CENTRAL LIMIT THEOREMS FOR COMPOUND PATHS ON THE TWO-DIMENSIONAL LATTICE

EVAN FANG, JONATHAN JENKINS, ZACK LEE, DANIEL LI, ETHAN LU, STEVEN J. MILLER, DILHAN SALGADO, AND JOSHUA M. SIKTAR

Abstract. Zeckendorf proved that every integer can be written uniquely as a sum of non-consecutive Fibonacci numbers $\{F_n\}$, with later researchers showing that the distribution of the number of summands needed for such decompositions of integers in $[F_n, F_{n+1})$ converges to a Gaussian as $n \to \infty$. Decomposition problems have been studied extensively for a variety of different sequences and notions of legal decompositions; for the Fibonacci numbers, a legal decomposition is one for which each summand is used at most once and no two consecutive summands may be chosen. Chen, et al. [11] generalized earlier work to d-dimensional lattices of positive integers; there, a legal decomposition was defined as a path such that every point chosen had each component strictly less than the same component of the previous chosen point in the path. They were able to prove Gaussianity results despite the lack of uniqueness of the decompositions; however, one would expect their results to hold in the more general case where some components are identical. The strictly decreasing assumption was needed in that work to obtain simple, closed form combinatorial expressions, which could then be well approximated and lead to the limiting behavior. In this work, we remove that assumption through inclusion-exclusion arguments. These lead to more involved combinatorial sums; using generating functions and recurrence relations, we obtain tractable forms in two dimensions and prove Gaussianity again. A more involved analysis should work in higher dimensions.

1. Introduction

Among the many fascinating properties of the Fibonacci numbers is the following classical observation, credited to Zeckendorf [40]: every positive integer admits a unique representation as a sum of nonadjacent Fibonacci numbers $\{F_n\}$, where[1] $F_1 = 1$, $F_2 = 2$, and $F_{n+1} = F_n + F_{n-1}$. We can treat this property as an equivalent definition of the Fibonacci numbers: they are the only sequence from which every positive integer can be decomposed uniquely as a sum of nonadjacent terms. It turns out that there is often a relationship between rules for legal decompositions and sequences $\{G_n\}$, and the literature is now filled with many results on properties of the summands in legal decompositions of numbers in intervals $[G_n, G_{n+1})$ as $n \to \infty$. These range from the mean number of summands growing linearly, with the factor related to the roots of the characteristic polynomial of the recurrence, to the distribution of the number of summands converging to a Gaussian, to the distribution of gaps between summands; see for example [4, 6, 7, 12, 13, 19, 20, 21, 22, 24, 23, 25, 29, 32, 31, 34, 35, 39, 38] and the references therein. It is the notions of distributions of summands of positive integer decompositions that most directly tie to our work in this paper, which introduces tools from enumerative combinatorics to count the number of possible decompositions on a lattice.

Most of the sequences studied to date have been one-dimensional. Additional sequences, such as those in [9, 10], appear two-dimensional but can be converted into one-dimensional sequences and attacked using existing techniques. The interest in truly higher-dimensional

This work was supported by NSF Grant DMS1561945, Carnegie Mellon and Williams College.

[1]If we started with $F_0 = 0$ and $F_1 = 1$, then $F_2 = 1$ and we trivially lose uniqueness.

sequences has been amplified by the ability to connect the counting of special lattice paths to the number of ways to decompose positive integers as sums of smaller positive integers with respect to the geometry of the lattice, not to mention the combinatorics that arise, which are interesting in their own right. This underlying principle motivated Chen, et al. [11] to consider a true multidimensional sequence by looking at paths among lattice points with non-negative integer coefficients. They defined a legal decomposition in d-dimensions to be a finite collection of lattice points for which

(1) each point is used at most once, and
(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 a point in the decomposition and the next one).

They called the path of chosen lattice points a **simple jump path**; at each step, every component was *strictly less than* the corresponding component of the previous step. One can construct a sequence on the lattice in many ways. For example, in two dimensions one can go along diagonal paths parallel to the line $y = -x$, adding at each lattice point the first number that cannot be legally represented. This situation is slightly more involved in higher dimensions, although for most of the problems studied, the values of the ordered points do not matter; what matters is the geometry of the lattice walks. In (1.1), we illustrate several diagonals' worth of entries for the given algorithm when $d = 2$. Unlike the one-dimensional case where the Fibonacci sequence is used, we now find that uniqueness of integer decompositions fails (for example, 25 has two legal decompositions: $20 + 5$ and $24 + 1$).

$$\begin{array}{cccccccccc}
280 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
157 & 263 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
84 & 155 & 259 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
50 & 82 & 139 & 230 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
28 & 48 & 74 & 123 & 198 & \cdots & \cdots & \cdots & \cdots & \cdots \\
14 & 24 & 40 & 66 & 107 & 184 & \cdots & \cdots & \cdots & \cdots \\
7 & 12 & 20 & 33 & 59 & 100 & 171 & \cdots & \cdots & \cdots \\
3 & 5 & 9 & 17 & 30 & 56 & 93 & 160 & \cdots & \cdots \\
1 & 2 & 4 & 8 & 16 & 29 & 54 & 90 & 159 & \cdots
\end{array} \qquad (1.1)$$

The restriction that every coordinate must decrease in a simple jump path is a severe one. For example, in (1.1), 5 had to be added at the location $(2, 2)$ because the sequence $(3, 1)$, $(1, 1)$ resulting in the sum $4 + 1$, only decreases horizontally and is invalid. However, this strictly decreasing condition was needed in [11] to obtain simple closed form combinatorial expressions whose approximations led to a proof that the limiting behavior of the distribution of the number of summands converges to a Gaussian. (The distribution of the gaps was accomplished by Borade, et al. in [5]; unlike other gap distribution results, for this problem there is an interesting twist as the gaps are vectors and not integers.) Thus, it is precisely the analysis of these decompositions that makes the enumeration of lattice paths a worthwhile problem to study, in the finite case, and in the asymptotic case where we look at how many terms are needed, on average, to decompose large integers.

Replacing the strict decreasing condition in [11] with the nonstrict decreasing condition, we obtain what we call a **generalized jump path**. Formally, we define a **generalized jump path** of length n starting at $(p_1, p_2, \dots, p_d)$ to be a sequence of points $\{(x_{i,1}, x_{i,2}, \dots, x_{i,d})\}_{i=0}^{n}$ such that:

- $(x_{0,1}, x_{0,2}, \dots, x_{0,d}) = (p_1, p_2, \dots, p_d)$,

- for all i and j, we have $x_{i,j} \geq x_{i+1,j}$,
- for all i, we have $(x_{i,1}, x_{i,2}, \ldots, x_{i,d}) \neq (x_{i+1,1}, x_{i+1,2}, \ldots, x_{i+1,d})$ [2],
- if $i < n$, for all j, $x_{i,j} > 0$, and
- $(x_{n,1}, x_{n,2}, \ldots, x_{n,d}) = (0, 0, \ldots, 0)$.

These conditions imply that for each point in the sequence, at least one coordinate must decrease, while the remaining coordinates cannot increase. For convenience, we include the requirement that all paths end at the origin outside of the lattice, meaning that the lattice in (1.2) has row and column indices starting at 1. With this requirement, moving directly to the origin is equivalent to not choosing any additional points in a path and is considered a legal option. Below is the number of generalized jump paths in two dimensions from the point (i, j) to $(0, 0)$ for $i, j \leq 4$.

$$\begin{array}{ccccc} \vdots & \vdots & \vdots & \vdots & \iddots \\ 4 & 20 & 76 & 252 & \cdots \\ 2 & 8 & 26 & 76 & \cdots \\ 1 & 3 & 8 & 20 & \cdots \\ 1 & 1 & 2 & 4 & \cdots \end{array} \tag{1.2}$$

We can now find a recursive formula that allows us to calculate the total number of generalized jump paths starting at $p = (p_1, p_2, \ldots, p_d)$. Letting $S((p_1, p_2, \ldots, p_d))$ represent the total number of paths starting at p, we may do the following: partition all the generalized jump paths from (p) by the location of their first step. Either the path goes directly from (p) to the origin, or (p) first goes to some other lattice point (a). Note that (a) must have at least one coordinate, say (a_i), such that $(a_i < p_i)$, hence, we know that $(a \in \{[0, p_1] \times [0, p_2] \times \cdots \times [0, p_d]\} \setminus \{p\})$, where $[0, p_i] = \{0, 1, \ldots, p_i\}$ for each $i \leq d$. By definition, there are $(S(a))$ paths from the point a to the origin. Summing over all possible (a), we find a formula for $(S(p))$:

$$S(p) \;=\; \sum_{\substack{a \neq p,\\ a \in [0,p_1] \times [0,p_2] \times \cdots [0,p_d]}} S(a). \tag{1.3}$$

Using this formula and some further analysis, we are able to asymptotically analyze the total number of generalized jump paths and the number of such paths with a fixed length k. Our main result is as follows.

Theorem 1.1. *Let $X_{p,q}$ be the random variable denoting the number of generalized jump paths starting at a point $(p, q) \in \mathbb{N}^2$ and ending at the origin. Suppose $p = m$ and $q = cm$ for some $m \in \mathbb{N}^+$ and $c \geq 1$ rational. Then $X_{p,q}$ converges to a Gaussian as $m \to \infty$, with mean $\frac{q+p+\sqrt{p^2+6pq+q^2}}{4}$ and variance $\frac{p+q}{8} + \frac{(p+q)^2}{8\sqrt{p^2+6pq+q^2}}$.*

In Section 2, we introduce some notation for our problem and prove some basic properties of unrestricted generalized jump paths. In Section 3, we use these properties in conjunction with various analytical and combinatorial methods to obtain a generating function for the number of paths to a fixed point as a function of path length. Using that result, we prove the Gaussianity in the limit of the number of summands in decompositions in Section 4 (Theorem 1.1). We use a method similar to that found in [11]; the difficulty in the argument is in determining a good count of the number of paths and, from that, a good estimate for the number of paths of a given length. We conclude with a brief discussion of future questions to study.

[2] A natural future question would be to remove this condition, which would allow the same point to be used multiple times in a decomposition. See Section 5 for more information.

2. Properties of Generalized Jump Paths

Recall that we define a **generalized jump path** of length n that starts at $(p_1, p_2, \ldots, p_d)$ to be a sequence of points $\{(x_{i,1}, x_{i,2}, \ldots, x_{i,d})\}_{i=0}^{n}$ such that:

- $(x_{0,1}, x_{0,2}, \ldots, x_{0,d}) = (p_1, p_2, \ldots, p_d)$,
- for all i and j, we have $x_{i,j} \geq x_{i+1,j}$,
- for all i, we have $(x_{i,1}, x_{i,2}, \ldots, x_{i,d}) \neq (x_{i+1,1}, x_{i+1,2}, \ldots, x_{i+1,d})$,
- If $i < n$, for all j, $x_{i,j} > 0$, and
- $(x_{n,1}, x_{n,2}, \ldots, x_{n,d}) = (0, 0, \ldots, 0)$.

Under these circumstances, a path can have a length as short as 1 or as long as nd. Let $g((p_1, p_2, \ldots, p_d), n)$ be the number of such paths of length n starting at $(p_1, p_2, \ldots, p_d)$.
We also define an **unrestricted generalized jump path** of length n to be a sequence of points $\{(x_{i,1}, x_{i,2}, \ldots, x_{i,d})\}_{i=0}^{n}$ such that:

- $(x_{0,1}, x_{0,2}, \ldots, x_{0,d}) = (p_1, p_2, \ldots, p_d)$,
- for all i and j we have $x_{i,j} \geq x_{i+1,j}$,
- for all i we have $(x_{i,1}, x_{i,2}, \ldots, x_{i,d}) \neq (x_{i+1,1}, x_{i+1,2}, \ldots, x_{i+1,d})$, and
- for all i, j, $x_{i,j} > 0$,

Analogously to the definition of g, we let $u((p_1, p_2, \ldots, p_d), n)$ be the number of *unrestricted* generalized jump paths of length n starting from the point $(p_1, p_2, \ldots, p_d)$. Note that unrestricted generalized jump paths are simply generalized jump paths with the last restriction lifted (i.e., the sequence does not need to end at the bottom left corner).

We now establish and prove two basic properties for g and u, the first of which was alluded to in our definition of u.

Lemma 2.1 (Unrestricted-Restricted Relationship). *Let $v = (p_1, p_2, \ldots, p_d)$. For all $n \in \mathbb{N}$,*

$$u(v, n) \;=\; g(v, n) + g(v, n + 1). \tag{2.1}$$

Proof. The set of unrestricted jump paths of length n that do not end at $(0, 0, \ldots, 0)$ bijects to the set of restricted jump paths of length $n + 1$ that end at $(0, 0, \ldots, 0)$. In particular, any path of length n not ending at the origin can have the origin appended to the end of the path; likewise, any path of length $n + 1$ ending at the origin can have the origin removed from the end of the path. This immediately implies the result. □

Theorem 2.2 (Two-Dimensional Path Recurrence). *For all $p, q, n \in \mathbb{N}^+$,*

$$\begin{aligned} u((p, q), n) \;&= u((p, q - 1), n) + u((p, q - 1), n - 1) \\ &\quad + u((p - 1, q), n) + u((p - 1, q), n - 1) \\ &\quad - u((p - 1, q - 1), n) - u((p - 1, q - 1), n - 1). \end{aligned} \tag{2.2}$$

Proof. Let the functions $\ell((p, q), n)$, $d((p, q), n)$, and $m((p, q), n)$ denote the number of unrestricted jump paths that take n jumps starting from the point (p, q), where the first jump of these paths is directly *left*, directly *down*, and both *left and down*, respectively. With this definition, we can partition the unrestricted jump paths with a fixed starting point based on the direction of the first step in the $\mathbb{N}^2$ lattice. From this we realize

$$\begin{aligned}
\ell((p,q),n) &= \ell((p-1,q),n) + u((p-1,q),n-1) \\
d((p,q),n) &= d((p,q-1),n) + u((p,q-1),n-1) \\
m((p,q),n) &= d((p-1,q),n) + m((p-1,q),n) \\
&= \ell((p,q-1),n) + m((p,q-1),n) \\
&= u((p-1,q-1),n-1) + u((p-1,q-1),n). \qquad (2.3)
\end{aligned}$$

Notice, by the definitions of the functions $\ell((p,q),n)$, $d((p,q),n)$, and $m((p,q),n)$, that

$$u((p,q),n) = \ell((p,q),n) + d((p,q),n) + m((p,q),n). \qquad (2.4)$$

So, by using the identity in (2.4), we can rewrite $u((p,q),n)$ as:

$$\begin{aligned}
u((p,q),n) &= \ell((p-1,q),n) + u((p-1,q),n-1) \\
&\quad + d((p,q-1),n) + u((p,q-1),n-1) \\
&\quad + d((p-1,q),n) + m((p-1,q),n) \\
&\quad + \ell((p,q-1),n) + m((p,q-1),n) \\
&\quad - (u((p-1,q-1),n-1) + u((p-1,q-1),n)) \\
&= \ell((p-1,q),n) + d((p-1,q),n) + m((p-1,q),n) \\
&\quad + d((p,q-1),n) + \ell((p,q-1),n) + m((p,q-1),n) \\
&\quad + u((p-1,q),n-1) + u((p,q-1),n-1) \\
&\quad - (u((p-1,q-1),n-1) + u((p-1,q-1),n)) \\
&= u((p-1,q),n) + u((p,q-1),n) \\
&\quad + u((p-1,q),n-1) + u((p,q-1),n-1) \\
&\quad - u((p-1,q-1),n-1) - u((p-1,q-1),n), \qquad (2.5)
\end{aligned}$$

which shows our desired result. □

This new recurrence relation will serve as an agent in the creation of an explicit generating function for the number of generalized jump paths starting at a fixed point $(p,q) \in \mathbb{N}^2$ and ending at the origin.

3. Two-Dimensional Generating Function

For every lattice point in $\mathbb{N}^2$, there is a two-dimensional generating function for the lengths of the paths starting at that point. Explicitly, we denote $F_{p,q}(x)$ to be

$$F_{p,q}(x) = \sum_{k=0}^{p+q} u((p,q),k)x^k. \qquad (3.1)$$

Our main result in this section is an alternative form for $F_{p,q}$ that is more readily studied using asymptotic techniques.

Theorem 3.1. *If $p \leq q$, then*

$$F_{p,q}(x) = (1+x)^p \sum_{k=0}^{q} \binom{q}{k}\binom{p+k}{k} x^k. \qquad (3.2)$$

We will provide two proofs. The first is a pure generating function approach that only proves the claim for the case $p = q = n$, whereas the second is purely enumerative and proves the theorem in generality. For technical convenience, one often analyzes all paths starting at a point on the main diagonal; thus, if we restrict our investigation to this case, the simpler first proof suffices.

3.1. **Pure Generating Functions Method.** Consider the generating function

$$B(x,y,z) \;=\; \sum_{p,q,k\in\mathbb{Z}^+} u((p,q),k)x^p y^q z^k. \tag{3.3}$$

Using Theorem 2.2, we see that

$$B(x,y,z) \;=\; 1 + (1+z)(x+y-xy)B(x,y,z), \tag{3.4}$$

which implies that

$$B(x,y,z) \;=\; \frac{1}{1-(1+z)(x+y-xy)}. \tag{3.5}$$

Now, we use a method adapted from [37] to determine the central terms of the infinite summation expansion of (3.5). The idea is to use a change of variables to make $B(x,y,z)$ easier to transform into a geometric series. To this end, we define

$$D(x,s,u) \;=\; B(s,x/s,z), \tag{3.6}$$

and let $u = 1+z$; thus from (3.6), we compute

$$D(x,s,u) \;=\; \frac{1}{1-u(s+x/s-x)} \;=\; \frac{-s/u}{s^2 - s(x+1/u)+x}. \tag{3.7}$$

Our generating function for the central terms will be the s^0 coefficient of $D(x,s,u)$. We denote the two solutions for s in the denominator of $D(x,s,u)$ as α and β, where, for ease of reading, we suppress the arguments of these functions as follows:

$$\alpha \;=\; \frac{ux+1-\sqrt{u^2x^2-4u^2x+2ux+1}}{2u}, \quad \beta \;=\; \frac{ux+1+\sqrt{u^2x^2-4u^2x+2ux+1}}{2u}. \tag{3.8}$$

Using the method of partial fractions on (3.7), we find that

$$D(x,s,u) \;=\; \frac{1}{u(\beta-\alpha)}\left(\frac{\alpha}{s-\alpha} - \frac{\beta}{s-\beta}\right). \tag{3.9}$$

We expand each term in (3.9) as a geometric series, which is applicable only within the radii of convergence of the parameters; this ensures that the geometric series produced converge and that the subsequent expressions we study are well-defined. We obtain

$$D(x,s,u) \;=\; \frac{1}{u(\beta-\alpha)}\left(\sum_{m\geq 1}\alpha^m s^{-m} + \sum_{m\geq 0}\beta^{-m}s^m\right). \tag{3.10}$$

The s^0 term in (3.10) is just $\frac{1}{u(\beta-\alpha)}$ as $\beta^{-0}s^0 = 1$. Because $\beta - \alpha = \frac{\sqrt{u^2x^2-4u^2x+2ux+1}}{u}$ and the generating function for the central terms will be the s^0 coefficient of $D(x,s,u)$, we can substitute into the $n = 0$ term of (3.10) into (3.7):

$$B(x,x,z) \;=\; \frac{1}{\sqrt{u^2x^2-4u^2x+2ux+1}} \;=\; \frac{1}{\sqrt{(ux+1)^2-4u^2x}}, \tag{3.11}$$

where we used that $u = z+1$.

We now represent $1/\sqrt{(ux+1)^2-4u^2x}$ as a power series of the form $\sum_{i=0} b_i(z)x^i$. Differentiating with respect to x, we obtain the recurrence relation

$$b_i \;=\; \begin{cases} 1 & i=0 \\ 1+3z+2z^2 & i=1 \\ \dfrac{(2i-1)(1+2z)(1+z)b_{i-1}-(1+z)^2(i-1)b_{i-2}}{i} & i\geq 2. \end{cases} \tag{3.12}$$

For sake of brevity, we define a_i such that $a_i(1+z)^i = b_i$; then, this sequence satisfies the recurrence

$$a_i \;=\; \begin{cases} 1 & i=0 \\ 1+2z & i=1 \\ \dfrac{(2i-1)(1+2z)a_{i-1}-(i-1)a_{i-2}}{i} & i\geq 2. \end{cases} \tag{3.13}$$

The solution to this recurrence relation is $a_i = P_i(2z+1)$, where P_i is the ith Legendre Polynomial (see [26]).

$$P_l(t) \;=\; \sum_{k=0}^{\ell}\binom{\ell}{k}\binom{-\ell-1}{k}\left(\frac{1-t}{2}\right)^k. \tag{3.14}$$

It follows that

$$a_n \;=\; P_n(2z+1) \;=\; \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}z^k. \tag{3.15}$$

Upon renaming the variables,

$$F_{n,n}(x) \;=\; b_n(x) \;=\; (1+x)^n a_n(x) \;=\; (1+x)^n\sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}x^k, \tag{3.16}$$

which proves Theorem 3.1 in the $p=q=n$ case, as desired.

3.2. **Combinatorial Method.** We attempt to obtain a more compact formula for $g(p,n)$ by first relaxing our constraint to allow for paths with stationary points (where consecutive points are allowed to be exactly the same), and then later correcting for the over-counting. We do this because it is significantly easier to count the total number of paths with this relaxation in place.

Let $r(p,n)$ denote the number of paths from p to the origin of length n with this relaxed constraint. Arguing as in the Stars and Bars problem[3], it follows that

$$r(p,n) \;=\; \prod_{i=1}^{d}\binom{p_i+n-1}{p_i}. \tag{3.17}$$

[3]The number of ways to divide N identical items into G groups, where all that matters is how many items are placed in a group, is $\binom{N+G-1}{G}$. To see this, consider $N+G-1$ items in a line; choosing $G-1$ of these partitions the remaining N items into G sets. The first set is all the elements up to the first chosen one, and so on. Thus, there is a one-to-one correspondence between the two counting problems. This method was first used for Zeckendorf decomposition problems in [27], and has been successfully used in many other works since then.

By the Principle of Inclusion-Exclusion, we have that the number of paths with no stationary points is

$$g(p,n) \;=\; \sum_{k=0}^{n-1}(-1)^k\binom{n}{k}r(p,n-k), \tag{3.18}$$

where each term in the sum represents the number of paths from p to the origin containing at least k stationary points, for $0 \le k \le n-1$. Now, the following identity and its proof serve as motivation for how to proceed in evaluating (3.18).

Lemma 3.2. *For $k,m,n \in \mathbb{N}$,*

$$\sum_{i=0}^{n}\binom{n}{i}\binom{m+n-i}{k-i}(-1)^i \;=\; \binom{m}{k}. \tag{3.19}$$

Recall that $[n]$ is shorthand for the set $\{1,2,\ldots,n\}$.

Proof. We view the inner term as counting the number of ordered pairs (S,T) such that $S \subseteq [n]$, $T \subseteq [m+n] \setminus S$, and $|S|+|T| = k$. Let the set of all such valid ordered pairs be V. We consider the sign-reversing involution $f\colon V\setminus E \mapsto V\setminus E$ (where E is the set of "exceptions," for which f is ill-defined) defined by toggling the smallest term in $S \cup T$ between S and T. For example, when $k = 5$,

$$f(\{2,3,5\},\{7,8\}) \;=\; (\{3,5\},\{2,7,8\}). \tag{3.20}$$

Given this definition, we see that f is its own inverse (hence, an involution) and always "flips" the parity of $|S|$, sending ordered pairs with a positive coefficient in our sum to pairs with negative coefficients (and vice versa). Therefore, for all the ordered pairs on which f is well-defined, our desired sum is 0. Thus, we have

$$\sum_{i=0}^{n}\binom{n}{i}\binom{m+n-i}{k-i}(-1)^i \;=\; |E|. \tag{3.21}$$

Proceeding, we also see that the only pairs $(S,T) \in E$ are those where $S = \emptyset$ and $T \subseteq [m+n]\setminus[n]$. There are clearly $\binom{m}{k}$ such choices for these sets, and the desired result follows. □

Viewed properly, our formula in (3.18) looks similar to the above lemma; we now rewrite it as

$$g(p,n) \;=\; \sum_{i=0}^{n}(-1)^i\binom{n}{i}\prod_{k=1}^{d}\binom{(p_k-1)+n-i}{(n-1)-i}. \tag{3.22}$$

We also note that when $d = 1$, the above formula agrees with (3.2), as expected. Although similar methods may be applied in higher dimensions, we now consider the special case $d = 2$ because our eventual goal is a Gaussianity result in exactly two dimensions.

Theorem 3.3. *For $p,q \in \mathbb{N}$, we can represent the number of generalized jump paths of length n from (p,q) to $(0,0)$ as*

$$g((p,q),n) \;=\; \sum_{i=0}^{n-1}\binom{p-1}{i}\binom{p-1+n-i}{p}\binom{q}{n-i-1}. \tag{3.23}$$

Proof. We first assume, without loss of generality, that $p \leq q$. This is a valid assumption because there is an immediate bijection between generalized jump paths starting at (p, q) and those starting at (q, p): simply reverse the step sizes in the x and y-coordinates for each step in a given path. By letting $d = 2$ in (3.22),

$$g(p, n) \;=\; \sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{(p-1)+n-i}{(n-1)-i}\binom{(q-1)+n-i}{(n-1)-i}. \tag{3.24}$$

We now proceed in a similar manner to the proof of Lemma 3.2. We view the inner term as counting the number of ordered pairs (S, T, U) such that $S \subseteq [n]$, $T \subseteq [p+n-1] \setminus S$, $U \subseteq [q+n-1] \setminus S$, and $|S| + |T| = |S| + |U| = n-1$. Let the set of all such valid ordered triples be V. We consider the sign-reversing involution $f\colon V \setminus E \to V \setminus E$ (where E is the set of "exceptions" for which f is ill-defined) defined by toggling the smallest term in $S \cup (T \cap U)$ between all three sets. Given this definition, we see that f is its own inverse (hence, an involution) and always flips the parity of $|S|$, sending ordered pairs with a positive coefficient in our sum to pairs with negative coefficients (and vice versa). The ordered pairs with a positive coefficient are precisely those where i is even; therefore, due to the factor of $(-1)^i$ in the summand (3.24), the even-indexed terms and odd-indexed terms cancel each other. Thus, for all the ordered pairs on which f is well-defined, our desired sum is 0, and we obtain

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{(p-1)+n-i}{(n-1)-i}\binom{(q-1)+n-i}{(n-1)-i} \;=\; |E|, \tag{3.25}$$

From here, our problem reduces to counting $|E|$. For f to be ill-defined, it is necessary and sufficient for $S = \emptyset$ and $T \cap U \cap [n] = \emptyset$.

To prove our desired result, we begin by indexing every tuple (S, T, U) by fixing $i = |T \cap U|$. With i fixed, we then choose the i terms that are common to T and U, which must be a subset of $[p+n-1] \setminus [n]$. Consequently, there are exactly $\binom{p-1}{i}$ ways to do this. We may now freely choose the remaining $n-i-1$ members of T (there are $\binom{p+n-i-1}{n-i-1}$ ways to do this) and the remaining $n-i-1$ members of U (for which there are $\binom{q}{n-1}$ valid configurations). Multiplying these three terms (and simplifying the second term), we find exactly the desired term, concluding the proof. □

Using this new result, we now present an alternative proof of Theorem 3.1.

Proof (Theorem 3.1). We use (2.1), combined with Theorem 3.3, to obtain

$$u((p, q), n) \;=\; \sum_{i=0}^{n}\binom{p}{i}\binom{p+n-i}{p}\binom{q}{n-i}. \tag{3.26}$$

By the definition of $F_{p,q}$ given in (3.1), we know that

$$\begin{aligned} F_{p,q}(x) \;&=\; \sum_{k=0}^{p+q}\sum_{i=0}^{k}\binom{p}{i}\binom{p+k-i}{p}\binom{q}{k-i}x^k \\ &=\; \left(\sum_{i=0}^{p}\binom{p}{i}x^i\right)\left(\sum_{k=0}^{q}\binom{q}{k}\binom{p+k}{p}x^k\right) \\ &=\; (1+x)^p\sum_{k=0}^{q}\binom{q}{k}\binom{p+k}{k}x^k, \end{aligned} \tag{3.27}$$

which proves Theorem 3.1. □

4. Proving Two-Dimensional Gaussianity

We begin this section with a little bit of notation. Let $X_{p,q}$ be a random variable that counts the length of a generalized jump path starting from the point (p, q) and ending at $(0, 0)$, when this path is chosen uniformly at random from all such paths. From Theorem 3.1, we have

$$X_{p,q} \ = \ A_{p,q} + B_{p,q}, \tag{4.1}$$

where A and B are independent random variables, proportional to a binomial coefficient and a product of binomial coefficients, respectively; in particular,

$$P(A = k) \ \propto \ \binom{p}{k}, \quad P(B = k) \ \propto \ \binom{q}{k}\binom{p+k}{k}. \tag{4.2}$$

Note that A is just the well-studied binomial random variable, which converges to a Gaussian with mean $p/2$ and variance $p/4$ as $p \to \infty$.

In this light, we restate Theorem 1.1 by using the notation above.

Theorem 4.1. *Suppose $p = m$ and $q = cm$ for $m \in \mathbb{N}^+$ and $c \geq 1$ fixed. Then, $X_{p,q}$ converges to a Gaussian as $m \to \infty$, with mean $\frac{q+p+\sqrt{p^2+6pq+q^2}}{4}$ and variance $\frac{p+q}{8} + \frac{(p+q)^2}{8\sqrt{p^2+6pq+q^2}}$.*

By our definition of $X_{p,q}$ in (4.1), it suffices to show B also converges to a Gaussian with mean $\frac{q-p+\sqrt{p^2+6pq+q^2}}{4}$ and variance $\frac{q-p}{8} + \frac{(p+q)^2}{8\sqrt{p^2+6pq+q^2}}$ for the aforementioned choices of p and q. The proof of this result will use similar techniques to those found in [11], although in our case, the algebra will be more involved because of the more complicated structure of the formulas for the number of paths.

First, we have that $P(B = k)$ is proportional to a ratio of factorials:

$$P(B = k) \ \propto \ \frac{q!}{p!}\frac{(p+k)!}{k!k!(q-k)!} \ = \ \frac{q!}{p!}\frac{(m+k)!}{k!k!(cm-k)!}. \tag{4.3}$$

Applying Stirling's formula to approximate the factorials in (4.3), we find for p, q, and k large that

$$P(B = k) \ \propto \ \frac{q!}{p!}\frac{(m+k)^{m+k+\frac{1}{2}}}{2\pi k^{2k+1}(cm-k)^{cm-k+\frac{1}{2}}}. \tag{4.4}$$

We now define

$$M \ = \ \frac{(m+k)^{m+k+\frac{1}{2}}}{k^{2k+1}(cm-k)^{cm-k+\frac{1}{2}}} \tag{4.5}$$

(i.e., we remove the terms in (4.4) that are constant with respect to k). Taking the logarithm of both sides of (4.5), we find

$$\log M \ = \ \left(m + k + \frac{1}{2}\right)\log(m+k) - \left(2k+1\right)\log(k) - \left(cm - k + \frac{1}{2}\right)\log(cm-k). \tag{4.6}$$

Now, we write k as $am + t\sqrt{m}$, where $a = \frac{c-1+\sqrt{c^2+6c+1}}{4}$. The factor of a is such that M is maximized near the point $k = an$, where this value of k is the positive solution to $\frac{(m+k)(cm-k)}{k^2} = 1$. This decomposition allows us to introduce a more natural variable for the number of steps in a path, where this number is written in terms of its distance (in natural units) from the mean. By the Taylor Expansion of $\log(1+x)$, we have

$$\log(u+x) \;=\; \log(u)+\log\left(1+\frac{x}{u}\right) \;=\; \log(u)+\frac{x}{u}-\frac{1}{2}\left(\frac{x}{u}\right)^2+O\left(\frac{x^3}{u^3}\right). \tag{4.7}$$

We do this expansion because later, in Lemma 4.2, we show that $k = am$ is the center of the distribution, and almost all (i.e., with probability 1 in the limit) the mass of the distribution is located where t is small. We compute

$$\begin{aligned}
\log M \;&=\; \Big(m+k+\frac{1}{2}\Big)\log\Big(m\Big(1+a+\frac{t}{\sqrt{m}}\Big)\Big)\\
&\quad -\Big(2k+1\Big)\log\Big(m\Big(a+\frac{t}{\sqrt{m}}\Big)\Big)\\
&\quad -\Big(cm-k+\frac{1}{2}\Big)\log\Big(m\Big(c-a-\frac{t}{\sqrt{m}}\Big)\Big)\\
&= (m-cm-1)\log(m)\\
&\quad +\Big(m+k+\frac{1}{2}\Big)\log\Big(1+a+\frac{t}{\sqrt{m}}\Big)\\
&\quad -\Big(2k+1\Big)\log\Big(a+\frac{t}{\sqrt{m}}\Big)\\
&\quad -\Big(cm-k+\frac{1}{2}\Big)\log\Big(c-a-\frac{t}{\sqrt{m}}\Big)\\
&= (m-cm-1)\log(m)\\
&\quad +\Big(m+k+\frac{1}{2}\Big)\Big(\log(1+a)+\frac{t}{(1+a)\sqrt{m}}-\frac{t^2}{2(1+a)^2m}+O\Big(\frac{t^3}{m^{3/2}}\Big)\Big)\\
&\quad -\Big(2k+1\Big)\Big(\log(a)+\frac{t}{a\sqrt{m}}-\frac{t^2}{2a^2m}+O\left(\frac{t^3}{m^{3/2}}\right)\Big)\\
&\quad -\left(cm-k+\frac{1}{2}\right)\left(\log(c-a)-\frac{t}{(c-a)\sqrt{m}}-\frac{t^2}{2(c-a)^2m}+O\left(\frac{t^3}{m^{3/2}}\right)\right).
\end{aligned} \tag{4.8}$$

After standard but tedious computations, the details of which are in Appendix A, we obtain

$$\log M \;=\; -\chi t^2+f(m)+O\left(\frac{t^3}{\sqrt{m}}\right), \tag{4.9}$$

where

$$\chi \;=\; \frac{2c^2+10c^3-10c^4-2c^5+2c^2\sqrt{1+c(6+c)}+4c^3\sqrt{1+c(6+c)}+2c^4\sqrt{1+c(6+c)}}{8c^4} \tag{4.10}$$

and $f(m)$ is independent of k and t. We also note that, because $f(m)$ is independent of k and t, it will necessarily be canceled when M is normalized to become a probability distribution, so we may disregard it in our future analysis.

We now prove the following lemma, which demonstrates that the tails of this distribution decay sufficiently quickly.

Lemma 4.2. *As* $m \to \infty$, $P(|t| > m^{0.1}) \to 0$.

Proof. We will demonstrate that $P(t > m^{0.1})$ goes to 0 as $n \to \infty$; the proof for $P(t < -m^{0.1})$ proceeds similarly.

Note that if we increase k by 1, then the number of paths of length k increases by a factor of approximately $\frac{(p+k)(q-k)}{k^2}$ (which is strictly decreasing in k). Plugging in $k = am + t\sqrt{m}$, we see that the proportion at which the number of paths increases is

$$1 - \left(\frac{8\sqrt{c^2+6c+1}}{(1+c)^2 + (c-1)\sqrt{c^2+6c+1}} \frac{t}{\sqrt{m}} \right) + O\left(\frac{t^2}{m}\right). \tag{4.11}$$

Let the coefficient of the $t/\sqrt{m}$ term be c'. Plugging in $t = 1$, and ignoring the higher order terms, for $k \geq an + \sqrt{n}$, this ratio will be at most $r = 1 - c'\sqrt{n}$. Let B_0 be the number of paths of length k when $t = 1$. The number of paths with $t > m^{0.1}$ is bounded above by a geometric series with first term $B_0 \cdot r^{m^{0.6} - m^{0.5}}$ and ratio r, so we have

$$r^{m^{0.5}} \;=\; \left(1 - \frac{c'}{\sqrt{m}}\right)^{\sqrt{m}} \approx e^{-c'} < 1. \tag{4.12}$$

Thus, as m goes to infinity, $B_0 \cdot r^{m^{0.6}-m^{0.5}} \;=\; B_0 \cdot \left(e^{-c'}\right)^{m^{0.1}-1}$, where the last exponent, $m^{0.1} - 1$, goes to infinity. Now note that the total number of paths of any length is strictly greater than B_0, so the probability that $t > m^{0.1}$ is at most

$$\frac{B_0 \cdot \left(e^{-c'}\right)^{m^{0.1}-1} / (1-r)}{B_0} \;=\; \frac{\sqrt{m}}{c'} \left(e^{-c'}\right)^{m^{0.1}-1}, \tag{4.13}$$

which goes to 0 as m grows. Thus, $P(t > m^{0.1})$ goes to 0 as $m \to \infty$, as desired. □

Consequently, we know that $t^3/\sqrt{m} \to 0$ as $m \to \infty$, so by sending p and q to ∞,

$$P(B = k) \;=\; Ce^{-\chi t^2} \;=\; Ce^{-\frac{(am-k)^2}{m/\chi}}, \tag{4.14}$$

where C is a real-valued constant. Furthermore, (4.14) is the equation for a Gaussian distribution with mean

$$am \;=\; \frac{q - p + \sqrt{p^2 + 6pq + q^2}}{4}, \tag{4.15}$$

and variance

$$\frac{m}{2\chi} \;=\; m\left(\frac{c-1}{8} + \frac{c^2+2c+1}{8\sqrt{1+6c+c^2}}\right) \;=\; \frac{q-p}{8} + \frac{q^2+2pq+p^2}{8\sqrt{p^2+6pq+q^2}}. \tag{4.16}$$

We conclude that $X_{p,q}$ is a Gaussian random variable with mean

$$E[X_{p,q}] \;=\; E[A_{p,q}] + E[B_{p,q}] \;=\; \frac{p}{2} + \frac{q-p+\sqrt{p^2+6pq+q^2}}{4} \;=\; \frac{p+q}{4} + \frac{\sqrt{p^2+6pq+q^2}}{4}, \tag{4.17}$$

and variance

$$\begin{aligned} \mathrm{Var}\left(X_{p,q}\right) \;&=\; \mathrm{Var}\left(A_{p,q}\right) + \mathrm{Var}\left(B_{p,q}\right) \\ &=\; \frac{p}{4} + \frac{q-p}{8} + \frac{q^2+2pq+p^2}{8\sqrt{p^2+6pq+q^2}} \;=\; \frac{p+q}{8} + \frac{(p+q)^2}{8\sqrt{p^2+6pq+q^2}}. \end{aligned} \tag{4.18}$$

This completes the proof of Theorem 1.1. □

5. Further Work and Open Questions

We conclude with some suggestions of future continuations of the work done in this paper, which are predominantly combinatorial in nature.

(1) Is it possible to generalize the generating function approach for points not on the diagonal? Although empirical results suggest similar behavior, the associated Taylor expansion becomes significantly harder to work with.
(2) How do our results generalize to higher dimensions? Our combinatorial approach admits a natural extension to higher dimensions, although the casework becomes significantly more cumbersome. Furthermore, note that the generating function approach cannot be used in three or more dimensions.
(3) How quickly does our distribution converge to a Gaussian?
(4) What is the behavior if we allow a point to be used more than once? If we can use a point arbitrarily many times it is unclear how to define the terms of our sequence; thus, the natural question would be, what happens if each point can be used at most T times, for some fixed T.
(5) There are even more combinatorial questions worth exploring. Many additional restrictions can be put on the generalized jump paths. One such restriction is prohibiting any path from visiting any points that lie above the line $y = x$.

Appendix A. Gaussianity Calculations

We give the simplification of $\log M$ from (4.5). We have that

$$\begin{aligned}\log M \;=\; & -\frac{t}{a\sqrt{m}} + \frac{t}{2(1+a)\sqrt{m}} + \frac{t}{2(-a+c)\sqrt{m}} - \frac{2kt}{a\sqrt{m}} + \frac{kt}{(1+a)\sqrt{m}} \\ & - \frac{kt}{(-a+c)\sqrt{m}} + \frac{\sqrt{m}t}{1+a} + \frac{c\sqrt{m}t}{-a+c} - \frac{t^2}{2(1+a)^2} + \frac{ct^2}{2(-a+c)^2} \\ & + \frac{t^2}{2a^2m} - \frac{t^2}{4(1+a)^2m} + \frac{t^2}{4(-a+c)^2m} + \frac{kt^2}{a^2m} - \frac{kt^2}{2(1+a)^2m} \\ & - \frac{kt^2}{2(-a+c)^2m} - \log(a) - 2k\log(a) + \frac{1}{2}\log(1+a) + k\log(1+a) \\ & + m\log(1+a) - \frac{1}{2}\log(-a+c) + k\log(-a+c) - cm\log(-a+c). \qquad \text{(A.1)}\end{aligned}$$

Substituting $a = \frac{c-1+\sqrt{c^2+6c+1}}{4}$, we find that this expression, for $\log M$, is equal to

$$
\begin{aligned}
&-\frac{4t}{\left(-1+c+\sqrt{1+6c+c^2}\right)\sqrt{m}}+\frac{t}{2\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)\sqrt{m}}\\
&+\frac{t}{2\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)\sqrt{m}}+\frac{c\sqrt{m}t}{c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)}\\
&+\frac{\sqrt{m}t}{1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)}+\frac{ct^2}{2\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)^2}\\
&-\frac{t^2}{2\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)^2}+\frac{8t^2}{\left(-1+c+\sqrt{1+6c+c^2}\right)^2 m}\\
&+\frac{t^2}{4\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)^2 m}-\frac{t^2}{4\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)^2 m}\\
&-\frac{8t\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)}{\left(-1+c+\sqrt{1+6c+c^2}\right)\sqrt{m}}-\frac{t\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)}{\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)\sqrt{m}}\\
&+\frac{t\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)}{\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)\sqrt{m}}+\frac{16t^2\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)}{\left(-1+c+\sqrt{1+6c+c^2}\right)^2 m}\\
&-\frac{t^2\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)}{2\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)^2 m}-\frac{t^2\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)}{2\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)^2 m}\\
&-\log\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)\\
&-2\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)\log\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)\\
&-\frac{1}{2}\log\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)-cm\log\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)\\
&+\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)\log\left(c+\frac{1}{4}\left(1-c-\sqrt{1+6c+c^2}\right)\right)\\
&+\frac{1}{2}\log\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)+m\log\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right)\\
&+\left(\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)m+\sqrt{m}t\right)\log\left(1+\frac{1}{4}\left(-1+c+\sqrt{1+6c+c^2}\right)\right). \qquad \text{(A.2)}
\end{aligned}
$$

Simplifying and collecting like terms, we have that $\log M$ equals

$$\begin{aligned}
&\frac{1}{8c^4 n}\Big(t^2+6ct^2+7c^2t^2+4c^3t^2-3c^4t^2-6c^5t^2-c^6t^2+\sqrt{1+c(6+c)}t^2\\
&\quad+3c\sqrt{1+c(6+c)}t^2+2c^2\sqrt{1+c(6+c)}t^2\\
&\quad-2c^3\sqrt{1+c(6+c)}t^2+3c^4\sqrt{1+c(6+c)}t^2+c^5\sqrt{1+c(6+c)}t^2\Big)\\
&+\frac{1}{8c^4\sqrt{m}}\Big(2c^2t+2c^3t+10c^4t+2c^5t+2c^2\sqrt{1+c(6+c)}t-4c^3\sqrt{1+c(6+c)}t\\
&\quad-2c^4\sqrt{1+c(6+c)}t-2t^3-12ct^3-6c^2t^3+8c^3t^3-6c^4t^3-12c^5t^3\\
&\quad-2c^6t^3-2\sqrt{1+c(6+c)}t^3-6c\sqrt{1+c(6+c)}t^3+4c^2\sqrt{1+c(6+c)}t^3\\
&\quad-4c^3\sqrt{1+c(6+c)}t^3+6c^4\sqrt{1+c(6+c)}t^3+2c^5\sqrt{1+c(6+c)}t^3\Big)\\
&+\frac{1}{8c^4}\Big(-2c^2t^2-10c^3t^2+10c^4t^2+2c^5t^2-2c^2\sqrt{1+c(6+c)}t^2-4c^3\sqrt{1+c(6+c)}t^2\\
&\quad-2c^4\sqrt{1+c(6+c)}t^2+8c^4\log(8)-4c^4\log\Big(4+12c-4\sqrt{1+c(6+c)}\Big)\\
&\quad-8c^4\log\Big(-1+c+\sqrt{1+c(6+c)}\Big)+4c^4\log\Big(3+c+\sqrt{1+c(6+c)}\Big)\Big)\\
&+\frac{\sqrt{m}}{8c^4}\Big(8c^4t\log\Big(1+3c-\sqrt{1+c(6+c)}\Big)-16c^4t\log\Big(-1+c+\sqrt{1+c(6+c)}\Big)\\
&\quad+8c^4t\log\Big(3+c+\sqrt{1+c(6+c)}\Big)\Big)\\
&+\frac{m}{8c^4}\Big(-8c^4\log(4)+8c^5\log(4)-2c^4\sqrt{1+c(6+c)}\log(4)-2c^4\log\Big(1+3c-\sqrt{1+c(6+c)}\Big)\\
&\quad-6c^5\log\Big(1+3c-\sqrt{1+c(6+c)}\Big)+2c^4\sqrt{1+c(6+c)}\log\Big(1+3c-\sqrt{1+c(6+c)}\Big)\\
&\quad+4c^4\log\Big(-1+c+\sqrt{1+c(6+c)}\Big)-4c^5\log\Big(-1+c+\sqrt{1+c(6+c)}\Big)\\
&\quad-4c^4\sqrt{1+c(6+c)}\log\Big(-1+c+\sqrt{1+c(6+c)}\Big)+6c^4\log\Big(3+c+\sqrt{1+c(6+c)}\Big)\\
&\quad+2c^5\log\Big(3+c+\sqrt{1+c(6+c)}\Big)+2c^4\sqrt{1+c(6+c)}\log\Big(4\Big(3+c+\sqrt{1+c(6+c)}\Big)\Big)\Big).
\end{aligned} \tag{A.3}$$

We now verify that the $t\sqrt{m}$ term has coefficient 0. Our current coefficient is

$$\begin{aligned}
&\log\Big(1+3c-\sqrt{1+c(6+c)}\Big)-2\log\Big(-1+c+\sqrt{1+c(6+c)}\Big)\\
&\quad+\log\Big(3+c+\sqrt{1+c(6+c)}\Big).
\end{aligned} \tag{A.4}$$

Exponentiating this expression, we obtain

$$\frac{\Big(1+3c-\sqrt{1+c(6+c)}\Big)\Big(3+c+\sqrt{1+c(6+c)}\Big)}{\Big(-1+c+\sqrt{1+c(6+c)}\Big)^2} \;=\; 1 \tag{A.5}$$

as desired, completing the computations.

Consequently, as claimed, we have that $\log M$ is of the form

$$\log M \;=\; -\chi t^2+f(m)+O\left(\frac{t^3}{\sqrt{n}}\right), \tag{A.6}$$

where χ is defined as

$$\chi \;=\; \frac{2c^2+10c^3-10c^4-2c^5+(2c^2+4c^3+2c^4)\sqrt{1+c(6+c)}}{8c^4}, \tag{A.7}$$

and $f(m)$ is independent of k or t with $f(m) = O\left(\frac{1}{\sqrt{m}}\right)$.

References

[1] H. Alpert, *Differences of multiple Fibonacci numbers*, Integers: Electronic Journal of Combinatorial Number Theory, **9** (2009), 745–749.

[2] O. Beckwith, A. Bower, L. Gaudet, R. Insoft, S. Li, S. J. Miller, and P. Tosteson, *The average gap distribution for generalized Zeckendorf decompositions*, The Fibonacci Quarterly, **51.1** (2013), 13–27.

[3] I. Ben-Ari and S. Miller, *A probabilistic approach to generalized Zeckendorf decompositions*, SIAM Journal on Discrete Mathematics, **30.2** (2016), 1302–1332.

[4] 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.

[5] N. Borade, D. Cai, D. Z. Chang, B. Fang, A. Liang, S. J. Miller, and W. Xu, *Gaps of summands of the Zeckendorf lattice*, to appear in The Fibonacci Quarterly.

[6] A. Bower, R. Insoft, S. Li, S. J. Miller, and P. Tosteson, *The distribution of gaps between summands in generalized Zeckendorf decompositions* (and an appendix on *Extensions to initial segments* with Iddo Ben-Ari), Journal of Combinatorial Theory, Series A, **135** (2015), 130–160.

[7] J. L. Brown, Jr., *Zeckendorf's theorem and some applications*, The Fibonacci Quarterly, **2.3** (1964), 163–168.

[8] M. Catral, P. Ford, P. E. Harris, S. J. Miller, and D. Nelson, *Generalizing Zeckendorf's theorem: The Kentucky sequence*, The Fibonacci Quarterly, **52.5** (2014), 68–90.

[9] M. Catral, P. Ford, P. E. Harris, S. J. Miller, and D. Nelson, *Legal decompositions arising from non-positive linear recurrences*, preprint.

[10] 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.

[11] E. Chen, R. Chen, L. Guo, C. Jiang, S. J. Miller, J. M. Siktar, and P. Yu, *Gaussian behavior in Zeckendorf decompositions from lattices*, The Fibonacci Quarterly, **57.3** (2019), 201–212.

[12] D. E. Daykin, *Representation of natural numbers as sums of generalized Fibonacci numbers*, J. London Mathematical Society, **35** (1960), 143–160.

[13] P. Demontigny, T. Do, A. Kulkarni, S. J. Miller, D. Moon, and U. Varma, *Generalizing Zeckendorf's Theorem to f-decompositions*, Journal of Number Theory, **141** (2014), 136–158.

[14] P. Demontigny, T. Do, A. Kulkarni, S. J. Miller, and U. Varma, *A generalization of Fibonacci far-difference representations and Gaussian behavior*, The Fibonacci Quarterly, **52.3** (2014), 247–273.

[15] R. Dorward, P. Ford, E. Fourakis, P. E. Harris, S. J. Miller, E. Palsson, and H. Paugh, *Individual gap measures from generalized Zeckendorf decompositions*, Uniform Distribution Theory, **12.1** (2017), 27–36, http://arxiv.org/pdf/1509.03029v1.pdf

[16] R. Dorward, P. Ford, E. Fourakis, P. E. Harris, S. J. 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.

[17] M. Drmota and J. Gajdosik, *The distribution of the sum-of-digits function*, J. Théor. Nombrés Bordeaux, **10.1** (1998), 17–32.

[18] S. Eger, *Stirling's approximation for central extended binomial coefficients*, American Mathematical Monthly, **121.4** (2014), 344–349, http://arxiv.org/pdf/1203.2122.pdf.

[19] P. Filipponi, P. J. Grabner, I. Nemes, A. Pethö, and R. F. Tichy, *Corrigendum to: "Generalized Zeckendorf expansions"*, Appl. Math. Lett., **7.6** (1994), 25–26.

[20] A. S. Fraenkel, *Systems of numeration*, American Mathematical Monthly, **92.2** (1985), 105–114.

[21] P. J. Grabner, R. F. Tichy, I. Nemes, and A. Pethö, *Generalized Zeckendorf expansions*, Appl. Math. Lett., **7.2** (1994), 25–28.

[22] N. Hamlin, *Representing positive Integers as a sum of linear recurrence sequences*, Abstracts of Talks, Fourteenth International Conference on Fibonacci Numbers and Their Applications, (2010), 2–3.

[23] N. Hamlin and W. A. Webb, *Representing positive integers as a sum of linear recurrence sequences*, The Fibonacci Quarterly **50.2** (2012), 99–105.
[24] V. E. Hoggatt, *Generalized Zeckendorf theorem*, The Fibonacci Quarterly **10.1** (1972), (special issue on representations), 89–93.
[25] T. J. Keller, *Generalizations of Zeckendorf's theorem*, The Fibonacci Quarterly, **10.1** (1972), (special issue on representations), 95–102.
[26] W. Koepf, *Hypergeometric Summation: An Algorithmic Approach to Summation and Special Function Identities*, Braunschweig, Germany: Vieweg, 1998.
[27] M. Koloğlu, G. Kopp, S. J. Miller, and Y. Wang, *On the number of summands in Zeckendorf decompositions*, The Fibonacci Quarterly **49.2** (2011), 116–130.
[28] M. Lamberger and J. M. Thuswaldner, *Distribution properties of digital expansions arising from linear recurrences*, Math. Slovaca, **53.1** (2003), 1–20.
[29] C. G. Lekkerkerker, *Voorstelling van natuurlyke getallen door een som van getallen van Fibonacci*, Simon Stevin, **29** (1951–1952), 190–195.
[30] T. Lengyel, *A counting based proof of the generalized Zeckendorf's theorem*, The Fibonacci Quarterly **44.4** (2006), 324–325.
[31] R. Li and S. J. Miller, *Central limit theorems for gaps of generalized Zeckendorf decompositions*, The Fibonacci Quarterly, **57.3** (2019), 213–230, http://arxiv.org/abs/1606.08110v1.pdf
[32] R. Li and S. J. Miller, *A collection of central limit type results in generalized Zeckendorf decompositions*, The Fibonacci Quarterly, **55.5** (2017), 105–114.
[33] S. J. Miller, *The Probability Lifesaver*, Princeton University Press, 2017, 752 pages.
[34] S. J. Miller and Y. Wang, *From Fibonacci numbers to central limit type theorems*, Journal of Combinatorial Theory, Series A, **119.7** (2012), 1398–1413.
[35] S. J. 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.
[36] A. Pethö and R. F. Tichy, *On digit expansions with respect to linear recurrences*, J. Number Theory, **33** (1989), 243–256.
[37] R. P. Stanley, *Enumerative Combinatorics* (Vol. 2), Cambridge studies in advanced mathematics, 2001.
[38] W. Steiner, *The joint distribution of greedy and lazy Fibonacci expansions*, The Fibonacci Quarterly **43.1** (2005), 60–69.
[39] W. Steiner, *Parry expansions of polynomial sequences*, Integers, **2** (2002), Paper A14.
[40] 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 Liége, **41** (1972), 179–182.

MSC2010: 11B01, 05A02, 60C02

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
Email address: evanfang@andrew.cmu.edu

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
Email address: jtjenkin@andrew.cmu.edu

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
Email address: ztlee@andrew.cmu.edu

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
Email address: drli@andrew.cmu.edu

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
Email address: ethanlu@andrew.cmu.edu

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, and Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267
Email address: sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213
Email address: dsalgado@andrew.cmu.edu

Department of Mathematics, University of Tennessee-Knoxville, Knoxville, TN 37916
Email address: jsiktar@vols.utk.edu