  $\gamma \in (\frac{1}{2}, \frac{2}{3})$ . We may restrict the summation in (20) to values of  $s$  satisfying  $|s - \frac{1}{4}m| < m^\gamma$ . To see this, note that

$$\sum_{s \geq \frac{1}{4}m + m^\gamma} \mathbb{P}(T_{m-2s-1} = s-1) \leq \sum_{s \geq \frac{1}{4}m + m^\gamma} \mathbb{P}(T_{m/2} \geq s-1),$$

which tends to 0 as  $m \rightarrow \infty$  by the moderate-deviation theorem of Cramér [2] (or, for a more modern treatment, see Feller [6, p. 549]<sup>†</sup>). We have used the fact that  $T_k$  is stochastically increasing in  $k$  in that, for  $t \geq 0$ ,  $\mathbb{P}(T_k > t)$  is non-decreasing in  $k$ . A similar argument applies for  $s \leq \frac{1}{4}m - m^\gamma$ .

Suppose  $|s - \frac{1}{4}m| < m^\gamma$ . By [14, Thm 1.2], there is an absolute constant  $C$  such that

$$\left| \mathbb{P}(T_{m-2s-1} = s-1) - \sqrt{\frac{2}{\pi(m-2s-1)}} \exp\left(-\frac{(4s-m-1)^2}{2(m-2s-1)}\right) \right| \leq \frac{C}{m^{2/3}}.$$

It follows that

$$\sum_{|s - \frac{1}{4}m| < m^\gamma} \mathbb{P}(T_{m-2s-1} = s-1)$$

deviates from

$$\Psi_m := \sum_{|s - \frac{1}{4}m| < m^\gamma} \sqrt{\frac{2}{\pi(m-2s-1)}} \exp\left(-\frac{(4s-m-1)^2}{2(m-2s-1)}\right)$$

by at most  $C(2m^\gamma + 1)/m^{2/3}$ , which tends to 0. Express  $\Psi_m$  as an integral, make the change of variable

$$\beta = \frac{4s-m-1}{\sqrt{m-2s-1}}$$

and let  $m \rightarrow \infty$  to obtain, by the dominated convergence theorem, that

$$\Psi_m \rightarrow \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\beta^2} d\beta = \frac{1}{2}.$$

Therefore, the terms  $\mathbb{P}(T_{m-2s-1} = s-1)$  in (20) are (asymptotically) concentrated in the interval  $\frac{1}{4}m \pm m^\gamma$ , with total weight  $\frac{1}{2}$ . Hence, by (18) and (20)–(21),

$$\pi_m \sim \frac{2}{8\sqrt{\pi m/4}} = \frac{1}{2\sqrt{\pi m}},$$

as claimed. □

*Outline proof of Theorem 2(c).* One may deduce part (c) from (15) by adapting the argument leading to (17). Alternatively, one may perform a direct calculation as above, and there follows a sketch of this. Let  $V_m$  be the set of vectors  $\omega \in \Omega_m$  such that  $S(\omega) = 0$ . Elements  $\omega \in V_m$  may be expressed as interleaved headruns and tailruns, with the counts of heads and tailruns being balanced by the condition  $S(\omega) = 0$  (see Remark 5). Such  $\omega$

---

<sup>†</sup>Feller's theorem may be stated as follows. Let  $(Z_i)$  be a sequence of independent identically distributed random variables with zero mean and unit variance, and with finite moment generating function in some neighborhood of 0, and write  $S_n = Z_1 + Z_2 + \dots + Z_n$ . The central limit theorem states that, for  $x \in \mathbb{R}$ ,  $\mathbb{P}(S_n > x\sqrt{n})/[1 - \Phi(x)] \rightarrow 1$  as  $n \rightarrow \infty$ , where  $\Phi$  denotes the standard Gaussian distribution function. The same limit is valid as  $x, n \rightarrow \infty$  in such a way that  $x = o(n^{1/6})$ .

may start with either **H** or **T**, and each case leads to two terms as in the first line of (19). Thus,  $|V_m|$  is the sum of four terms, each being the product of two binomial coefficients. The analysis continues as in the above proof.  $\square$

## ACKNOWLEDGMENTS

GRG thanks Daniel Litt for a sweet question, and Charles McGinnes for drawing his attention to it. He thanks Richard Weber, David Stirzaker, and Svante Janson for useful correspondence, and is grateful to the referees for their valuable reports which have led to corrections and improvements.

## REFERENCES

- [1] A.-L. Basdevant, O. Hénard, É. Maurel-Segala, and A. Singh, *On cases where Litt's game is fair*, (2024), <https://arxiv.org/abs/2406.20049>.
- [2] H. Cramér, *Sur un nouveau théorème-limite de la théorie des probabilités*, Actualités Scientifiques et Industrielles **736** (1938), 2–23, transl. H. Touchette, <https://arxiv.org/abs/1802.05988>.
- [3] A. de Moivre, *Approximatio ad Summam Terminorum Binomii  $\overline{a+b}^n$  in Seriem expansi*, (1733).
- [4] ———, *The Doctrine of Chances*, 3rd ed., 1756, <https://archive.org/details/doctrineofchance00moiv/page/n5/mode/2up>.
- [5] S. B. Ekhad and D. Zeilberger, *How to answer questions of the type: if you toss a coin  $n$  times, how likely is HH to show up more than HT?*, (2024), <https://arxiv.org/abs/2405.13561>.
- [6] W. Feller, *An Introduction to Probability Theory and its Applications. Vol. II*, 2nd ed., John Wiley & Sons., New York, 1971.
- [7] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 4th ed., Oxford University Press, 2020.
- [8] S. Janson, M. Nica, and S. Segert, *The generalized Alice HH vs Bob HT problem*, (2025), <https://arxiv.org/abs/2503.19035>.
- [9] D. Litt, (2024), <https://x.com/littmath/status/1769044719034647001>.
- [10] K. Pearson, *Historical note on the origin of the normal curve of errors*, Biometrika **16** (1924), 402–404.
- [11] K. Pearson, A. de Moivre, and R. C. Archibald, *A rare pamphlet of Moivre and some of his discoveries*, Isis **8** (1926), 671–683, <https://www.journals.uchicago.edu/doi/epdfplus/10.1086/358439>.
- [12] S. Ramesh, (2024), <https://x.com/RadishHarmers/status/1770217475960885661>.
- [13] S. Segert, *A proof that HT is more likely to outnumber HH than vice versa in a sequence of  $n$  coin flips*, (2024), <https://arxiv.org/abs/2405.16660>.
- [14] Z. Szewczak and M. Weber, *Classical and almost sure local limit theorems*, Dissertationes Math. **589** (2023), 97 pp.

CENTRE FOR MATHEMATICAL SCIENCES, CAMBRIDGE UNIVERSITY, WILBERFORCE ROAD, CAMBRIDGE CB3 0WB, UK

*Email address:* grg@statslab.cam.ac.uk