

GAUSSIAN BEHAVIOR IN ZECKENDORF DECOMPOSITIONS FROM LATTICES

ERIC CHEN, ROBIN CHEN, LUCY GUO, CINDY JIANG, STEVEN J. MILLER, JOSHUA M. SIKTAR,
AND PETER YU

ABSTRACT. Zeckendorf's Theorem states that any positive integer can be written uniquely as a sum of nonadjacent Fibonacci numbers. We consider higher-dimensional lattice analogues in a chosen dimension $d \geq 1$, where a legal decomposition of a number n is a collection of lattice points such that each point is included at most once. Once a point is chosen, all future points must have strictly smaller coordinates, and the pairwise sum of the values of the points chosen equals n . We prove that the distribution of the number of summands in these lattice decompositions converges to a Gaussian distribution in d dimensions. As an immediate corollary, we obtain a new proof for the asymptotic number of certain lattice paths.

1. INTRODUCTION

Zeckendorf's Theorem states that any positive integer can be uniquely written as the sum of nonconsecutive Fibonacci numbers $\{F_n\}$, defined by $F_1 = 1$, $F_2 = 2$, and $F_{n+1} = F_n + F_{n-1}$ for all $n \geq 2$ [30]. We call this sum a number's **Zeckendorf decomposition**, and this leads to an equivalent definition of the Fibonacci sequence: this is the only sequence such that every positive integer can be written uniquely as a sum of nonadjacent terms. This interplay between recurrence relations and notions of legal decompositions holds for other sequences and recurrence rules as well. We report on some of the previous work on properties of generalized Zeckendorf decompositions for certain sequences, and then discuss our new generalizations to two-dimensional sequences. There is now extensive literature on the subject (see for example [2, 3, 4, 5, 8, 9, 13, 14, 15, 16, 18, 17, 19, 26, 27, 28, 29] and the references therein).

Lekkerkerker [22] proved that the average number of summands in the Zeckendorf decompositions of $m \in [F_n, F_{n+1})$ is $\frac{n}{\varphi^2+1} + O(1) \approx .276n$ as $n \rightarrow \infty$. Later authors extended this to other sequences and higher moments (see the previous references, in particular [1, 9, 10, 11, 24, 21, 27]), proving that given any rules for decompositions, there is a unique sequence such that every number has a unique decomposition, and the average number of summands converges to a Gaussian.

To date, most of the sequences studied have been one-dimensional; many that appear to be higher dimensional (such as [6, 7]) can be converted to one-dimensional sequences. Our goal is to investigate decompositions that are truly higher dimensional. We do so by creating a sequence arising from two-dimensional lattice paths on ordered pairs of positive integers. A legal decomposition in d dimensions will be a finite collection of lattice points for which

- (1) each point is used at most once, and

This project was begun when Miller was visiting Carnegie Mellon in 2016 and completed when he visited again in 2019; he thanks the institution for its hospitality. We thank Dilhan Salgado, Evan Fang, Jonathan Jenkins, Joonha Lee, Daniel Li, Ray Li, and Ethan Lu for carefully reading the final manuscript and some discussions, and the referee for helpful comments. This work was supported by NSF grants DMS1265673, DMS1561945, and the Eureka Program.

- (2) if the point (i_1, i_2, \dots, i_d) is included, then all subsequent points $(i'_1, i'_2, \dots, i'_d)$ have $i'_j < i_j$ for all $j \in \{1, 2, \dots, d\}$ (i.e., *all* coordinates must decrease between any two points in the decomposition).

We call these sequences of points on the d -dimensional lattice **simple jump paths**. In Section 4, we discuss generalizations in which we allow only some of the coordinates to decrease between two consecutive points in the path; this adds combinatorial difficulties. Note that the number we assign to each lattice point depends on how we order the points (unless we are in one dimension). For example, if $d = 2$ we can order the points by going along diagonal lines, or L -shaped paths. Explicitly, the first approach gives the ordering

$$(1, 1), \quad (2, 1), (1, 2), \quad (3, 1), (2, 2), (1, 3), \quad \dots, \tag{1.1}$$

whereas the second yields

$$(1, 1), \quad (2, 1), (2, 2), (1, 2), \quad (3, 1), (3, 2), (3, 3), (2, 3), (1, 3), \quad \dots \tag{1.2}$$

For the purposes of this paper, however, it does not matter which convention we adopt as our results on the distribution in the number of summands of a legal decomposition depend only on the combinatorics of the problem, and not the values assigned to each tuple; the role of the values surfaces only in questions of what fraction of integers are represented, and how often. We call the labeling attached to any choice a **Simple Zeckendorf Sequence in d Dimensions**, and comment shortly on how this is done. If $d = 1$, then we denote the sequence as $\{y_a\}_{a=0}^\infty$ and construct it as follows.

- (1) Set $y_1 := 1$.
- (2) Iterate through the natural numbers. If we have constructed the first k terms of our sequence, the $(k + 1)$ th term is the smallest integer that cannot be written as a sum of terms in the sequence, with each term used at most once.

Note this sequence is just powers of 2,

$$1 \quad 2 \quad 4 \quad 8 \quad 16 \quad 32 \quad 64 \quad 128 \quad 256 \quad 512 \quad \dots, \tag{1.3}$$

and a legal decomposition of n is just its binary representation.

If $d = 2$, on the other hand, as remarked above we have choices. We describe the Simple Zeckendorf Diagonal Sequence $\{y_{a,b}\}_{a,b=0}^\infty$; its construction is analogous to the $d = 1$ case and uses the following procedure.

- (1) Set $y_{1,1} := 1$.
- (2) Iterate through the natural numbers. For each such number, check if any path of numbers in our sequence with a strict leftward and downward movement between each two points sums to the number. If no such path exists, add the number to the sequence so that it is added to the shortest unfilled diagonal moving from the bottom right to the top left.
- (3) If a new diagonal must begin to accommodate a new number, set the value $y_{k,1}$ to be that number, where k is minimized so that $y_{k,1}$ has not yet been assigned.

In (1.4), we illustrate several diagonals' worth of entries when $d = 2$, where the elements are always added in increasing order. Note that unlike the Fibonacci sequence, we immediately see that we have lost the uniqueness of decompositions (for example, 25 has two legal

decompositions: $20 + 5$ and $24 + 1$).

280
157	263
84	155	259
50	82	139	230
28	48	74	123	198
14	24	40	66	107	184
7	12	20	33	59	100	171
3	5	9	17	30	56	93	160
1	2	4	8	16	29	54	90	159	...

(1.4)

Analogous procedures to the one which creates (1.4) exist for higher dimensions, but the intended illustration is most intuitive in two dimensions. For the same reason as in the $d = 2$ case, there are clearly multiple procedures to generate the higher-dimensional sequences, even if one fixes restrictions on how to choose the summands in as many as $d - 2$ dimensions.

Numerical explorations (see Figure 1) suggest that, similar to sequences mentioned earlier, the distribution of the number of summands converges to a Gaussian.

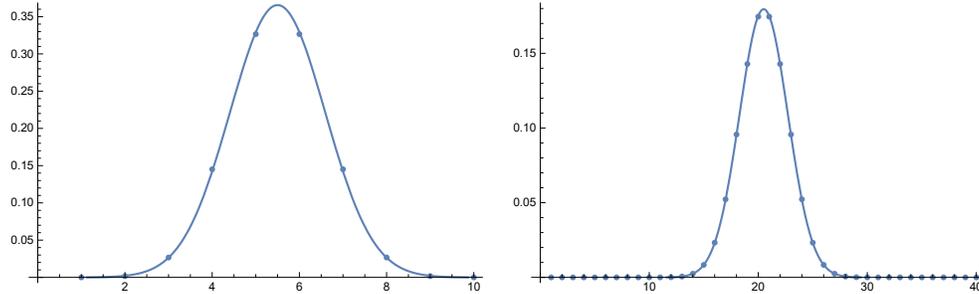


FIGURE 1. Gaussian distribution of the number of simple jump paths of varying lengths; in particular, if the paths have length (n, n) , then the mean is $\frac{1}{2}n + 1$ and the variance is $\frac{n^2}{4(2n-1)}$. Left: Starting at $(10, 10)$. Right: Starting at $(40, 40)$. In both cases, the horizontal axis is the number of summands and the vertical axis is the probability of obtaining a simple jump path with some number of summands when selecting one from all simple jump paths uniformly at random.

Our main result is that in any number of dimensions, we converge to a Gaussian as $n \rightarrow \infty$.

Theorem 1.1. (*d-dimensional Gaussianity*) *Let n be a positive integer, and consider the distribution of the number of summands among all simple jump paths of dimension d with starting point (i, i, \dots, i) where $1 \leq i \leq n$, and each distribution represents a (not necessarily unique) decomposition of some positive number. This distribution converges to a Gaussian as $n \rightarrow \infty$.*

In Section 2, we motivate our problem further, explore the notion of a simple jump path in more depth, and prove some needed lemmas. Then, we prove Theorem 1.1 in Section 3. The result is proven with the Central Limit Theorem for a binomial random variable if $d = 1$. If $d = 2$, it can be proved directly through combinatorial identities, but for larger d the combinatorial lemmas do not generalize, and we are forced to resort to analytic techniques. We show that the functional dependence is that of a Gaussian, and thus, as the probabilities add up to 1, the normalization constant, which depends on the number of paths, must have

a certain asymptotic formula. Thus, as an immediate consequence, we obtain new proofs for the asymptotic number of paths. We conclude the paper with a discussion of future work and generalizations of the simple jump paths.

2. PROPERTIES OF SIMPLE JUMP PATHS

We first set some notation for our simple jump paths. We have walks in d dimensions starting at some initial point (a_1, a_2, \dots, a_d) with each $a_j > 0$, and ending at the origin $(0, 0, \dots, 0)$. Note that our simple jump paths must always have movement in all dimensions at each step. We are just adding one extra point, at the origin, and saying every path must end there. Note that as we always change all of the indices during a step, we never include a point where only some of the coordinates are zero, and thus, there is no issue in adding one extra point and requiring all paths to end at the origin.

Our walks are sequences of points on the lattice grid with positive indices or the origin, and we refer to movements between two such consecutive points as **steps**. Thus, a simple jump path is a walk where each step has a strict movement in all d dimensions. More formally, a simple jump path of length k starting at (a_1, a_2, \dots, a_d) is a sequence of points $\{(x_{i,1}, \dots, x_{i,d})\}_{i=0}^k$ where the following hold:

- $(x_{0,1}, \dots, x_{0,d}) = (a_1, \dots, a_d)$,
- $(x_{k,1}, \dots, x_{k,d}) = (0, \dots, 0)$, and
- for each $i \in \{1, \dots, k - 1\}$ and $j \in \{1, \dots, d\}$, $x_{i,j} > x_{i+1,j}$.

For a fixed d and any choice of starting point $(n, n, \dots, n) \in \mathbb{R}^d$, we let $s_d(n)$ denote the number of simple jump paths from (n, n, \dots, n) to the origin, and $t_d(k, n)$ denote the subset of these paths with exactly k steps. As we must reach the origin, every path has at least one step, the maximum number of steps is n , and

$$s_d(n) = \sum_{k=1}^n t_d(k, n). \tag{2.1}$$

We now determine $t_d(k, n)$. In one dimension we have $t_d(k, n) = \binom{n-1}{k-1}$, because we must choose exactly $k - 1$ of the first $n - 1$ terms (we must choose the n th term as well as the origin, and thus, choosing $k - 1$ additional places ensures there are exactly k steps). The generalization to higher dimensions is immediate because we are looking at simple paths, and thus, there is movement in each dimension in each step; this is why we restrict ourselves to simple paths, because in the general case we do not have tractable formulas like the one below.

Lemma 2.1. *For a_1, \dots, a_d positive integers, let $t_d(k; a_1, \dots, a_d)$ denote the number of simple paths of length k starting at (a_1, \dots, a_d) and ending at $(0, \dots, 0)$. Then for $1 \leq k \leq \min(a_1, \dots, a_d)$,*

$$t_d(k; a_1, \dots, a_d) = \binom{a_1 - 1}{k - 1} \binom{a_2 - 1}{k - 1} \dots \binom{a_d - 1}{k - 1}. \tag{2.2}$$

If $a_1 = \dots = a_d = n$, we write $t_d(k, n)$ for $t_d(k; a_1, \dots, a_d)$. We have

$$s_d(n) = \sum_{k=1}^n t_d(k, n), \tag{2.3}$$

and $s_1(n) = 2^{n-1}$, $s_2(n) = \binom{2n-2}{n-1}$ (for higher d , we are unaware of the existence of any closed form expressions).

The proof is an immediate, repeated application of the one-dimensional result, with the two formulas (for $s_1(n)$ and $s_2(n)$) being well-known binomial identities (see for example [25]).

3. GAUSSIANTY IN d -DIMENSIONAL LATTICES

3.1. Mean and Variance. To prove Theorem 1.1, we start by determining the density, $p_d(k, n)$, for the number of simple jump paths of length k starting at (n, \dots, n) :

$$p_d(k, n) = \frac{t_d(k, n)}{s_d(n)}. \tag{3.1}$$

Much, though not all, of the proof when $d = 1$ carries over to general d . We therefore concentrate on $d = 1$ initially and then remark on what issues arise when we generalize and discuss the resolution of these problems.

We begin by determining the mean and standard deviation. The analysis for the mean holds for all d , but the combinatorial argument for the variance requires $d \leq 2$. Due to the presence of $n - 1$ in the formula for $t_d(k, n)$, we work with $n + 1$ below to simplify some of the algebra.

Lemma 3.1. *Consider all simple jump paths from $(n + 1, \dots, n + 1)$ to the origin in d dimensions. If K is the random variable denoting the number of steps in each path, then its mean $\mu_d(n + 1)$ and standard deviation $\sigma_d(n + 1)$ are*

$$\mu_d(n + 1) = \frac{1}{2}n + 1 \tag{3.2}$$

and

$$\sigma_1(n + 1) = \frac{\sqrt{n}}{2}, \quad \sigma_2(n + 1) = \frac{n}{2\sqrt{2n - 1}} \approx \frac{\sqrt{n}}{2\sqrt{2}}. \tag{3.3}$$

Further, we have

$$\sigma_d(n + 1) \leq \sigma_1(n + 1) \leq \sqrt{n}/2. \tag{3.4}$$

Proof. The results for $d = 1$ are well-known, because we have a binomial random variable. For $d = 2$, one can compute the mean and the variance by combinatorial arguments (see Appendix A); whereas these can be generalized to give the mean for any d , they do not generalize for the variance.

Because we must end at the origin, note that each path must have length at least 1. Thus, instead of studying the number of paths of length $k \in \{1, \dots, n + 1\}$, we instead study the number of paths of length $\kappa \in \{0, \dots, n\}$ and then add 1 to obtain the mean (there is no need to add 1 for the variance because the variance of K and $K - 1$ are the same).

Because

$$t_d(k; n + 1) = \frac{\binom{n}{k}^d}{s_d(n + 1)}, \tag{3.5}$$

the symmetry of the binomial coefficients about $n/2$ implies the mean of $K - 1$ is $n/2$. All that remains is to prove the variance bound for $d \geq 2$. Note that the variance of $K - 1$ is

$$\sigma_d(n + 1) = \sum_{\kappa=0}^n (\kappa - n/2)^2 \frac{\binom{n}{\kappa}^d}{s_d(n + 1)}. \tag{3.6}$$

By symmetry, it suffices to investigate $\kappa \geq n/2$. Since the binomial coefficients are strictly decreasing as we move further from the mean, for such κ we find that

$$\frac{p_d(\kappa)}{p_d(\kappa + 1)} = \frac{\binom{n}{\kappa}^d}{\binom{n}{\kappa+1}^d} \geq 1, \tag{3.7}$$

and thus for every $g > 0$, we see that the probability of $K - 1$ being within g of the mean increases as d increases. Thus, the variance is smallest at $d = 1$, completing the proof. \square

Next, we show that K is close to the mean with high probability.

Lemma 3.2. *Consider all simple jump paths from $(n + 1, \dots, n + 1)$ to the origin in d dimensions. If K is the random variable denoting the number of steps in each path, then the probability that K is at least $n^\epsilon n^{1/2}/2$ from the mean is at most $n^{-2\epsilon}$.*

Proof. By Chebyshev’s Inequality,

$$\text{Prob}(|K - (n/2 + 1)| \geq n^\epsilon \sigma_d(n + 1)) \leq \frac{1}{n^{2\epsilon}}. \tag{3.8}$$

Because $\sigma_d(n + 1) \leq n^{1/2}/2$ by Lemma 3.1, we only decrease the probability on the left if we replace $\sigma_d(n + 1)$ with $n^{1/2}/2$, and thus the claim follows. \square

One important consequence of the above lemma is that if we write k as $\mu_d(n + 1) + \ell n^{1/2}/2$, then with probability tending to 1, we may assume $|\ell| \leq n^\epsilon$.

3.2. Gaussianity. The proof of Theorem 1.1 in general proceeds similar to the $d = 1$ case. For $d \leq 2$, we have closed formulas for the variance and $s_d(n + 1)$, which simplify the proof. For general d , we show that the resulting distribution has the same functional form as a Gaussian, and from this we obtain asymptotics for the variance and the number of paths.

Proof of Theorem 1.1. From Lemma 3.2, if we write

$$k = \mu_d(n + 1) + \ell n^{1/2}/2 \tag{3.9}$$

then the probability of $|\ell|$ being at least $n^{1/9}$ is at most $n^{2/9}$, so in the arguments below, we assume $|\ell| \leq n^{1/9}$. In particular, this means that k and $n - k$ are close to $n/2$ with probability tending to 1 as $n \rightarrow \infty$. We are using $n^{1/2}/2$ and not $\sigma_d(n + 1)$, because this way a quantity below will perfectly match the $d = 1$ case.

For m large, Stirling’s Formula states that

$$m! = m^m e^{-m} \sqrt{2\pi m} \left(1 + O\left(\frac{1}{m}\right)\right). \tag{3.10}$$

Thus,

$$\begin{aligned} p_d(k, n + 1) &= \frac{\binom{n}{k}^d}{s_d(n + 1)} = \frac{1}{s_d(n + 1)} \left(\frac{n!}{k!(n - k)!}\right)^d \\ &= \frac{1}{s_d(n + 1)} \left(\frac{\sqrt{2\pi n} n^n}{\sqrt{4\pi^2 k(n - k)} k^k (n - k)^{n - k}} \cdot \frac{(1 + O(\frac{1}{n}))}{(1 + O(\frac{1}{n - k})) (1 + O(\frac{1}{k}))}\right)^d, \end{aligned} \tag{3.11}$$

and the ratio of the big-Oh terms is $1 + O(1/n)$, since k and $n - k$ are approximately $n/2$ (note the big-Oh constant here is allowed to depend on d , which is fixed).

We now turn to the other part of the above expression. If we divide the rest of the quantity in parentheses by 2^n , then we have the probability in one dimension, whose analysis is well-known; thus,

$$p_d(k, n+1) = \frac{2^{nd}n^{d/2}}{s_d(n+1)} \left(\frac{n^n}{2^n k^k (n-k)^{n-k} \sqrt{2\pi k(n-k)}} \right)^d \cdot (1 + O(1/n)). \quad (3.12)$$

The quantity to the d th power converges (up to the normalization factor) to a Gaussian by the Central Limit Theorem for a binomial random variable; for completeness we sketch the proof.

Using that n and $n-k$ are close to $n/2$, we find

$$\begin{aligned} p_{\text{main},1}(k) &:= \frac{n^n}{2^n k^k (n-k)^{n-k} \sqrt{2\pi k(n-k)}} \\ &= \frac{1}{\sqrt{\frac{1}{2}\pi n^2}} \cdot \frac{1}{\left(1 - \frac{\ell\sqrt{n}}{n/2}\right)^{n/2 - \frac{\ell\sqrt{n}}{2} + \frac{1}{2}} \left(1 + \frac{\ell\sqrt{n}}{n/2}\right)^{n/2 + \frac{\ell\sqrt{n}}{2} + \frac{1}{2}}}. \end{aligned} \quad (3.13)$$

Let q_n be the denominator of the second fraction above. We approximate $\log(q_n)$ and then exponentiate to estimate q_n . As $|\ell| \leq n^{1/9}$, when we take the logarithms of the terms in q_n , only the first two terms in the Taylor expansion of $\log(1+u)$ contribute as $n \rightarrow \infty$. Thus,

$$\begin{aligned} \log q_n &= \left(\frac{n}{2} - \frac{\ell\sqrt{n}}{2} + \frac{1}{2}\right) \left(-\frac{\ell}{\sqrt{n}} - \frac{\ell^2}{2n} + O\left(\frac{\ell^3}{n^{3/2}}\right)\right) \\ &\quad + \left(\frac{n}{2} + \frac{\ell\sqrt{n}}{2} + \frac{1}{2}\right) \left(\frac{\ell}{\sqrt{n}} - \frac{\ell^2}{2n} + O\left(\frac{\ell^3}{n^{3/2}}\right)\right) \\ &= \frac{\ell^2}{2} + O\left(n \cdot \frac{n^{1/3}}{n^{3/2}} - \frac{\ell^2}{2n}\right) = \frac{\ell^2}{2} + O\left(n^{-1/6}\right), \end{aligned} \quad (3.14)$$

which implies (since $k = \mu_d(n+1) + \ell\sqrt{n}/2$)

$$q_n = e^{\frac{(k-\mu_d(n+1))^2}{n/2}} e^{O(n^{-1/6})}. \quad (3.15)$$

Thus, collecting our expansions yields, for $|\ell| \leq n^{1/9}$,

$$p_d(k, n+1) = \frac{2^{nd}n^{d/2}}{s_d(n+1)(\pi n^2/2)^{d/2}} e^{-\frac{d(k-\mu_d(n+1))^2}{n/2}} \cdot e^{O(n^{-1/6})}. \quad (3.16)$$

Note the second exponential is negligible as $n \rightarrow \infty$, and the first exponential is that of a Gaussian with mean $\mu_d(n+1)$ and variance $\sigma_d(n+1)^2 = n/4d$. Because this is a probability distribution, it must sum to 1 (the terms with $|\ell|$ large contribute negligibly in the limit), and thus, $2^{nd}/(s_d(n+1)(\pi n/2)^{d/2})$ must converge to the normalization constant of this Gaussian, which is $1/\sqrt{2\pi s_d(n+1)^2}$. In particular, we obtain

$$s_d(n+1) \sim \frac{2^{nd}n^{d/2}}{(\pi n^2/2)^{d/2}} \cdot \sqrt{2\pi n/4d} = 2^{nd} \left(\frac{\pi n}{2}\right)^{-\frac{d}{2} + \frac{1}{2}} d^{-1/2}. \quad (3.17)$$

□

4. FUTURE WORK AND CONCLUDING REMARKS

We could also consider the **Compound Zeckendorf Diagonal Sequence in d dimensions**, which is constructed in a similar way to (1.3) and (1.4), but allows more paths to be legal (explicitly, each step is no longer required to move in all of the dimensions). Whereas the $d = 1$ Compound Zeckendorf Diagonal Sequence is the same as the simple one, the two notions of paths give rise to different sequences when $d = 2$. In that case, the Compound Zeckendorf Diagonal Sequence is denoted $\{z_{a,b}\}_{a=0,b=0}^\infty$, and is constructed as follows.

- (1) Set $z_{1,1} := 1$.
- (2) Iterate through the natural numbers. For each such number, check if any path of distinct numbers without upward or rightward movements sums to the number. If no such path exists, add the number to the sequence so that it is added to the shortest unfilled diagonal, moving from the bottom right to the top left.
- (3) If a new diagonal must begin to accommodate a new number, set the value $z_{k,1}$ to be that number, where k is minimized so that $z_{k,1}$ has not yet been assigned.

The difference between this and the Simple Zeckendorf Diagonal Sequence is that we now allow movement in just one direction. This greatly complicates the combinatorial analysis because now the simultaneous movements in different dimensions depend on each other. In particular, if a step contains a movement in one direction, it no longer needs to contain a movement in other directions to be regarded as a legal step. In (4.1), we illustrate several diagonals' worth of entries, where the elements are always added in increasing order.

6992
2200	6054
954	2182	5328
364	908	2008	5100
138	342	862	1522	4966
44	112	296	520	1146	2952
16	38	94	184	476	1102	2630
4	10	22	56	168	370	1052	2592
1	2	6	18	46	140	366	1042	2270

(4.1)

Just as in (1.4), uniqueness of decompositions does not hold in the compound case. For instance, $112 + 38 + 10$ and $140 + 18 + 2$ are both legal decompositions of 160 in (4.1). Moreover, just like the Simple Zeckendorf Diagonal Sequences (1.3) and (1.4), Compound Zeckendorf Diagonal Sequences can be built in higher dimensions with multiple ways of formulating how to add terms to the sequence.

Many of the articles in the literature use combinatorial methods and manipulations of binomial coefficients to obtain similar results (see, for instance, [12, 23, 27]). Thus, a question worth future study is to extend the combinatorial variance calculation to d dimensions (see Lemma A.3).

Finally, similar to [3, 20] and related work, we can investigate the distribution of gaps between summands in legal paths. One can readily obtain explicit combinatorial formulas for the probability of a given gap; the question is whether or not nice limits exist in this case, as they do for the one-dimensional recurrences previously studied.

APPENDIX A. DERIVATION OF MEAN AND STANDARD DEVIATION FOR SIMPLE JUMP PATHS

Before we compute the mean and the standard deviation, we record a simple combinatorial lemma that will be needed later.

Lemma A.1. *For n a positive integer,*

$$\sum_{k=0}^n k^2 \binom{n}{k}^2 = n^2 \binom{2n-2}{n-1}. \quad (\text{A.1})$$

Proof. Expanding the binomial coefficient and canceling ks yields

$$\sum_{k=0}^n k^2 \binom{n}{k}^2 = \sum_{k=0}^n k^2 \left(\frac{n!}{k!(n-k)!} \right)^2 = \sum_{k=1}^n n^2 \left(\frac{(n-1)!}{(k-1)!(n-k)!} \right)^2. \quad (\text{A.2})$$

Shifting indices, we can rewrite the above as

$$\sum_{k=0}^n k^2 \binom{n}{k}^2 = n^2 \sum_{\ell=0}^{n-1} \binom{n-1}{\ell}^2 = n^2 \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} \binom{n-1}{n-1-\ell}, \quad (\text{A.3})$$

and the sum equals $\binom{2n-2}{n-1}$ (it is the number of ways to choose $n-1$ objects from $2n-2$, where we consider $n-1$ of the items to be in one set and the remaining $n-1$ in another). \square

Lemma A.2 (Mean for Simple Jump Path Distribution). *If $\mu_d(i)$ denotes the mean number of steps in a d -dimensional simple jump path from (i, i, \dots, i) to the origin, then*

$$\mu_d(n+1) = \frac{1}{2}n + 1. \quad (\text{A.4})$$

Proof. By the definition of the first moment,

$$\begin{aligned} \mu_d(n+1) &= \frac{\sum_{k=1}^{n+1} k \cdot t_d(k, n+1)}{s_d(n+1)} \\ &= \frac{\sum_{k=0}^n (k+1) t_d(k+1, n+1)}{s_d(n+1)} = \frac{\sum_{k=0}^n k \binom{n}{k}^d + s_d(n+1)}{s_d(n+1)}. \end{aligned} \quad (\text{A.5})$$

We complete the proof based on the parity of n . We first assume n is odd. Then,

$$\sum_{k=0}^n k \binom{n}{k}^d = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \left[k \binom{n}{k}^d + (n-k) \binom{n}{n-k}^d \right] = n \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{k}^d. \quad (\text{A.6})$$

Notice that by the symmetry of binomial coefficients,

$$\sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{k}^d = \sum_{k=\lceil \frac{n}{2} \rceil}^n \binom{n}{k}^d, \quad (\text{A.7})$$

so

$$n \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{k}^d = \frac{1}{2}n \sum_{k=0}^n \binom{n}{k}^d = \frac{1}{2}n s_d(n+1), \quad (\text{A.8})$$

and substituting into (A.5) completes the proof in this case.

Now, we consider n even. A similar analysis as in the previous case works, except we need to address the term where $k = n/2$, which is matched with itself:

$$\begin{aligned} \sum_{k=0}^n k \binom{n}{k}^d &= \frac{n}{2} \binom{n}{n/2}^d + \sum_{k=0}^{\frac{n}{2}-1} k \binom{n}{k}^d + \sum_{k=\frac{n}{2}+1}^n k \binom{n}{k}^d \\ &= \frac{n}{2} \binom{n}{n/2}^d + \sum_{k=0}^{\frac{n}{2}-1} \left[k \binom{n}{k}^d + (n-k) \binom{n}{n-k}^d \right] \\ &= \frac{n}{2} \binom{n}{n/2}^d + n \sum_{k=0}^{\frac{n}{2}-1} \binom{n}{k}^d. \end{aligned} \tag{A.9}$$

Again, utilizing the symmetry of binomial coefficients,

$$\sum_{k=0}^{\frac{n}{2}-1} \binom{n}{k}^d = \sum_{k=\frac{n}{2}+1}^n \binom{n}{k}^d, \tag{A.10}$$

so (A.9) is equivalent to

$$\frac{n}{2} \binom{n}{n/2}^d + \frac{n}{2} \sum_{k \in \{0,1,\dots,n\} \setminus \{n/2\}} \binom{n}{k}^d = \frac{n}{2} \sum_{k=0}^n \binom{n}{k}^d = \frac{n}{2} s_d(n+1), \tag{A.11}$$

completing the proof. □

Lemma A.3 (Standard Deviation for Two-Dimensional Simple Jump Paths). *If $\sigma_2(i)$ represents the standard deviation for the number of steps in a simple jump path in d dimensions from (i, i) to the origin, then*

$$\sigma_2(n+1) = \frac{n}{2\sqrt{2n-1}}. \tag{A.12}$$

Because the variance in the one-dimensional case is well-known (it is the variance of a binomial random variable), we provide details only for $d = 2$. As remarked earlier, the combinatorial approach taken below does not generalize to higher d .

Proof. We use the simple closed form expression for $s_2(n+1)$, namely that it equals $\binom{2n}{n}$. By the definition of the second standardized moment and use of (A.4) where $d = 2$, we have

$$\sigma_2(n+1)^2 = \frac{\sum_{k=1}^{n+1} k^2 \binom{n}{k-1}^2}{\binom{2n}{n}} - \left(\frac{1}{2}n+1\right)^2. \tag{A.13}$$

Shifting the index of summation to start at $k = 0$ and expanding yields

$$\begin{aligned} \sigma_2(n+1)^2 &= \frac{\sum_{k=0}^n (k+1)^2 \binom{n}{k}^2}{\binom{2n}{n}} - \left(\frac{1}{2}n+1\right)^2 \\ &= \frac{\sum_{k=0}^n k^2 \binom{n}{k}^2}{\binom{2n}{n}} + \frac{\sum_{k=0}^n 2k \binom{n}{k}^2}{\binom{2n}{n}} + \frac{\sum_{k=0}^n \binom{n}{k}^2}{\binom{2n}{n}} - \left(\frac{1}{2}n+1\right)^2 \\ &= \frac{\sum_{k=0}^n k^2 \binom{n}{k}^2}{\binom{2n}{n}} + 2\mu_2(n+1) - \frac{\sum_{k=0}^n \binom{n}{k}^2}{\binom{2n}{n}} - \left(\frac{1}{2}n+1\right)^2. \end{aligned} \tag{A.14}$$

Using (3.2) for the mean and recalling that $\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}$, we have

$$\begin{aligned} \sigma_2(n+1)^2 &= \frac{\sum_{k=0}^n k^2 \binom{n}{k}^2}{\binom{2n}{n}} + 2 \left(\frac{1}{2}n + 1 \right) - 1 - \left(\frac{1}{2}n + 1 \right)^2 \\ &= \frac{\sum_{k=0}^n k^2 \binom{n}{k}^2}{\binom{2n}{n}} - \frac{n^2}{4}. \end{aligned} \tag{A.15}$$

Substituting the identity (A.1) from Lemma A.1 into (A.15) gives

$$\begin{aligned} \sigma_2(n+1)^2 &= \frac{n^2 \binom{2n-2}{n-1}}{\binom{2n}{n}} - \frac{n^2}{4} \\ &= \frac{n^3}{4n-2} - \frac{n^2}{4} = \frac{n^2}{8n-4}. \end{aligned} \tag{A.16}$$

Taking the square root of both sides of (A.16) gives the desired result. \square

We remark on the difficulty in generalizing the above argument to arbitrary d . The problem is in (A.1). There it was crucial that $d = 2$, because we then canceled the k^2 with the two factors of k in the denominator. In higher dimensions, we do not have such perfect alignment.

REFERENCES

- [1] I. Ben-Ari and S. Miller, *A probabilistic approach to generalized Zeckendorf decompositions*, SIAM Journal on Discrete Mathematics, **30** (2016), no. 2, 1302–1332.
- [2] A. Best, P. Dynes, X. Edelsbrunner, B. McDonald, S. Miller, K. Tor, C. Turnage-Butterbaugh, and M. Weinstein, *Gaussian behavior of the number of summands in Zeckendorf decompositions in small intervals*, The Fibonacci Quarterly, **52.5** (2014), 47–53.
- [3] A. Bower, R. Insoft, S. Li, S. Miller, and P. Tosteson, *The distribution of gaps between summands in generalized Zeckendorf decompositions*, Journal of Combinatorial Theory, **135** (2015), 130–160.
- [4] J. L. Brown, Jr., *Zeckendorf’s Theorem and some applications*, The Fibonacci Quarterly, **2.3** (1964), 163–168.
- [5] M. Catral, P. Ford, P. Harris, S. Miller, and D. Nelson, *Generalizing Zeckendorf’s Theorem: The Kentucky sequence*, The Fibonacci Quarterly, **52.5** (2014), 68–90.
- [6] M. Catral, P. Ford, P. E. Harris, S. J. Miller, and D. Nelson, *Legal decompositions arising from non-positive linear recurrences*, The Fibonacci Quarterly, **54.4** (2016), 348–365.
- [7] M. Catral, P. Ford, P. E. Harris, S. J. Miller, D. Nelson, Z. Pan, and H. Xu, *New behavior in legal decompositions arising from non-positive linear recurrences*, The Fibonacci Quarterly, **55.3** (2017), 252–275 (expanded arXiv version: <http://arxiv.org/pdf/1606.09309>).
- [8] D. E. Daykin, *Representation of natural numbers as sums of generalized Fibonacci numbers*, J. London Mathematical Society, **35** (1960), 143–160.
- [9] P. Demontigny, T. Do, A. Kulkarni, S. Miller, D. Moon, and U. Varma, *Generalizing Zeckendorf’s Theorem to f -decompositions*, Journal of Number Theory, **141** (2014), 136–158.
- [10] R. Dorward, P. Ford, E. Fourakis, P. Harris, S. Miller, E. Palsson, and H. Paugh, *New behavior in legal decompositions arising from non-positive linear recurrences*, The Fibonacci Quarterly, **55.3** (2017), 252–275.
- [11] M. Drmota and J. Gajdosik, *The distribution of the sum-of-digits function*, J. Théor. Nombres Bordeaux, **10** (1998), no. 1, 17–32.
- [12] S. Eger, *Stirling’s approximation for central extended binomial coefficients*, American Mathematical Monthly, **121** (2014), no. 4, 344–349, <http://arxiv.org/pdf/1203.2122.pdf>.
- [13] P. Filipponi, P. J. Grabner, I. Nemes, A. Pethö, and R. F. Tichy, *Corrigendum to: “Generalized Zeckendorf expansions”*, Appl. Math. Lett., **7** (1994), no. 6, 25–26.
- [14] A. S. Fraenkel, *Systems of numeration*, Amer. Math. Monthly **92** (1985), no. 2, 105–114.
- [15] P. J. Grabner, R. F. Tichy, I. Nemes, and A. Pethö, *Generalized Zeckendorf expansions*, Appl. Math. Lett., **7** (1994), no. 2, 25–28.

THE FIBONACCI QUARTERLY

- [16] N. Hamlin, *Representing positive integers as a sum of linear recurrence sequences*, The Fibonacci Quarterly, **50.2** (2012), 99–105.
- [17] N. Hamlin and W. A. Webb, *Representing positive integers as a sum of linear recurrence sequences*, The Fibonacci Quarterly, **50.2** (2012), 99–105.
- [18] V. E. Hoggatt, *Generalized Zeckendorf Theorem*, The Fibonacci Quarterly, **10.1** (1972), (Special Issue on Representations), 89–93.
- [19] T. J. Keller, *Generalizations of Zeckendorf's Theorem*, The Fibonacci Quarterly, **10.1** (1972), (Special Issue on Representations), 95–102.
- [20] M. Kologlu, G. Kopp, S. Miller, and Y. Wang, *On the number of summands in Zeckendorf decompositions*, Journal of Number Theory, **49** (2011), no. 2, 116–130.
- [21] M. Lamberger and J. M. Thuswaldner, *Distribution properties of digital expansions arising from linear recurrences*, Math. Slovaca **53** (2003), no. 1, 1–20.
- [22] C. G. Lekkerkerker, *Voorstelling van natuurlyke getallen door een som van getallen van Fibonacci*, Simon Stevin, **29** (1951–1952), 190–195.
- [23] T. Lengyel, *A counting based proof of the generalized Zeckendorf's Theorem*, The Fibonacci Quarterly, **44.4** (2006), 324–325.
- [24] R. Li and S. J. Miller, *A collection of Central Limit Type results in generalized Zeckendorf decompositions*, Proceedings of the Seventeenth International Conference on Fibonacci Numbers and Their Applications, The Fibonacci Quarterly, **55.5** (2017), 105 – 114.
- [25] S. J. Miller, *The Probability Lifesaver*, Princeton University Press, 2017, pp. 752.
- [26] S. Miller and Y. Wang, *From Fibonacci numbers to Central Limit Type Theorems*, Journal of Combinatorial Theory, Series A **119** (2012), no. 7, 1398–1413.
- [27] S. Miller and Y. Wang, *Gaussian behavior in generalized Zeckendorf decompositions*, Combinatorial and Additive Number Theory, CANT 2011 and 2012 (Melvyn B. Nathanson, editor), Springer Proceedings in Mathematics & Statistics, (2014), 159–173.
- [28] W. Steiner, *Parry expansions of polynomial sequences*, Integers, **2** (2002), Paper A14.
- [29] W. Steiner, *The joint distribution of greedy and lazy Fibonacci expansions*, The Fibonacci Quarterly, **43.1** (2005), 60–69.
- [30] E. Zeckendorf, *Représentation des nombres naturels par une somme des nombres de Fibonacci ou de nombres de Lucas*, Bulletin de la Société Royale des Sciences de Liège, **41** (1972), 179–182.

MSC2010: 11B02 (primary), 05A02 (secondary)

HIGH SCHOOL AFFILIATED TO RENMIN UNIVERSITY, ZHONGGUANCUN ROAD 37, HAIDIAN DISTRICT, BEIJING, CHINA

Email address: 2821524834@qq.com

THE EUREKA PROGRAM

Email address: chenziyang20010420@126.com

THE EUREKA PROGRAM

Email address: 530824689@qq.com

CHELTENHAM LADIES' COLLEGE, GLOUCESTERSHIRE, UNITED KINGDOM GL50 3EP

Email address: xycindyjiang@gmail.com

DEPARTMENT OF MATHEMATICS AND STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA 01267

Current address: Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213

Email address: sjm1@williams.edu

DEPARTMENT OF MATHEMATICAL SCIENCES, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PA 15213

Email address: jsiktar@andrew.cmu.edu

THE TAFT SCHOOL, WATERTOWN, CT 06795

Email address: peter0201yu@gmail.com