On cases where Litt's game is fair

A-L. Basdevant [*]   Olivier Hénard [†]   Édouard Maurel-Ségala [†]&   Arvind Singh [†]

December 27, 2024

Abstract

A fair coin is flipped n times, and two finite sequences of heads and tails with the same length are given, say A and B. Each time the word A appears in the sequence of fair coin flips, Alice gets a point, and each time the word B appears, Bob gets a point. Who is more likely to win? This puzzle is a slight extension of Litt's game [8] that recently set Twitter abuzz. We show that Litt's game is fair for any value of n and any two words that have the same auto-correlation structure by building up a bijection that exchanges Bob and Alice scores; the fact that the inter-correlations do not come into play in this case may come up as a surprise.

1   Introduction and main results

In [8], Litt came up with the following puzzle: a fair coin is tossed n times, Alice gets a point each time the sequence HH appears, while Bob scores a point each time the sequence HT appears (and these sequences may be overlapping). The game may result in a win for Alice, a win for Bob or a tie (the possibilities being exclusive). Who is more likely to win, Alice or Bob? The perhaps surprising answer (if one judges by the Twitter poll) is that:

1. for any size $n \geqslant 3$, Bob is more likely to win than Alice,

2. for n large, it holds:

$$2 \times \big(\mathbb{P}_n(\text{Bob wins}) - \mathbb{P}_n(\text{Alice wins}) \big) \sim \mathbb{P}_n(\text{Tie}) \sim \frac{1}{\sqrt{\pi n}}.$$

Several elements of proofs for the first fact quickly emerged on Twitter [13, 14, 17, 18], then more formal proofs of both facts appeared on the arXiv [3, 16]. Nica [10] recorded a YouTube video to introduce the problem to a wider audience. Zeilberger online journal [19] maintains a list of

---

[*]Sorbonne Université.

[†]Université Paris-Saclay.

anne.laure.basdevant@normalesup.org
olivier.henard@universite-paris-saclay.fr
edouard.maurel-segala@universite-paris-saclay.fr
arvind.singh@universite-paris-saclay.fr

contributions and questions on the problem; among the proofs given, we would like to advertise an early probabilistic proof : in [13], Ramesh builds a "fair majorant" for the score difference of Alice and Bob that consists in a delayed simple random walk; the score difference of Alice and Bob is a deterministic function of this walk, at distance at most 1 below it, see right after the bibliography for more details. A rigorous proof of both facts 1 and 2 is then at an easy reach. Another by-product of his approach is that giving Alice an initial advantage of only one point reverses the statement of fact 1 in a strong sense: for any size $n \geqslant 0$, Alice then has probability $>1/2$ of winning; in particular, Alice is more likely to win than Bob. This approach has been detailed in Grimmett [5], who recently rediscovered the arguments of Ramesh.

Considering a contest between HT and TH is also possible, but arguably less interesting: flipping the sequence of tosses, or reading the sequence in the reverse order exchanges Alice and Bob points, which results in a fair game. Our aim in this note is to uncover more hidden symmetries within the words ensuring the game is again tied up.

For $\ell \geqslant 0$, a *word* of length ℓ is a sequence $A = (a_1, \ldots, a_\ell)$ in $\{\mathrm{H}, \mathrm{T}\}^\ell$. For $X_n := (\varepsilon_k)_{1 \leqslant k \leqslant n}$ a finite sequence in $\{\mathrm{H}, \mathrm{T}\}^n$, we denote by

$$N_A(X_n) := |\{\ell \leqslant k \leqslant n, (\varepsilon_{k-\ell+1}, \ldots, \varepsilon_k) = (a_1, \ldots, a_\ell)\}|$$

the number of occurrences of the word A in the sequence X_n. Let A and B be two words of size ℓ. In the generalized Litt's game, Alice wins if $N_A(X_n) > N_B(X_n)$, Bob wins if $N_B(X_n) > N_A(X_n)$, and this is a tie otherwise. A key quantity encoding the intersections of A and B is the correlation of A and B, which may be represented as a subset of integers or simply as a number (the base 2 expansion of the subset):

$$[A|B] = \sum_{k \in \mathrm{Cor}(A,B)} 2^k \quad \text{where} \quad \mathrm{Cor}(A, B) = \{1 \leqslant k \leqslant \ell - 1, \ (a_{\ell-k+1}, \ldots, a_\ell) = (b_1, \ldots, b_k)\}.$$

To be more specific, we shall call *inter-correlation* the correlation of two distinct words (beware the order matters), and *auto-correlation* the correlation of a word with itself. Our main result in this note is that Litt's game, which is played with a uniform X_n, is fair for words A and B with the same auto-correlation, regardless of their inter-correlation.

Theorem 1. *Let A, B two words of length ℓ such that $[A|A] = [B|B]$. Assume that, under $\mathbb{P}_n$, the letters $(\varepsilon_i)_{1 \leqslant i \leqslant n}$ of X_n form an i.i.d. sequence with the uniform distribution on $\{H, T\}$. Then for each $n \geqslant 1$, $(N_A(X_n), N_B(X_n))$ and $(N_B(X_n), N_A(X_n))$ have the same distribution. In particular, for any $n \geqslant 1$,*

$$\mathbb{P}_n(\text{Bob wins}) = \mathbb{P}_n(\text{Alice wins}).$$

For instance, the words $A = \mathrm{HHTHTH}$ and $B = \mathrm{HTTTHH}$ have the same auto-correlation 2 hence Litt's game played with A and B is fair despite the apparent lack of symmetry between these words. It is a remarkable fact, specific to the case of words with the same autocorrelation, that the law of $(N_A(X_n), N_B(X_n))$ does not depend on the inter-correlations, and this feature is at the heart of our proof.

2

It is then possible to estimate the quantity in the last equality, or equivalently the probability of a tie $\mathbb{P}_n(\text{Tie})$, using a Local Central Limit theorem for sums of weakly dependent random variables. We do not venture into this for the following reason : In a recent breakthrough announced online [11], Nica (together with Janson) pointed to a generic method based on Edgeworth expansions to tackle the case of words A and B with possibly distinct auto-correlations; the sketch of proof, together with some basic moment computations, hints to the following asymptotic estimates: for $A \neq B$, as n gets large, the following asymptotics:

$$\mathbb{P}_n(\text{Bob wins}) - \mathbb{P}_n(\text{Alice wins}) = \frac{[A|A] - [B|B]}{\sqrt{2^\ell + [A|A] + [B|B] - [A|B] - [B|A]}} \cdot \frac{1}{2\sqrt{\pi n}} + o\left(\frac{1}{\sqrt{n}}\right),$$

$$\mathbb{P}_n(\text{Tie}) = \frac{2^\ell}{\sqrt{2^\ell + [A|A] + [B|B] - [A|B] - [B|A]}} \cdot \frac{1}{2\sqrt{\pi n}} + o\left(\frac{1}{\sqrt{n}}\right)$$

hold as soon as the denominators on both RHS are non null (the four[1] pair of words giving a null denominator are associated with degenerated cases). Shortly after posting a first version of this work, Segert communicated to us a proof [15], available online as an addendum to his arXiv paper, of a general statement that entails both estimates. In case $[A|A] = [B|B]$, Theorem 1 refines on the first formula by stating that the quantity is, in fact, identically null for all n.

A key feature of the first formula is that $\mathbb{P}_n(\text{Bob wins}) - \mathbb{P}_n(\text{Alice wins})$ has the sign of $[A|A] - [B|B]$. In light of this observation together with some extensive numerical computations[2], we are lead to state the following remarkable conjecture for a *fixed length* n of the underlying word X_n. We use Δ to denote the symmetric difference of two sets.

Conjecture 1. *Assume* $[A|A] > [B|B]$. *Then, for each* $n \geqslant n_0 := 2\ell - \max\{\text{Cor}(A)\Delta\text{Cor}(B)\}$,

$$\mathbb{P}_n(\text{Alice wins}) < \mathbb{P}_n(\text{Bob wins}) < \frac{1}{2} < \mathbb{P}_n(\text{Alice wins}) + \mathbb{P}_n(\text{Tie}).$$

Some comments are in order. First, the restriction to $n \geqslant n_0$ is necessary for the first inequality only (the most important one indeed), and we have equality for the $n < n_0$; also, the last two inequalities are equivalent (and valid for each $n \geqslant 0$), but we like to phrase the conjecture this way, because of the following vivid interpretation of this set of inequalities: even if Alice looses in the standard version of the game, if we were to give Alice *one additional point* [3], then her new winning probability would exceed $1/2$ for each n according to the last inequality; in particular, Alice would be advantaged over Bob at each subsequent time. Conjecture 1 also suggests that the advantage of Bob versus Alice may be put in bijection with a subset of the Tie event, a key feature of the proof by Ramesh indeed. Conjecture 1 is valid for the original Litt's game by Ramesh's proof; also, it is in line with the asymptotics above: to see this, notice the last two inequalities in the conjecture are equivalent to

$$|\mathbb{P}_n(\text{Alice wins}) - \mathbb{P}_n(\text{Bob wins})| < \mathbb{P}_n(\text{Tie}),$$

---

[1] $(\text{TH}^{\ell-1}, \text{H}^{\ell-1}\text{T})$ and its siblings, for which $N_A(X_n) - N_B(X_n)$ belongs to $\{-1, 0, 1\}$.

[2] the conjecture has been checked for all words A and B of size $\leqslant 10$ and for $n \leqslant 35$.

[3] beware this is distinct from starting the random word X_n with an A

which holds asymptotically because of the inequality $|[A|A] - [B|B]| \leqslant 2^\ell - 2$, that holds for every pair of words A,B.

We should also point that the topic of pattern matching and overlaps has been the subject of many investigations, starting with the non-transitive *Penney Ante game* named after Penney [12] (and famously solved by Conway) in which one is asked to compute the probability that A appears before B in a sequence of fair coin flips, see [6, 7] and the references therein.

Our fixed length conjecture is reminiscent of an important fixed length result in the area of pattern matching : the fact that the number of words X of length n avoiding A is larger than the number of words avoiding B if $[A, A] > [B, B]$ for each $n \geqslant n_0$ defined in term of A and B in Conjecture 1, and with equality before. For n large enough (larger than an indefinite constant), the statement has been stated and proved by Guibas and Odlyzko [6] and independently by Blom [1], while the result for $n \geqslant n_0$ has been established by Mannson [9] after a series of works on the topic [4, 2].

2   Proof of Theorem 1

The auto-correlation and inter-correlation of two words are quantities that appear naturally when we look at the probability of one word appearing before another in a sequence of coin flips. The formal definition, which we repeat below, is the following.

Definition 2. *Let A, B two words of length ℓ. We define the indices of inter-correlation of A and B by*

$$\mathrm{Cor}(A, B) = \{1 \le k \le \ell - 1, (a_{\ell-k+1}, \ldots, a_\ell) = (b_1, \ldots, b_k)\}.$$

We write $\mathrm{Cor}(A)$ to denote $\mathrm{Cor}(A, A)$. We define the inter-correlation $[A|B]$ of (A, B) by

$$[A|B] = \sum_{k \in \mathrm{Cor}(A,B)} 2^k$$

and the auto-correlation of A as $[A|A]$.

Let us make some remarks about this definition

- The number $[A|B]$ is not in general equal to $[B|A]$.

- For all A, B of length ℓ, $[A|B] \in [\![0; 2^\ell - 2]\!]$.

- For every words A, B, C, D of length ℓ, $[A|B] = [C|D]$ if and only if $\mathrm{Cor}(A, B) = \mathrm{Cor}(C, D)$.

Fix any two words A and B with length ℓ and the same auto-correlation. To prove that $(N_A(X_n), N_B(X_n))$ has the same law as $(N_B(X_n), N_A(X_n))$, we prove the existence of a bijection ϕ from $\{0,1\}^n$ to $\{0,1\}^n$ such that, for any sequence $X_n \in \{0,1\}^n$, $(N_A(X_n), N_B(X_n)) = (N_B(\phi(X_n)), N_A(\phi(X_n)))$.

We introduce some additional notation. If C, D are two words, CD will be the concatenation of C and D. Besides, for two words C, D of length ℓ and $m \in \mathrm{Cor}(C, D)$, we denote $C^m D$

the word of length $2\ell - m$ beginning by C and ending by D. For example, if $C = \underline{\text{HTTHTH}}$, $D = \overline{\text{THTHHH}}$, we have $C^2 D = \underline{\text{HTTHTHTHHH}}$. We extend this notation to k words $C_1, \ldots, C_k$ of length ℓ and $m_1, \ldots, m_k \geqslant 1$ in the obvious way: if $m_i \in \text{Cor}(C_i, C_{i+1})$ for each i, we define the word of length $k\ell - \sum m_i$

$$Y = C_1{}^{m_1} C_2{}^{m_2} C_3{}^{m_3} \ldots C_{k-1}{}^{m_{k-1}} C_k. \tag{1}$$

Definition 3. *Let A, B be two words of length ℓ. We call* overlap *of A and B a word Y of the form* (1) *where $C_1, \ldots, C_k \in \{A, B\}^k$. We denote by $\mathcal{E}(A, B)$ the set of all* overlaps *of A and B.*

Note that do not accept $m_i = 0$ in (1). In particular, the concatenation $Y = AB$ of A and B may not be in $\mathcal{E}(A, B)$. We notice also that, for $Y \in \mathcal{E}(A, B)$, the expression of Y in the form given by (1) may be not unique. We call the maximal decomposition of Y the one such that, if $Y = C_1{}^{m_1} C_2{}^{m_2} C_3{}^{m_3} \ldots C_{k-1}{}^{m_{k-1}} C_k$, we have

$$N_A(Y) = |\{1 \leq i \leq k, C_i = A\}| \qquad N_B(Y) = |\{1 \leq i \leq k, C_i = B\}|.$$

We have the following decomposition of a word.

Definition 4. *Let Y a word. There exists a unique way to write Y in the form*

$$Y = X_0 E_1 X_1 E_2 \ldots X_{k-1} E_k X_k \tag{2}$$

such that

- *the words $E_1, \ldots, E_k$ belong to $\mathcal{E}(A, B)$.*

- *For all $i \in [\![0, k]\!]$, neither A nor B appears in X_i (word X_i may be empty).*

- *$N_A(Y) = \sum_{i=1}^{k} N_A(E_i)$ and $N_B(Y) = \sum_{i=1}^{k} N_B(E_i)$.*

We call pattern *of the word Y with respect to A and B the words $(E_1, \ldots, E_k)$ which appears in* (2) *and write $\boldsymbol{Patt}_{A,B}(Y) = (E_1, E_2, \ldots, E_k)$.*

Proposition 1. *Let A, B be two words with the same auto-correlation, and let $\phi : \mathcal{E}(A, B) \to \mathcal{E}(A, B)$ be defined in the following way. If $Y := C_1{}^{m_1} C_2{}^{m_2} \ldots C_{k-1}{}^{m_{k-1}} C_k$ with $C_i \in \{A, B\}$, we set*

$$\phi(Y) = \bar{C}_k{}^{m_{k-1}} \bar{C}_{k-1}{}^{m_{k-2}} \ldots \bar{C}_2{}^{m_1} \bar{C}_1$$

where $\bar{C}_i = A$ if $C_i = B$ and $\bar{C}_i = B$ if $C_i = A$. Then ϕ is well-defined, it is independent of the decomposition chosen for Y, it is an involution and $\phi(Y)$ has the same length as Y. Moreover, we have $(N_A(Y), N_B(Y)) = (N_B(\phi(Y)), N_A(\phi(Y)))$.

Proof. We first prove that $\phi(Y)$ is well defined *i.e.* if $m_i \in \text{Cor}(C_i, C_{i+1})$, then $m_i \in \text{Cor}(\bar{C}_{i+1}, \bar{C}_i)$. Recall that we assume that A, B have the same auto-correlation, *i.e.* $\text{Cor}(A) = \text{Cor}(B)$. We have two cases:

- Either $C_i = C_{i+1}$, for example $C_i = A$. Then $\bar{C}_i = \bar{C}_{i+1} = B$. Hence, we have $\text{Cor}(C_i, C_{i+1}) = \text{Cor}(A) = \text{Cor}(B) = \text{Cor}(\bar{C}_{i+1}, \bar{C}_i)$.

- Or $C_i \neq C_{i+1}$, for example $(C_i, C_{i+1}) = (A, B)$. Then $(\bar{C}_{i+1}, \bar{C}_i)$ is also equal to (A, B) and so $\mathrm{Cor}(C_i, C_{i+1}) = \mathrm{Cor}(\bar{C}_{i+1}, \bar{C}_i)$.

Note that ϕ does not depend of the decomposition chosen for Y. To justify this claim, it is enough to consider the case of words Y with two decompositions $Y = C_1{}^{m_1} C_2{}^{m_2} C_3$ and $Y = C_1{}^{m} C_3$. Note that $|Y| = 2\ell - m = 3\ell - m_1 - m_2$. Applying ϕ with the first decomposition we get $\phi(Y) = \bar{C}_3{}^{m_2} \bar{C}_2{}^{m_1} \bar{C}_1$ and since $|\phi(Y)| = |Y| < 2\ell$, necessarily, $\bar{C}_3$ and $\bar{C}_1$ overlap in the writing of $\phi(Y)$ and thus we also have $\phi(Y) = \bar{C}_3{}^{m} \bar{C}_1$. The fact that ϕ is an involution is clear.

We write now Y with its maximal decomposition. By construction, we directly get that $N_A(\phi(Y)) \geqslant N_B(Y)$ and $N_B(\phi(Y)) \geqslant N_A(Y)$. We claim that, by maximality of the decomposition, there is equality in these two inequalities. Indeed, consider the case of words Y whose maximal decomposition consist in two words: $C_1{}^{m} C_2$. Among those words Y, only the words where $C_1 = C_2$ have to be considered. If $Y_0 = C_1{}^{m} C_1$, it holds $\phi(Y_0) = \bar{C}_1{}^{m} \bar{C}_1$. Now, assume that $\phi(Y_0) = \bar{C}_1{}^{m_1} C_3{}^{m_2} \bar{C}_1$ for some $C_3 \in \{A, B\}$. The map ϕ being independent of the decomposition, we get $Y_0 = \phi(\phi(Y_0)) = C_1{}^{m_1} \bar{C}_3{}^{m_2} C_1$, which has a distinct count of the $\bar{C}_3$-word; this negates the fact that $Y_0 = C_1{}^{m} C_1$ is the maximal decomposition in the first place. $\qquad \square$

The function ϕ defined in the previous proposition can be extended to a function on patterns. For $(E_1, \ldots, E_k) \in \mathcal{E}(A, B)^k$, we set

$$\phi(E_1, \ldots, E_k) = (\phi(E_k), \ldots, \phi(E_1))$$

which defines an involution on $\mathcal{E}(A, B)^k$. Given a pattern $M = (E_1, \ldots, E_k) \in \mathcal{E}(A, B)^k$ and $n \geqslant 1$, we define $L_M(n)$ as the number of words of length n with pattern M:

$$L_M(n) = |\{\text{word } Y, \ |Y| = n \text{ and } \mathbf{Patt}_{A,B}(Y) = M\}|.$$

To prove Theorem 1, it is sufficient to show that, for any pattern M and any $n \geqslant 1$, we have

$$L_M(n) = L_{\phi(M)}(n). \tag{3}$$

Indeed, if Y is a word such that $\mathbf{Patt}_{A,B}(Y) = M$ and Z is a word such that $\mathbf{Patt}_{A,B}(Z) = \phi(M)$, we have

$$(N_A(Y), N_B(Y)) = \left(\sum_{i=1}^{k} N_A(E_i), \sum_{i=1}^{k} N_B(E_i)\right) = \left(\sum_{i=1}^{k} N_B(\phi(E_i)), \sum_{i=1}^{k} N_A(\phi(E_i))\right) = (N_B(Z), N_A(Z)).$$

In view of Equality (3), we conclude that

$$P\big((N_A(X_n), N_B(X_n)) = (a, b)\big) = \frac{1}{2^n} \sum_{\substack{\text{pattern } M \text{ s.t.} \\ (N_A(M), N_B(M)) = (a, b)}} L_M(n) = \frac{1}{2^n} \sum_{\substack{\text{pattern } M \text{ s.t.} \\ (N_A(M), N_B(M)) = (a, b)}} L_{\phi(M)}(n)$$

$$= P\big((N_B(X_n), N_A(X_n)) = (a, b)\big).$$

In order to establish (3), we prove the more slightly more precise result:

Lemma 1. *For any pattern $M = (E_1, \ldots, E_k)$, for any $I = (i_0, \ldots, i_k)$, set*

$$L^M(I) = |\{Y = X_0 E_1 X_1 \ldots E_k X_k : \boldsymbol{Patt}_{A,B}(Y) = M; \forall j, |X_j| = i_j\}|.$$

Then

$$|L^M(I)| = |L^{\phi(M)}(I')|$$

where $I' = (i_k, \ldots, i_0)$.

Proof. Let us remark that we have

$$L^M(I) = L^{E_1}(i_0, 0) \left(\prod_{j=1}^{k-1} L^{(E_j, E_{j+1})}(0, i_j, 0) \right) L^{E_k}(0, i_k)$$

and

$$L^{\phi(M)}(I') = L^{\phi(E_k)}(i_k, 0) \left(\prod_{j=1}^{k-1} L^{(\phi(E_{j+1}), \phi(E_j))}(0, i_j, 0) \right) L^{\phi(E_1)}(0, i_0).$$

Moreover, the value of $L^{(E_j, E_{j+1})}(0, i_j, 0)$ only depends on i_j, on the word ending E_j and on the word beginning E_{j+1}. For example, if E_j ends with an A and E_{j+1} starts with a B, we have

$$L^{(E_j, E_{j+1})}(0, i_j, 0) = L^{(A,B)}(0, i_j, 0).$$

Now, if E_j ends with an A and E_{j+1} starts with a B, then $\phi(E_{j+1})$ ends with an A and $\phi(E_j)$ starts with an B. Thus, in this case, we directly get

$$L^{(\phi(E_{j+1}), \phi(E_j))}(0, i_j, 0) = L^{(E_j, E_{j+1})}(0, i_j, 0) = L^{(A,B)}(0, i_j, 0).$$

The situation is more involved when E_j ends with the same word than E_{j+1} starts with, let say the word A. Then indeed

$$L^{(\phi(E_{j+1}), \phi(E_j))}(0, i_j, 0) = L^{(B,B)}(0, i_j, 0) \text{ while } L^{(E_j, E_{j+1})}(0, i_j, 0) = L^{(A,A)}(0, i_j, 0).$$

Besides, if E_1 starts with A, then $\phi(E_1)$ ends with a B, and we have

$$L^{\phi(E_1)}(0, i_0) = L^{(B)}(0, i_0) \text{ while } L^{E_1}(i_0, 0) = L^{(A)}(i_0, 0),$$

and a similar assertion holds for $L^{E_k}(0, i_k)$ and $L^{\phi(E_k)}(i_k, 0)$. Combining all these remarks, we see that Lemma 1 will be proved as soon as we establish the following equality: for each $i \geqslant 0$,

$$L^{(A,A)}(0, i, 0) = L^{(B,B)}(0, i, 0) \qquad \text{and} \qquad L^{(B)}(0, i) = L^{(A)}(i, 0). \tag{4}$$

We prove (4) by induction on i. The proposition clearly holds for $i = 0$. Assume it holds for $k \leq i - 1$. Thus, for any pattern M, if $I = (i_0, \ldots, i_k)$ with $i_j < i$ for all j, we get

$$|L^M(I)| = |L^{\phi(M)}(I')|.$$

Let us now prove that $L^{(B)}(0,i) = L^{(A)}(i,0)$. Writing $|I| = \sum_{j=0}^{k} i_j$ if $I = (i_0, \ldots, i_k)$ and $|M| = \sum_{i=1}^{k} |E_i|$ for the length of the pattern $M = (E_1, \ldots, E_k)$, we find that

$$
\begin{aligned}
L^{(A)}(i,0) = &= |\{\text{words } XA: \ |X| = i \text{ and } \mathbf{Patt}_{A,B}(XA) = A\}| \\
&= 2^i - \sum_{k \geqslant 1} \sum_{\substack{M=(E_1,\ldots,E_k)\neq(A) \\ E_k \text{ ends with } A}} \sum_{\substack{I=(i_0,\ldots,i_{k-1},0) \\ |I|+|M|=i+|A|}} |L^M(I)| \\
&= 2^i - \sum_{k \geqslant 1} \sum_{\substack{M=(E_1,\ldots,E_k)\neq(A) \\ E_k \text{ ends with } A}} \sum_{\substack{I=(i_0,\ldots,i_{k-1},0) \\ |I|+|M|=i+|A|}} |L^{\phi(M)}(I')|.
\end{aligned}
$$

At this point, we use the recurrence assumption noticing that $|I| < i$ since $|M| > |A|$. Recalling that if M ends with A, then $\phi(M)$ starts with B, we see that the last line is also equal to

$$
2^i - \sum_{k \geqslant 1} \sum_{\substack{M=(E_1,\ldots,E_k)\neq(B) \\ E_1 \text{ starts with } B}} \sum_{\substack{I=(0,i_1,\ldots,i_k) \\ |I|+|M|=i+|B|}} |L^M(I)| \ = \ L^{(B)}(0,i).
$$

The equality $L^{(A,A)}(0,i,0) = L^{(B,B)}(0,i,0)$ is proved with a similar argument. $\qquad\square$

References

[1] Gunnar Blom. On the stochastic ordering of waiting times for patterns in sequences of random digits. *Journal of applied probability*, 21(4):730–737, 1984.

[2] Isa Cakir, Ourania Chryssaphinou, and Marianne Månsson. On a conjecture by eriksson concerning overlap in strings. *Combinatorics, Probability and Computing*, 8(5):429–440, 1999.

[3] Shalosh B. Ekhad and Doron 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? *arxiv:2405.13561*, 2024.

[4] Kimmo Eriksson. Autocorrelation and the enumeration of strings avoiding a fixed string. *Combinatorics, Probability and Computing*, 6(1):45–48, 1997.

[5] Geoffrey R. Grimmett. Alice and Bob on **X**: reversal, coupling, renewal, 2024.

[6] Leonidas J Guibas and Andrew M Odlyzko. String overlaps, pattern matching, and nontransitive games. *Journal of Combinatorial Theory, Series A*, 30(2):183–208, 1981.

[7] Shuo-Yen Robert Li. A martingale approach to the study of occurrence of sequence patterns in repeated experiments. *Annals of Probability*, 8(6):1171–1176, 1980.

[8] Daniel Litt. https://x.com/littmath/status/1769044719034647001, March 16 2024.

[9] Marianne Månsson. Pattern avoidance and overlap in strings. *Combinatorics, Probability and Computing*, 11(4):393–402, 2002.

[10] Mihai Nica. The Coin Flip Game that Stumped Twitter: Alice HH vs Bob HT. https://www.youtube.com/watch?v=BAiuFOwhAWw&t=13s.

[11] Mihai Nica. Alice HH vs Bob HT using an Edgeworth expansion. https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/MihaiNicaAliceBob.pdf, June 11 2024.

[12] Walter Penney. Penney-Ante, Problem 95. *Journal of Recreational Mathematics*, 2(4), October 1969.

[13] Sridhar Ramesh. https://x.com/RadishHarmers/status/1770217475960885661, March 19 2024.

[14] Sridhar Ramesh. https://x.com/RadishHarmers/status/1770530578179198981, March 20 2024.

[15] Simon Segert. https://github.com/SimonSegert/simonsegert.github.io/blob/main/misc/coin_tossing.pdf, 2024.

[16] Simon Segert. A proof that HT is more likely to outnumber HH than vice versa in a sequence of n coin flips. *arxiv:2405.16660*, 2024.

[17] Alex Selby. https://x.com/alexselby1770/status/1769795239667994765, March 20 2024.

[18] D.L. Yonge-Mallo. https://x.com/dlyongemallo/status/1772280170096775170, March 25 2024.

[19] Doron Zeilberger. Personal (online) journal. https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/litt.html.

For the ease of reference, we reproduce here[4] *verbatim* the tweets [13] with the permission of their author Sridhar Ramesh.

---

[4]"Consider a random walk in which one takes equally likely steps of one unit up or one unit down, but with different distributions of speeds. (E.g., maybe up steps take one hour, while down steps have probability 1/2 of taking 2 hours, 1/4 of taking 3 hours, 1/8 of 4 hours, etc). The time it takes to return to the origin is independent of whether the first step is up and last step is down or vice versa, as doing the same steps in reverse order has the same probability. Thus, for any fixed walk time, the last step away from the origin begun before the time limit is equally likely to be up or down. Thus, at the end when "the buzzer goes off", one is equally likely to be above or below the origin (possibly in the middle of an uncompleted step). Applied to our game, with HH as a step up in one unit of time and HT^nH as a step down over $n+1$ units of time, this says we are equally likely to end above the origin (Alice wins or we are in the middle of an HT^nH step which has tied the game) or below it (Bob wins). Since it is indeed possible to end in the middle of a game-tying HT^nH step (e.g., if the game consists of HHT followed by all T's), Alice is less likely to win than Bob. QED. The salient difference is that the "buzzer" can cut off HT^nH in the middle (after awarding Bob a game-tying point but before returning to H), which it cannot do for HH. The random walk framing perhaps allows some ready generalization to other interesting problems."