AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **77(2)** (2020), Pages 184–223

On the number of reachable pairs in a digraph

Eric Swartz

Department of Mathematics
College of William & Mary
Williamsburg, VA
U.S.A.
easwartz@wm.edu

Nicholas J. Werner

Department of Mathematics, Computer and Information Science
State University of New York College at Old Westbury
Old Westbury, NY
U.S.A.
wernern@oldwestbury.edu

Abstract

A pair (u, v) of (not necessarily distinct) vertices in a directed graph D is called a reachable pair if there exists a directed path from u to v. We define the weight of D to be the number of reachable pairs of D, which equals the sum of the number of vertices in D and the number of directed edges in the transitive closure of D. In this paper, we study the set $W(n)$ of possible weights of directed graphs on n labeled vertices. We prove that $W(n)$ can be determined recursively and describe the integers in the set. Moreover, if $b(n) \geqslant n$ is the least integer for which there is no digraph on n vertices with exactly $b(n) + 1$ reachable pairs, we determine $b(n)$ exactly through a simple recursive formula and find an explicit function $g(n)$ such that $|b(n) - g(n)| < 2n$ for all $n \geqslant 3$. Using these results, we are able to approximate $|W(n)|$ — which is quadratic in n — with an explicit function that is within $30n$ of $|W(n)|$ for all $n \geqslant 3$, thus answering a question of Rao. Since the weight of a directed graph on n vertices corresponds to the number of elements in a preorder on an n element set and the number of containments among the minimal open sets of a topology on an n point space, our theorems are applicable to preorders and topologies.

ISSN: 2202-3518

1 Introduction

Let D be a directed graph (digraph) with vertex set V and directed edge set E. We will assume that digraphs do not contain loops or parallel edges (although these assumptions are irrelevant for this problem), but we do allow for pairs of oppositely oriented directed edges, e.g., given vertices x and y, we allow both (x, y) and (y, x) to be in the directed edge set. A *reachable pair* is an ordered pair (u, v) of vertices such that, for some nonnegative integer k, there exists a sequence of vertices $x_0 = u, x_1, \ldots, x_k = v$ with the property that there is a directed edge from x_i to x_{i+1} for each i, $0 \leqslant i \leqslant k-1$. In other words, (u, v) is a reachable pair if there exists a directed path from u to v. We allow $k = 0$ for this, and so (u, u) is considered a reachable pair for each vertex u.

The determination of reachability in digraphs has been the object of considerable study in algorithmic design. It is readily seen that the problem of determining the number of reachable pairs in a given digraph D is equivalent to finding the size of the *transitive closure* $\overline{D}$ of D: $\overline{D}$ has the same vertex set as D, but $\overline{D}$ contains the directed edge (u, v) if and only if there is a directed path from u to v in D. This problem has clear implications to communication within a network and is also important for many database problems, such as database query optimization. For a discussion of this problem, see [9, Section 15.5].

Figure 1: A digraph D (left) with its transitive closure $\overline{D}$ (right)

A natural question along these lines is the following: given a positive integer n, what are the possible numbers of reachable pairs in a digraph on n vertices? Obviously, by our definition, there must always be at least n reachable pairs, since (u, u) is a reachable pair for each vertex u; and there are at most n^2 total pairs, so the number of reachable pairs is at most n^2, which occurs in a complete directed graph.

Definition 1.1. We define the *weight* of a digraph D, denoted $w(D)$, to be the number of reachable pairs in D, which is equal to the sum of the number of vertices of D and the number of directed edges in the transitive closure of D. For each $n \in \mathbb{N}$, we define

$$W(n) = \{k \in \mathbb{N} : \text{ there exists a digraph } D \text{ on } n \text{ vertices of weight } k\}$$

to be the set of all possible weights of an n vertex digraph. We call $W(n)$ a *weight set*.

n	$W(n)$	$[n, n^2] \setminus W(n)$
1	1	$\varnothing$
2	$[2, 4]$	$\varnothing$
3	$[3, 7]$, 9	8
4	$[4, 13]$, 16	14, 15
5	$[5, 19]$, 21, 25	20, $[22, 24]$
6	$[6, 28]$, 31, 36	29, 30, $[32, 35]$
7	$[7, 35]$, $[37, 39]$, 43, 49	36, $[40, 42]$, $[44, 48]$
8	$[8, 52]$, 57, 64	$[53, 56]$, $[58, 63]$
9	$[9, 61]$, 63, $[65, 67]$, 73, 81	62, 64, $[68, 72]$, $[74, 80]$
10	$[10, 77]$, 79, $[82, 84]$, 91, 100	78, 80, 81, $[85, 90]$, $[92, 99]$
11	$[11, 95]$, 97, $[101, 103]$, 111, 121	96, $[98, 100]$, $[104, 110]$, $[112, 120]$
12	$[12, 109]$, $[111, 115]$, 117, $[122, 124]$, 133, 144	110, 116, $[118, 121]$, $[125, 132]$, $[134, 143]$

Table 1: Possible numbers of reachable pairs in a digraph with n vertices.

For example, the digraphs shown in Figure 1 each have exactly 15 reachable pairs, and so $w(D) = w(\overline{D}) = 15$.

Our terminology here is inspired by the standard vocabulary used for weighted graphs. If each directed edge of $\overline{D}$ is assigned a weight of 1, then the weight of D is simply n plus the total weight of the directed graph $\overline{D}$. Usually, we will assume that D itself is transitive, i.e., that $D = \overline{D}$.

Weight sets for directed graphs can be easily determined for small values of n. For example, we have $W(2) = \{2, 3, 4\}$, since, if $V = \{u, v\}$, we may choose E to be $\varnothing$, $\{(u, v)\}$, or $\{(u, v), (v, u)\}$. So, when $n = 2$, all values between n and n^2 occur as possible weights. On the other hand, when $n = 3$, it is impossible for there to be exactly eight reachable pairs: if $V = \{u, v, x\}$ and (u, v) is the unique pair that is not reachable, then both (u, x) and (x, v) are reachable pairs, meaning (u, v) is also reachable by transitivity, a contradiction. It is not difficult to see that $W(3) = [3, 7] \cup \{9\}$, where $[a, b]$ denotes the set of integers k such that $a \leqslant k \leqslant b$. Table 1 lists the values in $W(n)$ for $n \leqslant 12$, which may be determined through brute force calculations.

From Table 1, one can see that $W(n)$ becomes more fragmented as n increases. Nevertheless, there are intriguing patterns in this data. For instance, $W(n)$ never includes integers in the range $[n^2 - n + 2, n^2 - 1]$, and for $n \geqslant 5$, $W(n)$ does not meet $[n^2 - 2n + 5, n^2 - n]$ (see [6, Corollary 2]). Moreover, $W(n)$ always begins with a single interval that contains the majority of the elements of the set, which motivates the following definition.

Definition 1.2. For each $n \in \mathbb{N}$, we define $b(n)$ to be the least integer such that

$b(n) \geqslant n$ and there does not exist a digraph on n vertices with exactly $b(n)+1$ reachable pairs. Equivalently, $b(n)$ is the largest positive integer such that $[n, b(n)] \subseteq W(n)$.

The set $W(n)$ and the integer $b(n)$ were studied previously in [6], although there the reachable pairs did not include pairs of the form (u, u). Hence, the set $S(n)$ studied in [6] is related to $W(n)$ by

$$S(n) = \{k - n : k \in W(n)\},$$

and the function $f(n)$ studied in [6] is related to $b(n)$ by $f(n) = b(n) - n$. Indeed, [6, Theorem 6], when translated into our notation, gives a lower bound of

$$b(n) \geqslant n^2 - n \cdot \lfloor n^{0.57} \rfloor + \lfloor n^{0.57} \rfloor.$$

However, this bound is not asymptotically tight, as it is noted in [6] that the lower bound holds for large enough n if the exponent 0.57 is replaced by 0.53. Techniques for determining the set $S(n)$ for $n \leqslant 208$ are given in [6], although an efficient method for calculating this set in general or even estimating its size is left as an open problem.

The purpose of this paper is to study the set $W(n)$ and the function $b(n)$. First, we establish methods to determine $W(n)$ exactly. It follows from Rao's result [6, Theorem 6], that $b(n) \geqslant (3/4)n^2$ for all $n \geqslant 8$, and hence $[n, \lceil (3/4)n^2 \rceil] \subseteq W(n)$. We then prove that the larger values in $W(n)$ can realized by transitive digraphs with a particularly nice form. A *mother vertex* of D is a vertex u such that, for all vertices $v \neq u$, there is a directed edge from u to v.

Theorem 1.3. *Let D be a transitive digraph on n vertices with vertex set V and directed edge set E. If $w(D) > (3/4)n^2$, then there exists a transitive digraph D' on n vertices such that $w(D) = w(D')$ and D' has at least one mother vertex.*

Theorem 1.3 allows us to compute $W(n)$ recursively (see Corollary 2.6). While this is an interesting result, the real power of Theorem 1.3 becomes apparent when studying the function $b(n)$. We will prove that $b(n)$ can be determined exactly—and independently of $W(n)$—via its own recursive formula.

Theorem 1.4. *Define $\ell(z) := b(z) - z + 3$. Let $z \geqslant 1$ and let n be such that $\ell(z) \leqslant n < \ell(z+1)$. If $n \neq 8$, then $b(n) = n^2 - zn + b(z)$.*

While the recursive formula is quite effective in practice, for large values of n, it requires knowledge of the values of $b(m)$ for $m < n$. It would be beneficial to have a good estimate for $b(n)$ based only on n. Evidently from Theorem 1.4, this requires an accurate estimate for the integer z such that $\ell(z) \leqslant n < \ell(z+1)$. We are able to provide such an approximation for z (see Definitions 4.1, 4.5 and Theorem 4.6), which in turn allows us to provide a very good estimate for $b(n)$.

Theorem 1.5. *Define $N := \lfloor \log_2 \log_5 n \rfloor + 1$, and define*

$$g(n) := n^2 - \left(\sum_{j=1}^{N} \frac{n^{1+\frac{1}{2^j}}}{2^{j-1}} \right) + \left(2 - \frac{1}{2^{N-1}} \right) n.$$

For all $n \geqslant 3$, $|b(n) - g(n)| < 2n$.

Finally, we are able to use the theory we have built up to obtain an estimate for $|W(n)|$ with an error that is bounded by a constant times n, which answers a question of Rao [6]. Note that $|W(n)|$ tells us the number of different integers k for which there exists a transitive digraph on n vertices with exactly k directed edges.

Theorem 1.6. *For each $n \geqslant 3$, let $N := \lfloor \log_2 \log_5 n \rfloor + 1$, and define*

$$\mathfrak{r}(n) := n^2 - \sum_{k=1}^{N} \left(\frac{2^k}{\prod_{i=1}^{k}(2^i+1)} \right) n^{1+\frac{1}{2^k}}.$$

For all $n \geqslant 3$, $||W(n)| - \mathfrak{r}(n)| < 30n$.

We remark that $\mathfrak{r}(n)$ is often far closer to $|W(n)|$ than $30n$, and we did not attempt to optimize the constant that is multiplied by n in this inequality. For example, our methods allow us to determine computationally that $|W(5000)| = 24746694$, whereas $\mathfrak{r}(5000) \approx 24752227$, a difference on the order of 5000.

This paper is organized as follows. In Section 2, we establish some terminology and basic results and prove Theorem 1.3. Section 3 is devoted to the proof of Theorem 1.4. Sections 4 and 5 focus on the constructions of the functions $g(n)$ and $\mathfrak{r}(n)$, respectively, and the proofs of Theorems 1.5 and 1.6. Finally, included in Appendix A at the end of the paper are data and Mathematica [11] code used for calculations that occur in certain proofs.

We point out that the theorems of this paper have applications in other areas. Recall that a *preorder* or *quasi-order* on a set S is a relation on S that is reflexive and transitive. When D has vertex set $[1, n]$, the set of reachable pairs of D constitutes a preorder on $[1, n]$. Conversely, given a preorder $\mathcal{U}$ on $[1, n]$, the directed graph D with vertex set $[1, n]$ and directed edge set $\{(i, j) \in \mathcal{U} : i \neq j\}$ is transitive. Thus, the set of reachable pairs of D equals $\mathcal{U}$, and finding $W(n)$ is equivalent to determining the possible sizes of a preorder on an n-element set.

Preorders also correspond to topologies on finite sets. We sketch this relationship here; details can be found in most references on finite topologies such as [1, 4, 5, 7, 8, 10]. Given a preorder $\mathcal{U}$ on $[1, n]$, for each $i \in [1, n]$ let $X_i = \{j \in [1, n] : (i, j) \in \mathcal{U}\}$. One may then construct the topology $\mathcal{T}$ on $[1, n]$ that has $\{X_i\}_{1 \leqslant i \leqslant n}$ as its open basis. Conversely, beginning with a topology $\mathcal{T}$ on $[1, n]$, we can recover the corresponding preorder $\mathcal{U}$. To do this, for each $i \in [1, n]$ we let U_i be the minimal open set of $\mathcal{T}$ containing i. Then, $\mathcal{U}$ is defined by the rule $(i, j) \in \mathcal{U}$ if and only if $U_i \supseteq U_j$. In the topological formulation, the weight corresponds to $\sum_{i=1}^{n} |X_i| = \sum_{i=1}^{n} |U_i|$ (or, equivalently, the number of containments $U_i \supseteq U_j$), which is an invariant of $\mathcal{T}$. In this way, knowledge of weights and $W(n)$ provides information on these topological spaces and their corresponding preorders.

2 Elements of the weight set

We begin this section with a discussion of basic terminology and ideas that will prove useful. Let D be a transitive digraph with vertex set V and directed edge set E.

A *clique* or *complete directed subgraph* of D is a subset $A \subseteq V$ such that, for all $u \neq v \in A$, there is a directed edge from u to v in D. A *mother vertex* is a vertex u such that there is a path of directed edges from u to any other vertex in the digraph. In a transitive digraph, a mother vertex is a vertex u such that there is a directed edge from u to every other vertex. Given a subset $A \subseteq V$, the *induced subgraph* $D[A]$ has vertex set A and directed edge set $E[A] := \{(u_1, u_2) \in E : u_1, u_2 \in A\}$. It follows immediately from transitivity that, if D is a transitive digraph with vertex set V, and $A \cup B$ is a partition of V, then: both $D[A]$ and $D[B]$ are transitive digraphs; the digraph with vertex set V and directed edge set $E[A] \cup E[B]$ is transitive; and, if $E' := \{(u, v) : u \in A, v \in B\}$, the digraph with vertex set V and directed edge set $E[A] \cup E[B] \cup E'$ is transitive.

We first demonstrate a simple recursive method to produce subsets of $W(n)$.

Lemma 2.1. *For each $n \geqslant 1$, define the following sets $\mathcal{B}_n$ and $\mathcal{C}_n$:*

$$\mathcal{B}_n := \{n(n-k) + d : d \in W(k), 1 \leqslant k \leqslant n-1\}$$
$$\mathcal{C}_n := \{c + d : c \in W(n-k), d \in W(k), 1 \leqslant k \leqslant n-1\}.$$

Then, $W(n) \supseteq \mathcal{B}_n \cup \mathcal{C}_n \cup \{n^2\}$.

Proof. Certainly, $n^2 \in W(n)$. The weights in $\mathcal{C}_n$ correspond to a partition $V = A \cup B$ such that $|A| = n - k$; $|B| = k$; and the edge set E of D is the disjoint union of $E[A]$ and $E[B]$ (that is, there are no edges between A and B in either direction). Then, $w(D[A]) \in W(n-k)$, $w(D[B]) \in W(k)$, and $w(D) = w(D[A]) + w(D[B])$. The weights in $\mathcal{B}_n$ correspond to a partition $V = A \cup B$ such that $|A| = n - k$; $|B| = k$; each vertex in A is a mother vertex; no vertex in B is a mother vertex; and there are no edges from a vertex in B to a vertex in A. In this case, each vertex in A has weight n, $w(D[B]) \in W(k)$, and $w(D) = n(n-k) + w(D[B])$. □

We now provide a coarse lower bound on $b(n)$, where (recalling Definition 1.2) $b(n)$ is the largest integer such that $[n, b(n)] \subseteq W(n)$. While Rao's result [6, Thm. 6] is better asymptotically, Proposition 2.2 provides a more convenient lower bound.

Proposition 2.2. *Let $n \geqslant 1$ such that $n \neq 7$. Then, $b(n) \geqslant (3/4)n^2$, and $b(7) = 35$.*

Proof. When $n \neq 7$, this follows by inspection for $n \leqslant 18$ (see Appendix A.1, in which the data is calculated using Lemma 2.1) and Rao's lower bound $b(n) \geqslant (n - \lfloor n^{0.57} \rfloor)(n-1) + n$ [6, Thm. 6] when $n \geqslant 19$. For the case $n = 7$, we get $b(7) \geqslant 35$ by inspection (see again Appendix A.1), and $36 \notin W(7)$ by [6, p. 1597], so in fact $b(7) = 35$. □

Proposition 2.2 shows that for each $n \neq 7$ and each k between n and $\lceil (3/4)n^2 \rceil$, there always exists a digraph on n vertices with exactly k reachable pairs. However, in general there will be digraphs on n vertices with weight strictly between $\lceil (3/4)n^2 \rceil$ and n^2. In Theorem 1.3, we will prove that weights in $W(n)$ larger than $(3/4)n^2$ can be realized by digraphs that contain at least one mother vertex, which reduces the problem to considering the induced transitive digraph obtained after removing the mother vertices. Consequently, the weight sets can be determined recursively.

We will now discuss the structure of graphs containing mother vertices. When a transitive digraph D on a vertex set V has at least one mother vertex, then V admits a partition $V = A \cup B$, where A is the set of all mother vertices. Assuming that B is nonempty, i.e., assuming that D is not a complete digraph, then there are no directed edges from any vertex in B to any vertex in A (if there were an edge from a vertex v in B to any vertex in A, then, since D is a transitive digraph, v would itself be a mother vertex, a contradiction to the definition of B).

Of course, when each vertex in A is a mother vertex, A comprises a clique. So, a starting point for studying transitive digraphs with mother vertices is to examine how the vertex set can be partitioned into cliques.

Lemma 2.3. *Let D be a transitive digraph with directed edge set E, and let $A \cup B$ be a partition of the vertex set. If both A and B are cliques, then either there is a directed edge from each vertex in A to each vertex in B; or, there are no directed edges from any vertex in A to any vertex in B.*

Proof. Let $u, u' \in A$ and $v, v' \in B$. If there is a directed edge from u to v, then, by transitivity, there is a directed edge from u' to v'. The result follows. □

The following theorem, which is a consequence of [2, Proposition 2.3.1], shows that there is a nice partition of the vertices into cliques.

Theorem 2.4. *Let D be a transitive digraph. Then there exists a partition $\{V_i : 1 \leqslant i \leqslant t\}$ of the vertex set such that each V_i is a clique, and the following hold.*

(i) If $i < j$, then there is no directed edge from any vertex in V_j to any vertex in V_i.

(ii) If $i < j$, then either there is a directed edge from each vertex in V_i to each vertex in V_j; or, there are no directed edges from any vertex in V_i to any vertex in V_j.

Proof. We proceed by induction on the number of vertices of the digraph $D = (V, E)$. The result is obvious if there are only one or two vertices, so we assume that $n \geqslant 3$ and that the result is true if there are fewer than n vertices. Let D be a transitive digraph with vertex set V and directed edge set E on n vertices. Choose a vertex u such that the number of directed edges in D starting at u is maximum. Define

$$V_1 := \{v \in V : v = u \text{ or } (v, u) \in E\}.$$

First, V_1 is nonempty, since $u \in V_1$. Moreover, if $u' \in V_1$, then by transitivity every out-neighbor of u is an out-neighbor of u'. Since the number of directed edges in D starting at u is a maximum, there must also be a directed edge from u to u', and hence V_1 is a clique. Moreover, if $x \notin V_1$, then by construction there are no directed edges from x to any vertex in V_1. Thus, V_1 is a clique of the desired type, and we may remove the vertices of V_1 from D and partition the remaining vertices in the desired fashion by inductive hypothesis. Using also Lemma 2.3, the result follows. □

Using the partition guaranteed by Theorem 2.4, we can bound the weight of D in terms of the size of the largest clique.

Proposition 2.5. *Let D be a transitive digraph on n vertices, and let ω be the size of the largest clique in D. Then, $w(D) \leqslant n(n+\omega)/2$. In particular, if $\omega \leqslant n/2$, then $w(D) \leqslant (3/4)n^2$.*

Proof. Let D be a transitive digraph with partition $\{V_i : 1 \leqslant i \leqslant t\}$ as given by Theorem 2.4 and let $n_i := |V_i|$; note that $\omega = \max\{n_i : 1 \leqslant i \leqslant t\}$. For a fixed k and all $j < k$, there are no directed edges from any vertex in V_k to any vertex in V_j, and, for all $j > k$, there are no directed edges from any vertex in V_j to any vertex in V_k. This means there at least

$$n_k \left(\sum_{j \neq k} n_j \right) = n_k(n - n_k)$$

pairs that are not reachable that include vertices in the clique V_k. If we were to sum over all k, we will have counted each such pair twice, so the total number of pairs that are *not* reachable is at least

$$\frac{1}{2} \sum_{k=1}^{t} n_k(n - n_k) = \frac{1}{2} n \sum_{k=1}^{t} n_k - \frac{1}{2} \sum_{k=1}^{t} n_k^2.$$

Now, $\sum_{k=1}^{t} n_k = n$ and $n_k \leqslant \omega$ for all k, so

$$\frac{1}{2} n \sum_{k=1}^{t} n_k - \frac{1}{2} \sum_{k=1}^{t} n_k^2 \geqslant \frac{1}{2} n^2 - \frac{1}{2} \sum_{k=1}^{t} \omega n_k = \frac{1}{2} n^2 - \frac{1}{2} n\omega.$$

Hence, an upper bound on the number of pairs that *are* reachable is $n^2 - (\frac{1}{2}n^2 - \frac{1}{2}n\omega) = n(n+\omega)/2$, as desired. □

The converse of Proposition 2.5 tells us that when $w(D) > (3/4)n^2$, there must be a clique of size greater than $n/2$.

We are now able to prove Theorem 1.3, which says that any sufficiently large weight in $W(n)$ can be realized by a digraph with at least one mother vertex, and hence the larger values in $W(n)$ can be found by examining $W(k)$ for $k < n$.

Proof of Theorem 1.3. The result is obvious if $w(D) = n^2$, so assume that $(3/4)n^2 < w(D) < n^2$ and that the vertices of D are partitioned in the form given by Theorem 2.4, that is, assume that $\{V_i : 1 \leqslant i \leqslant t\}$ is a partition of the vertex set such that each V_i is a clique and, if $i < j$, there are no directed edges from any vertex in V_j to any vertex in V_i and there is either a directed edge from each vertex in V_i to each vertex in V_j or there are no directed edges from any vertex in V_i to any vertex in V_j. Let V_k be the largest clique in D, let $\omega = |V_k|$, and let $d = n - \omega$. Since $w(D) > (3/4)n^2$, we have $d < n/2$ by Proposition 2.5.

Let $D_1 = D[V \setminus V_k]$; then, D_1 is a transitive digraph on d vertices. Consider a vertex u in $V \setminus V_k$. We say that u is *weakly adjacent* to a vertex v if either (u, v) or (v, u) is a directed edge. If u is weakly adjacent to a vertex in V_k, then since V_k is a clique, u is weakly adjacent to every vertex in V_k. If u is the initial vertex of a directed edge to some vertex in V_k and is the terminal vertex of a directed edge from some vertex in V_k, then u is part of the clique. As this is not the case, we conclude that if u is weakly adjacent to a vertex in V_k, then there are exactly $\omega = |V_k|$ directed edges between u and V_k.

Let c be the number of vertices in $V \setminus V_k$ that are weakly adjacent to a vertex in V_k. Then, there are exactly

$$\omega^2 + \omega c = (n - d)^2 + (n - d)c$$

reachable pairs of D that contain a vertex from V_k. Thus,

$$w(D) = (n - d)^2 + (n - d)c + w(D_1).$$

Next, since $c \leqslant d < n/2$, we have $d - c < n - d$. Let B be any subset of V obtained by deleting $(n - d) - (d - c)$ vertices from V_k, and let $D_2 = D[B]$. Then, D_2 is transitive and has $n - (n - 2d + c) = 2d - c$ vertices, and

$$w(D_2) = (d - c)^2 + (d - c)c + w(D_1).$$

Some basic manipulation then shows that

$$\begin{aligned} w(D) &= (n - d)^2 + (n - d)c + w(D_1) \\ &= n(n - 2d + c) + (d - c)^2 + (d - c)c + w(D_1) \\ &= n(n - (2d - c)) + w(D_2). \end{aligned}$$

Thus, we may form the transitive digraph D' on n vertices by starting with the transitive digraph D_2 on $2d - c$ vertices and adding $n - (2d - c)$ mother vertices. The vertex set of D' is $A \cup B$, where A contains the newly added mother vertices and B is the vertex set of D_2. Then, $w(D') = w(D)$, and D' has at least one mother vertex. □

When D has at least one mother vertex, the weight of D depends entirely on the weight of the subgraph induced by the non-mother vertices. Thus, we now have a method to determine the sets $W(n)$ recursively.

Corollary 2.6. *Let $n \geqslant 1$ and let $\mathcal{B}_n := \{n(n-k) + d : d \in W(k), 1 \leqslant k \leqslant n-1\}$. If $n = 7$, then*

$$W(7) = [7, 35] \cup \mathcal{B}_7 \cup \{49\},$$

and if $n \neq 7$ then

$$W(n) = [n, \lceil (3/4)n^2 \rceil] \cup \mathcal{B}_n \cup \{n^2\}.$$

Proof. We know that $\mathcal{B}_n \cup \{n^2\} \subseteq W(n)$ for all n from Lemma 2.1. By Proposition 2.2, we have $[n, \lceil(3/4)n^2\rceil] \subseteq W(n)$ for $n \neq 7$ and $b(7) = 35$. So,

$$W(7) \supseteq [7, 35] \cup \mathcal{B}_7 \cup \{49\},$$

and when $n \neq 7$,

$$W(n) \supseteq [n, \lceil(3/4)n^2\rceil] \cup \mathcal{B}_n \cup \{n^2\}.$$

For the reverse containments, first assume that $n \neq 7$ and that D is a transitive digraph on n vertices that is not a complete graph. If $w(D) \leqslant (3/4)n^2$, then $w(D) \leqslant b(n)$, and we are done. Otherwise, $w(D) > (3/4)n^2$, and by Theorem 1.3, $w(D) = w(D^*)$ for some digraph D^* whose number of reachable pairs is an element of $\mathcal{B}_n$, and we are done.

When $n = 7$, we have $(3/4) \cdot 7^2 = 36.75$, so the arguments of the previous paragraph apply to transitive digraphs with at least 37 reachable pairs. We also know that $[7, 35] \subseteq W(7)$, so the only remaining question is whether there is a transitive digraph on 7 vertices with exactly 36 reachable pairs. However, this is ruled out by computations performed in [6, p. 1597]. Hence, $W(7) = [7, 35] \cup \mathcal{B}_7 \cup \{49\}$, and the proof is complete. □

3 A recursive formula for $b(n)$

In this section, we demonstrate the power of Theorem 1.3 and more closely examine the function $b(n)$, which is equal to the largest positive integer such that $[n, b(n)] \subseteq W(n)$, and thus bounds the beginning values in $W(n)$. By Proposition 2.2 and Theorem 1.3, we know that $b(n) \geqslant (3/4)n^2$ for all $n \neq 7$, and that any number of reachable pairs greater than $(3/4)n^2$ can be realized via a digraph D with vertex set $V = A \cup B$, where A is the nonempty set of mother vertices of V. We note also that the induced subgraph $D[B]$ is a transitive digraph on fewer than n vertices. By examining transitive digraphs with mother vertices, we can find integers that are in $[n, n^2] \setminus W(n)$, and since $b(n) + 1$ is the smallest such integer, this gives us useful information about $b(n)$.

Let us consider one way in which a gap could appear in the weight set $W(n)$. If $|B| = z$, then it is possible that $w(D) = n^2 - zn + b(z)$, but $w(D) \neq n^2 - zn + b(z) + 1$, because $w(D[B]) \neq b(z) + 1$. If, in addition, $n^2 - zn + b(z) + 1$ is not the weight of any digraph with $z - 1$ non-mother vertices, then it is plausible that $n^2 - zn + b(z) + 1 \notin W(n)$. This will occur if

$$n^2 - zn + b(z) + 2 \leqslant n^2 - (z - 1)n + z - 1.$$

Solving this inequality for n yields $n \geqslant b(z) - z + 3$. This inspires the next definition.

Definition 3.1. For all $z \in \mathbb{N}$, we define $\ell(z) := b(z) - z + 3$.

Clearly, we must know $b(z)$ to be able to calculate $\ell(z)$. However, it turns out that knowing $\ell(z)$ allows us to compute $b(n)$ for some values of $n \geqslant \ell(z)$. The relationship (barring some small exceptions) between $b(n)$ and $\ell(z)$ is that if $\ell(z) \leqslant n < \ell(z+1)$, then $b(n) = n^2 - zn + b(z)$. Thus, if we know $b(k)$ for $k \in [1, n-1]$, then we can calculate each $\ell(k)$, find the appropriate z, and then calculate $b(n)$. See Appendix A.2, which provides pseudocode and Mathematica code for computing b, ℓ, and z, and lists the values of these functions for $1 \leqslant n \leqslant 24$.

Most of this section is dedicated to proving the statements of the previous paragraph. The main theorem is Theorem 1.4, and the majority of the work is done in Propositions 3.4 and 3.6. Computational lemmas are introduced as they are needed to prove the propositions.

Lemma 3.2. *For all $n \geqslant 1$ and all $m \geqslant 1$, $b(n) + m \leqslant b(n+m)$.*

Proof. For any $n \geqslant 1$, we can form a transitive digraph D on $n+m$ vertices by taking a transitive digraph D' on n vertices with between n and $b(n)$ reachable pairs and adding m isolated vertices, yielding transitive digraphs on $n+m$ vertices with between $n+m$ and $b(n)+m$ reachable pairs. The result follows. □

Lemma 3.3. *Let $z \geqslant 6$. Then,*

(1) $\ell(z) > 4z$,

(2) if $n \geqslant \ell(z)$, then $n^2 - zn + b(z) \geqslant (3/4)n^2$.

Proof. (1) We have $\ell(6) = 25$ and $\ell(7) = 31$, so assume that $z \geqslant 8$. By Proposition 2.2, $b(z) \geqslant (3/4)z^2$, so

$$\ell(z) = b(z) - z + 3 \geqslant (3/4)z^2 - z + 3$$

and it is routine to verify that this is greater than $4z$.

For (2), assume $n \geqslant \ell(z)$ and let $h(x) = (1/4)x^2 - zx + b(z)$. Solving $h(x) = 0$ for x in terms of z yields $x = 2(z \pm \sqrt{z^2 - b(z)})$. We have

$$2(z + \sqrt{z^2 - b(z)}) < 4z < \ell(z) \leqslant n$$

so the desired inequality holds. □

Proposition 3.4. *Let $z \geqslant 4$ and let n be such that $\ell(z) \leqslant n < \ell(z+1)$. Then, $b(n) \geqslant n^2 - zn + b(z)$.*

Proof. Since $z \geqslant 4$, $n \geqslant \ell(4) = 12$, so by Proposition 2.2, $b(n) \geqslant (3/4)n^2$. Moreover, by Lemma 3.3(2), $n^2 - zn + b(z) \geqslant (3/4)n^2$. To get the stated result, we need to show that there exists a digraph on n vertices with exactly m reachable pairs for every integer m satisfying $(3/4)n^2 \leqslant m \leqslant n^2 - zn + b(z)$.

Consider transitive digraphs on n vertices that have exactly $n-k$ mother vertices, and let D_k be the subgraph on k vertices induced by the set of non-mother vertices. We can vary the choice of D_k so that the number of reachable pairs in D_k is any integer between k and $b(k)$ (inclusive). This allows us to produce transitive digraphs on n vertices with numbers of reachable pairs in the interval

$$I_k := [n(n-k)+k, \quad n(n-k)+b(k)],$$

and every weight in I_k is achievable. We will prove that the union $\bigcup_{k=z}^{n-1} I_k$ covers the entire interval $[\lceil (3/4)n^2 \rceil, n^2 - zn + b(z)]$. (Note that, if we were to order the intervals by their left endpoint, then, by our choice of notation, I_{n-1} is the leftmost interval and I_z is the rightmost interval.)

We claim that when $z+1 \leqslant k$, the lower endpoint of I_{k-1} is at most one more than the upper endpoint of I_k. That is, we seek to show that

$$n(n-(k-1)) + (k-1) \leqslant 1 + n(n-k) + b(k) \tag{1}$$

which is equivalent to showing

$$n + k - 1 \leqslant b(k) + 1.$$

Now, by assumption, $n \leqslant \ell(z+1)-1$, and by definition $\ell(z+1) = b(z+1)-(z+1)+3$. From this, we get that $n + z \leqslant b(z+1)+1$. Using Lemma 3.2, we have

$$\begin{aligned} n + k - 1 &= n + z + (k - z - 1) \\ &\leqslant b(z+1) + 1 + (k - (z+1)) \\ &\leqslant b(z+1+k-(z+1)) + 1 \\ &= b(k) + 1. \end{aligned}$$

Thus, (1) holds. This means that the union of the intervals I_k comprises a single interval, and, in particular, contains the interval $[2n-1, n^2 - zn + b(z)]$, which goes from the lower endpoint of I_{n-1} to the upper endpoint of I_z. Clearly, $2n-1 \leqslant (3/4)n^2$, so we conclude that

$$[n, \ \lceil (3/4)n^2 \rceil] \cup [2n-1, \ \ n(n-z)+b(z)] = [n, \ \ n^2 - zn + b(z)]$$

and the stated result follows. □

Lemma 3.5. *Assume that $z \geqslant 4$, $n \geqslant \ell(z)$, and $z+1 \leqslant d < n/2$. Then,*

$$n^2 - dn + d^2 \leqslant n^2 - zn + b(z).$$

Before we prove the result, we note that we need z to be at least 4 in order to have $\ell(z)/2 > z+1$. If $z = 3$, then $\ell(z) = 7$, and $z+1 = 4 > 7/2$. However, once $z \geqslant 4$, one can use Proposition 2.2 to easily show that $\ell(z) > 2z+2$. Hence, our assumption on z, given that $n \geqslant \ell(z)$ and $z+1 \leqslant d < n/2$, is necessary and can be satisfied.

Proof. First, let $h(d) = n^2 - dn + d^2$. Then, h is decreasing for $d < n/2$, so

$$n^2 - dn + d^2 \leqslant n^2 - (z+1)n + (z+1)^2. \tag{2}$$

Next, we claim that $2b(z) \geqslant z^2 + 3z - 2$. This is clear when $4 \leqslant z \leqslant 7$, and for $z \geqslant 8$ one may apply Proposition 2.2 and verify that

$$2b(z) \geqslant (3/2)z^2 \geqslant z^2 + 3z - 2 = (z+1)^2 + z - 3,$$

which implies that

$$n + b(z) \geqslant \ell(z) + b(z) = 2b(z) - z + 3 \geqslant (z+1)^2. \tag{3}$$

Combining (3) and (2) yields

$$n^2 - dn + d^2 \leqslant n^2 - (z+1)n + (z+1)^2 \leqslant n^2 - zn + b(z),$$

as desired. □

The next proposition strengthens Theorem 1.3, and is the key to the proof of Theorem 1.4.

Proposition 3.6. *Let $z \geqslant 6$ and let n be such that $n \geqslant \ell(z)$. Let D be a transitive digraph on n vertices with at least $n^2 - zn + b(z) + 1$ reachable pairs. Then, there exists a transitive digraph D' such that $w(D') = w(D)$ and D' has vertex set $A \cup B$, where A is the set of mother vertices in D', B is the set of non-mother vertices, and $|B| = s$ for some $0 \leqslant s \leqslant z$.*

Proof. The proposition is obvious if D is complete, so assume that $w(D) < n^2$. Since $z \geqslant 6$, we have $w(D) > n^2 - zn + b(z) \geqslant (3/4)n^2$ by Lemma 3.3(2). Hence, Theorem 1.3 can be applied. Let V_k be the largest clique of D with $n_k := |V_k|$, $d := n - n_k$, and, given $u \in V_k$, define

$$c := |\{x \in V : x \notin V_k, (u,x) \text{ or } (x,u) \in E\}|,$$

that is, c is the number of vertices in $V \setminus V_k$ that are weakly adjacent to a vertex in V_k. From the proof of Theorem 1.3, we know that there is a transitive digraph D' such that $w(D') = w(D)$ and the vertex set of D' can be partitioned into $A \cup B$, where A is the set of all mother vertices in D', B is the set of non-mother vertices, and B consists of $s := 2d - c$ vertices. It remains to show that $s \leqslant z$.

As noted above, $w(D) > (3/4)n^2$, so $n_k > n/2$ by Proposition 2.5 and hence $d = n - n_k < n/2$. Suppose that $z + 1 \leqslant d < n/2$. Then, the maximum number of reachable pairs in D is

$$w(D) \leqslant n(n-d) + d^2,$$

which by Lemma 3.5 is strictly less than $n^2 - zn + b(z) + 1$. So, we must have $d \leqslant z$.

By Lemma 3.3(1), $n > 4z$ and we are assuming that $d \leqslant z$, so we get $d < n/4$. This gives $n - 2d + c > n/2$ and $s < n/2$. If $z + 1 \leqslant s < n/2$, then using Lemma 3.5 shows that

$$\begin{aligned} w(D') &= n(n - 2d + c) + w(D'[B]) \\ &\leqslant n(n-s) + s^2 \\ &< n^2 - zn + b(z) + 1. \end{aligned}$$

This contradicts the fact that $w(D') = w(D)$. Thus, $s \leqslant z$, as required. □

We are now ready to prove Theorem 1.4.

Proof of Theorem 1.4. For $z \in [1, 5]$, the theorem can be proved by inspection using Corollary 2.6 and Table 1; see Appendix A.2. Note that the case $n = 8$ occurs when $z = 3$; see the remark following this proof for an explanation of why the result does not hold in this instance.

Assume that $z \geqslant 6$. We know that $b(n) \geqslant n^2 - zn + b(z)$ by Proposition 3.4. Let D be a transitive digraph on n vertices with $w(D) > n^2 - zn + b(z)$. By Proposition 3.6, we may assume that the vertex set of D is $A \cup B$, where A is the set of all mother vertices in D, B is the set of non-mother vertices, and $|B| = s$ for some $0 \leqslant s \leqslant z$. We will argue that $w(D) \geqslant n^2 - zn + b(z) + 2$.

We have $w(D) = n(n - s) + w(D[B])$. If $s = z$, then $w(D[B]) > b(z)$. As $b(z) + 1 \notin W(z)$, we must have $w(D[B]) \geqslant b(z) + 2$ and we are done. So, assume that $s \leqslant z - 1$. Then, a lower bound for $w(D)$ is

$$w(D) = n(n - s) + w(D[B]) \geqslant n(n - s) + s. \tag{4}$$

Now, $n(n - s) + s$ is decreasing as a function of s, so (4) and the fact that $n \geqslant \ell(z)$ yields

$$\begin{aligned} w(D) &\geqslant n(n - (z - 1)) + z - 1 \\ &= n^2 - zn + z - 1 + n \\ &\geqslant n^2 - zn + z - 1 + (b(z) - z + 3) \\ &= n^2 - zn + b(z) + 2. \end{aligned}$$

Hence, $n^2 - zn + b(z) + 1 \notin W(n)$ and therefore $b(n) = n^2 - zn + b(z)$. □

Remark 3.7. When $n = 8$, the corresponding z is $z = 3$, for which $\ell(z) = 7$. Computing $n^2 - zn + b(z)$ in this case yields 47, but $b(8) = 52$ by Table 1. What differs in this situation is that, when $n = 8$,

$$(3/4)n^2 = n^2 - zn + b(z) + 1,$$

and there exists a digraph on 8 vertices of weight 48: namely, start with a complete digraph on 4 vertices and add to it exactly 4 mother vertices. Lemma 3.3(2) shows that such coincidences cannot happen when z (and n) are sufficiently large.

Our first application of Theorem 1.4 is to use it to improve Corollary 2.6 for $n \geqslant 8$, since we now need only check the values of k through z (as opposed to $n - 1$).

Corollary 3.8. *Let $n \geqslant 8$ and let $\mathcal{B}_n := \{n(n - k) + d : d \in W(k), 1 \leqslant k \leqslant z\}$, where z is such that $\ell(z) \leqslant n < \ell(z + 1)$. Then,*

$$W(n) = [n, b(n)] \cup \mathcal{B}_n \cup \{n^2\}.$$

Proof. Certainly, $[n, b(n)] \cup \mathcal{B}_n \cup \{n^2\} \subseteq W(n)$ by Corollary 2.6. For $8 \leqslant n \leqslant 11$, the result follows by direct calculation and comparing the sets listed in Corollaries 2.6 and 3.8. By Proposition 2.2 and Theorem 1.4, $b(n) = n^2 - zn + b(z) \geqslant (3/4)n^2$ when $n \geqslant 12$, and, by Proposition 2.5, if $w(D) > (3/4)n^2$, this means $\omega > n/2$. Thus, if $w(D) > (3/4)n^2$, then $w(D) = n(n-k) + d$ for some $k < n/2$ and $d \in W(k)$. Since $n \geqslant 12$, we have $z \geqslant 4$, so, by Lemma 3.5, if $z+1 \leqslant k < n/2$ and $w(D) = n(n-k)+d$ for some $k < n/2$ and $d \in W(k)$, we have

$$w(D) \leqslant n(n-k) + k^2 \leqslant n^2 - zn + b(z) = b(n),$$

and so we only need to check values of k that are at most z. The result follows. □

4 Estimating $b(n)$

The purpose of this section is to estimate $b(n)$: that is, given n, can we get a reasonably accurate estimate of $b(n)$ without any knowledge of $b(m)$ for $m < n$? In Section 5, we will consider the same question for $|W(n)|$. By Corollary 3.8, such an estimate for $b(n)$ would be useful toward finding an estimate for $|W(n)|$, and, in light of Theorem 1.4, finding an estimate for $b(n)$ more or less reduces to having an accurate estimate for the integer z such that $\ell(z) \leqslant n < \ell(z+1)$.

Definition 4.1. For each $n \in \mathbb{N}$, $n \geqslant 3$, we define $\zeta(n)$ to be the unique positive integer such that $\ell(\zeta(n)) \leqslant n < \ell(\zeta(n)+1)$.

For computation of some values of $\zeta(n)$ for small values of n, see Appendix A.2.

We will often need to iterate the function ζ. When n is clear from context, we let $z_1 := \zeta(n)$ and $z_k := \zeta^k(n) = \zeta(z_{k-1})$ for each $k \geqslant 2$. With this notation, z_1 has the same definition as the integer z that appeared throughout Section 3.

The majority of this section is devoted to developing an explicit function $r(n)$ to estimate $\zeta(n)$ (Definition 4.5), and proving that r does in fact accurately approximate ζ (Theorem 4.6). Our first lemma lists some simple observations about the functions b, ℓ, and ζ. All of these follow from Theorem 1.4 and the definitions of the functions, and we shall use them freely in our subsequent work.

Lemma 4.2. *Let $x \in \mathbb{N}$, $x \geqslant 9$. Then,*

(1) $b(x) = x^2 - \zeta(x)x + b(\zeta(x))$.

(2) $\ell(x) = x^2 - (\zeta(x)+1)x + b(\zeta(x)) + 3$.

(3) If $x < \ell(\zeta(x)+1) - 1$, *then* $\zeta(x+1) = \zeta(x)$. *If* $x = \ell(\zeta(x)+1) - 1$, *then* $\zeta(x+1) = 1 + \zeta(x)$.

In order to approximate $b(n) = n^2 - z_1 n + b(z_1)$ in terms of n, we require bounds on z_1 and $b(z_1)$. These are obtained below in Lemmas 4.3 and 4.4(3), respectively. Usually, small cases must be checked by hand; this can be accomplished by using Theorem 1.4 to calculate b and ℓ recursively.

Lemma 4.3. *When $n \geqslant 12$, $\sqrt{n} \leqslant z_1 < \sqrt{2n}$, and the lower bound is strict for $n \neq 16$.*

Proof. When $12 \leqslant n \leqslant 46$, the lower bound holds by inspection (see Appendix A.3), and $\sqrt{n} = \zeta(n)$ only when $n = 16$. So, assume $n \geqslant 47$ (which implies that $z_1 \geqslant 8$) and let $t := \zeta(z_1+1)$; then, $t \geqslant 3$. One may compute that $\ell(z_1+1) = z_1^2 - (t-1)z_1 + \ell(t)$. Since $n < \ell(z_1+1)$ and $\ell(t) \leqslant z_1+1$, we obtain

$$n < z_1^2 - (t-1)z_1 + \ell(t) \leqslant z_1^2 - (t-1)z_1 + z_1 + 1 < z_1^2$$

where the last inequality holds because $t \geqslant 3$.

For the upper bound, we apply induction. The bound holds for $12 \leqslant n \leqslant 99$ by inspection, so assume $n \geqslant 100$, which means that $z_1 \geqslant 12$. Using this and the inductive hypothesis, we see that

$$2(z_2+1) < 2(\sqrt{2z_1}+1) < z_1.$$

Now, since $n \geqslant \ell(z_1)$, we have

$$n \geqslant \ell(z_1) = z_1^2 - (z_2+1)z_1 + b(z_2) + 3 > z_1^2 - (z_2+1)z_1 > z_1^2 - \frac{z_1^2}{2}.$$

The result follows. □

Lemma 4.4. *Let $n \in \mathbb{N}$.*

(1) For all $n \geqslant 8$, $n+1 \leqslant b(n+1) - b(n) \leqslant 2n+1$.

(2) If $n \geqslant 4$, then $\ell(n) < b(n) < \ell(n+1)$,

(3) If $n \geqslant 192$, then $|n - b(z_1)| \leqslant z_1 - 3$.

Proof. For (1), we use induction. When $n \leqslant 46$, the lemma holds by inspection (see Appendix A.4). If $n \geqslant 47$ and $n < \ell(z_1+1) - 1$, then $b(n+1) - b(n) = 2n + 1 - z_1$ and the bounds hold. Finally, if $n \geqslant 47$ and $n = \ell(z_1+1)-1$, then one may compute that $b(n+1) - b(n) = n - z_1 + b(z_1+1) - b(z_1)$ and apply the inductive hypothesis to obtain the desired bounds.

Part (2) is true by inspection for $4 \leqslant n \leqslant 8$ (see Appendix A.2), and follows from part (1) and the definition of ℓ for $n \geqslant 9$. For (3), the bounds hold by inspection for $192 \leqslant n \leqslant 480$ (see Appendix A.5). So, assume that $n \geqslant 481$, which means that $z_1 \geqslant 25$ and $z_2 \geqslant 6$. If $n \leqslant b(z_1)$, then the results follows from the definition of $\ell(z_1)$. So, assume that $n > b(z_1)$. We claim that $b(z_1+1) - b(z_1) \leqslant 2z_1 + 1 - z_2$. Indeed, the two expressions are equal when $z_1 < \ell(z_2+1) - 1$, and if $z_1 = \ell(z_2+1) - 1$, then one may check that

$$b(z_1+1) - b(z_1) = 2z_1 + 1 - z_2 - (\ell(z_2) - 1) < 2z_1 + 1 - z_2.$$

Using this inequality and the fact that $n < \ell(z_1+1) = b(z_1+1) - (z_1+1) + 3$, we obtain

$$n - b(z_1) < b(z_1+1) - b(z_1) - (z_1+1) + 3 \leqslant 2z_1 + 1 - z_2 - (z_1+1) + 3 \leqslant z_1 - 3,$$

as desired. □

At this point, we know that (for sufficiently large n), $n \approx b(z_1)$ and $z_1 \approx b(z_2)$. So,

$$n \approx b(z_1) = z_1^2 - z_2 z_1 + b(z_2) \approx z_1^2 - (z_2 - 1)z_1.$$

From this, we obtain the estimate $z_1 \approx n^{\frac{1}{2}} + \frac{1}{2}(z_2 - 1)$, which suggests that

$$\zeta(n) \approx n^{\frac{1}{2}} + \frac{1}{2}\big(\zeta(n^{\frac{1}{2}}) - 1\big).$$

Thus, to approximate $\zeta(n)$, we should construct a function $r(n)$ that satisfies

$$r(n) = n^{\frac{1}{2}} + \frac{1}{2}\big(r(n^{\frac{1}{2}}) - 1\big). \tag{5}$$

Definition 4.5. Given a real number $n \geqslant 3$, define $N := \lfloor \log_2 \log_5 n \rfloor + 1$, and define

$$r(n) := \left(\sum_{j=1}^{N} \frac{1}{2^{j-1}} n^{\frac{1}{2^j}} \right) - \frac{2^{N-1} - 1}{2^{N-1}}.$$

We briefly discuss our motivation for our choice of N. Supposing that $r(n) \approx z_1$, $r(n^{\frac{1}{2}}) \approx z_2$, etc., the question then becomes how often we would need to iterate until $z_N = 1$. If $z_N \approx n^{\frac{1}{2^N}}$, then this implies that N is the least integer such that $n^{\frac{1}{2^N}} < \ell(2) = 5$, i.e., $N = \lfloor \log_2 \log_5 n \rfloor + 1$. Moreover, direct calculation shows that $r(n)$ does satisfy (5), and we will prove that $r(n)$ is a good approximation for $\zeta(n)$.

Theorem 4.6. *For all $n \geqslant 3$, $|\zeta(n) - r(n)| < 1.985$.*

To prove Theorem 4.6, we use the triangle inequality and (5) multiple times to split $|\zeta(n) - r(n)|$ into six summands, each of which can be bounded individually. Letting $t = \zeta(\lceil n^{\frac{1}{4}} \rceil)$, one may obtain

$$\begin{aligned}
|\zeta(n) - r(n)| &= \left| z_1 - \left(n^{\frac{1}{2}} + \frac{r(n^{\frac{1}{2}}) - 1}{2} \right) \right| \\
&\leqslant \left| z_1 - \left(n^{\frac{1}{2}} + \frac{z_2 - 1}{2} \right) \right| + \frac{1}{2} \left| z_2 - r(n^{\frac{1}{2}}) \right| \\
&= \left| z_1 - \left(n^{\frac{1}{2}} + \frac{z_2 - 1}{2} \right) \right| + \frac{1}{2} \left| z_2 - \left(n^{\frac{1}{4}} + \frac{r(n^{\frac{1}{4}}) - 1}{2} \right) \right| \\
&\leqslant \left| z_1 - \left(n^{\frac{1}{2}} + \frac{z_2 - 1}{2} \right) \right| + \frac{1}{2} \left| z_2 - \left(z_1^{\frac{1}{2}} + \frac{z_3 - 1}{2} \right) \right| + \frac{1}{2} \left| z_1^{\frac{1}{2}} - n^{\frac{1}{4}} \right| \\
&\quad + \frac{1}{4} \left| z_3 - r(n^{\frac{1}{4}}) \right| \\
&\leqslant \left| z_1 - \left(n^{\frac{1}{2}} + \frac{z_2 - 1}{2} \right) \right| + \frac{1}{2} \left| z_2 - \left(z_1^{\frac{1}{2}} + \frac{z_3 - 1}{2} \right) \right| + \frac{1}{2} \left| z_1^{\frac{1}{2}} - n^{\frac{1}{4}} \right| \\
&\quad + \frac{1}{4} |z_3 - t| + \frac{1}{4} \left| t - r(\lceil n^{\frac{1}{4}} \rceil) \right| + \frac{1}{4} \left| r(\lceil n^{\frac{1}{4}} \rceil) - r(n^{\frac{1}{4}}) \right|.
\end{aligned} \tag{6}$$

We now present a series of lemmas that allow us to bound each of these terms.

Lemma 4.7. *When* $n \geqslant 33808$,

$$\left| z_1 - \left(n^{\frac{1}{2}} + \frac{z_2 - 1}{2} \right) \right| < \frac{\sqrt{2}}{\sqrt{z_1}} + \frac{\frac{5}{4} - \frac{3}{z_1} + \left(\frac{1}{8z_1}\right)^{\frac{1}{4}}}{2\sqrt{1 - \frac{\sqrt{2}}{\sqrt{z_1}} - \frac{1}{z_1}}}.$$

Proof. First, the condition on n implies that $z_1 \geqslant 192$, so by Lemma 4.4(3) we have $|n - b(z_1)| \leqslant z_1 - 3$ and $|z_1 - b(z_2)| \leqslant z_2 - 3$. Second, from $b(z_1) = z_1^2 - z_2 z_1 + b(z_2)$, we obtain

$$b(z_1) + \frac{(z_2 - 1)^2}{4} = \left(z_1 - \frac{z_2 - 1}{2} \right)^2 + (b(z_2) - z_1). \tag{7}$$

For readability, let $c = b(z_1) + \frac{1}{4}(z_2 - 1)^2$. Then, (7) yields

$$\left| \left(z_1 - \frac{z_2 - 1}{2} \right)^2 - c \right| \leqslant |z_1 - b(z_2)| \leqslant z_2 - 3.$$

Next, we use the triangle inequality, rationalize the numerators, and use the bounds we have already discussed in this proof:

$$\begin{aligned} \left| z_1 - \left(n^{\frac{1}{2}} + \frac{z_2 - 1}{2} \right) \right| &\leqslant \left| \left(z_1 - \frac{z_2 - 1}{2} \right) - \sqrt{c} \right| + \left| \sqrt{c} - n^{\frac{1}{2}} \right| \\ &= \frac{\left| \left(z_1 - \frac{z_2 - 1}{2} \right)^2 - c \right|}{\left(z_1 - \frac{z_2 - 1}{2} \right) + \sqrt{c}} + \frac{|c - n|}{\sqrt{c} + n^{\frac{1}{2}}} \\ &\leqslant \frac{z_2 - 3}{\left(z_1 - \frac{z_2 - 1}{2} \right) + \sqrt{c}} + \frac{(z_1 - 3) + \frac{(z_2 - 1)^2}{4}}{\sqrt{c} + n^{\frac{1}{2}}}. \end{aligned}$$

Using Lemma 4.3, we have $\sqrt{z_1} < z_2 < \sqrt{2z_1}$, and so we bound the first term by

$$\frac{z_2 - 3}{\left(z_1 - \frac{z_2 - 1}{2} \right) + \sqrt{c}} < \frac{z_2}{z_1} < \frac{\sqrt{2}}{\sqrt{z_1}}.$$

Note that $z_1 \geqslant \ell(z_2) > z_2^2 - (z_3 + 1) z_2$. Via Lemma 4.3, this gives $(z_2 - 1)^2 < z_1 + z_2 z_3 < z_1 + 2^{\frac{5}{4}} z^{\frac{3}{4}}$. Moreover, both c and n are bounded below by $\ell(z_1) > z_1^2 - (z_2 + 1) z_1$. By applying these bounds and canceling factors of z_1, we get

$$\frac{(z_1 - 3) + \frac{(z_2 - 1)^2}{4}}{\sqrt{c} + n^{\frac{1}{2}}} < \frac{(z_1 - 3) + \frac{z_1}{4} + \frac{z_2 z_3}{4}}{2\sqrt{z_1^2 - (z_2 + 1) z_1}} < \frac{\frac{5}{4} - \frac{3}{z_1} + \left(\frac{1}{8z_1}\right)^{\frac{1}{4}}}{2\sqrt{1 - \frac{\sqrt{2}}{\sqrt{z_1}} - \frac{1}{z_1}}},$$

which completes the proof. □

Lemma 4.8. *When* $n \geqslant 3350194786$, $\frac{1}{2}|z_1^{\frac{1}{2}} - n^{\frac{1}{4}}| < 0.13$.

Proof. When $n \geqslant 3350194786$, we have $z_1 \geqslant 58006$, and so $|z_1 - n^{\frac{1}{2}}| < \frac{z_2-1}{2} + 0.652$ by Lemma 4.7, which implies that

$$\frac{1}{2}\left|z_1^{\frac{1}{2}} - n^{\frac{1}{4}}\right| = \frac{|z_1 - n^{\frac{1}{2}}|}{2|z_1^{\frac{1}{2}} + n^{\frac{1}{4}}|} < \frac{\frac{z_2-1}{2} + 0.652}{2|z_1^{\frac{1}{2}} + n^{\frac{1}{4}}|} \leqslant \frac{\frac{z_2-1}{2} + 0.652}{2|\ell(z_2)^{\frac{1}{2}} + \ell(\ell(z_2))^{\frac{1}{4}}|}.$$

This fraction is decreasing as z_2 increases, and, when $n \geqslant 3350194786$, $z_2 \geqslant 250$, so $\ell(z_2) \geqslant 58006$ and $\ell(\ell(z_2)) \geqslant 3350194786$. The result follows. □

Lemma 4.9. *Let* $t = \zeta(\lceil n^{\frac{1}{4}} \rceil)$*. If* $n \geqslant 3350194786$*, then* $\frac{1}{4}|z_3 - t| \leqslant 0.25$*.*

Proof. The assumption on n guarantees that $z_3 \geqslant 18$. As always, $z_1 \geqslant \ell(z_2)$ and $z_2 \geqslant \ell(z_3)$, so

$$\begin{aligned} z_1 &\geqslant \ell(z_3)^2 - (z_3+1)\ell(z_3) + b(z_3) + 3 \\ &> \left(\ell(z_3) - \frac{z_3+1}{2}\right)^2 > \left(b(z_3-1) - \frac{z_3+1}{2}\right)^2, \end{aligned}$$

where the last inequality is true by Lemma 4.4(2). Applying this and Lemma 4.8, we see that

$$\begin{aligned} n^{\frac{1}{4}} &= z_1^{\frac{1}{2}} - \left(z_1^{\frac{1}{2}} - n^{\frac{1}{4}}\right) \\ &> b(z_3-1) - \frac{z_3+1}{2} - 1 > b(z_3-1) - (z_3-1) + 3 = \ell(z_3-1), \end{aligned}$$

where the second inequality is valid because $z_3 \geqslant 18$. For an upper bound on $\lceil n^{\frac{1}{4}} \rceil$, we use Lemma 4.3 to get

$$\lceil n^{\frac{1}{4}} \rceil \leqslant \lceil z_1^{\frac{1}{2}} \rceil \leqslant z_2 < \ell(z_3+1).$$

Thus, $\ell(z_3-1) \leqslant \lceil n^{\frac{1}{4}} \rceil < \ell(z_3+1)$, which means that $z_3 - 1 \leqslant t \leqslant z_3$, as required. □

Lemma 4.10. *Let* $x \geqslant 4$ *be an integer, let* $N = \lfloor \log_2 \log_5 x \rfloor + 1$*, and let* y *be a real number such that* $x - 1 < y < x$*.*

(1) If $x \neq 5^{2^d}$ *for any* $d \in \mathbb{N}$*, then* $r(x) - r(y) < \dfrac{2^N - 1}{2^N\sqrt{y}}$*.*

(2) If $x = 5^{2^d}$ *for some* $d \in \mathbb{N}$*, then* $r(x) - r(y) < \dfrac{2^d - 1}{2^d\sqrt{y}} + \dfrac{\sqrt{5}-1}{2^d}$*.*

Proof. Note that $a^{\frac{1}{2^k}} - c^{\frac{1}{2^k}} < a^{\frac{1}{2}} - c^{\frac{1}{2}}$ whenever $a, c, k > 1$. If $x \neq 5^{2^d}$ for all $d \in \mathbb{N}$, then $\lfloor \log_2 \log_5 y \rfloor = \lfloor \log_2 \log_5 x \rfloor$, so

$$\begin{aligned} r(x) - r(y) &= \sum_{j=1}^{N} \frac{x^{\frac{1}{2^j}} - y^{\frac{1}{2^j}}}{2^{j-1}} \leqslant \left(x^{\frac{1}{2}} - y^{\frac{1}{2}}\right) \sum_{j=1}^{N} \frac{1}{2^{j-1}} \\ &= \frac{\left(2 - \frac{1}{2^{N-1}}\right)(x-y)}{\left(x^{\frac{1}{2}} + y^{\frac{1}{2}}\right)} < \frac{2^N - 1}{2^N\sqrt{y}}. \end{aligned}$$

This proves (1). For (2), we have $N = d + 1$ and $\lfloor \log_2 \log_5 y \rfloor = \lfloor \log_2 \log_5 x \rfloor - 1$. Using the above work, we see that

$$\begin{aligned} r(x) - r(y) &= \sum_{j=1}^{N} \frac{1}{2^{j-1}} x^{\frac{1}{2^j}} - \left(1 - \frac{1}{2^{N-1}}\right) - \sum_{j=1}^{N-1} \frac{1}{2^{j-1}} y^{\frac{1}{2^j}} + \left(1 - \frac{1}{2^{N-2}}\right) \\ &= \sum_{j=1}^{N-1} \frac{x^{\frac{1}{2^j}} - y^{\frac{1}{2^j}}}{2^{j-1}} + \frac{1}{2^{N-1}} - \frac{1}{2^{N-2}} + \frac{1}{2^{N-1}} x^{\frac{1}{2^N}} \\ &< \frac{2^{N-1} - 1}{2^{N-1}\sqrt{y}} + \frac{\sqrt{5} - 1}{2^{N-1}}, \end{aligned}$$

which is the desired bound. □

We now have what we need to prove that the difference between $\zeta(n)$ and $r(n)$ is bounded absolutely.

Proof of Theorem 4.6. We proceed by induction. First, to show that the result holds for $n < 3350194786$, we note that there are only four possibilities for $N = \lfloor \log_2 \log_5 n \rfloor + 1$ in these cases: namely, $N = 1$ when $n < 25$; $N = 2$ when $25 \leqslant n < 625$; $N = 3$ when $625 \leqslant n < 5^8 = 390625$; and $N = 4$ for $5^8 \leqslant n < 3350194786$. When $n < 3350194786$, $z_1 \leqslant 58005$, and, with the exceptions of when $n = 5^{2^d}$ for $1 \leqslant d \leqslant 3$, $r(n)$ is an increasing function for each fixed z_1. This means, for nearly every $z_1 \leqslant 58005$, one only needs to check the extreme possibilities $n = \ell(z_1)$ and $n = \ell(z_1 + 1) - 1$, which is approximately 116000 cases (as opposed to more than 3 billion). We have verified these cases computationally, so the theorem indeed holds for $n < 3350194786$; see Appendix A.6.

Assume now for some fixed $n \geqslant 3350194786$ that the result holds for all integers at least 3 and less than n. It suffices to provide bounds for each summand in (6), which can be accomplished via Lemmas 4.7, 4.8, 4.9, the inductive hypothesis, and Lemma 4.10. Let $t = \zeta(\lceil n^{\frac{1}{4}} \rceil)$. By the inductive hypothesis, we have $\frac{1}{4}|t - r(\lceil n^{\frac{1}{4}} \rceil)| < 0.5$. Since $n \geqslant 3350194786$, we have $z_1 \geqslant 58006$ and $z_2 \geqslant 250$. If $\lceil n^{\frac{1}{4}} \rceil \neq 5^{2^d}$ for any $d \in \mathbb{N}$, then we can bound $|r(\lceil n^{\frac{1}{4}} \rceil) - r(n^{\frac{1}{4}})|$ by using Lemma 4.10 with $N = 2$ and $y = 3350194786^{\frac{1}{4}}$. This gives

$$|\zeta(n) - r(n)| < 0.652 + 0.4091 + 0.13 + 0.25 + 0.5 + 0.0121 = 1.9532.$$

On the other hand, if $\lceil n^{\frac{1}{4}} \rceil = 5^{2^d}$ for some d, then $d \geqslant 2$ and $n \geqslant 624^4 + 1$. By Lemma 4.3, $z_1 > \sqrt{n}$ and $z_2 > \sqrt{z_1}$, so $z_1 \geqslant 624^2 + 1$ and $z_2 \geqslant 625$. In this case, applying the same lemmas yields

$$|\zeta(n) - r(n)| < 0.640 + 0.380 + 0.13 + 0.25 + 0.5 + 0.085 = 1.985,$$

and, therefore, the result holds for all $n \geqslant 3$. □

In practice, the biggest difference we have seen between $\zeta(n)$ and $r(n)$ is about 1.45175. Indeed, very often the terms we bounded will be much smaller; for example, we see that

$$\lim_{n\to\infty} \left| r(\lfloor n^{\frac{1}{2}} \rfloor) - r(n^{\frac{1}{2}}) \right| = 0,$$

and, based on the proof of Lemma 4.7, we have

$$\limsup_{n\to\infty} \left| z_1 - \left(n^{\frac{1}{2}} + \frac{z_2 - 1}{2} \right) \right| = \frac{5}{8}, \quad \limsup_{n\to\infty} \frac{1}{2} \left| z_2 - \left(z_1^{\frac{1}{2}} + \frac{z_3 - 1}{2} \right) \right| = \frac{5}{16}$$

whereas these quantities should also be much smaller than this upper bound infinitely often. On the other hand, even if we were to use the triangle inequality to expand (6) out indefinitely to more and more summands, the above limits superior mean that it is likely impossible to bound $|\zeta(n) - r(n)|$ by anything less than $5/8 + 5/16 + 5/32 + \cdots = (5/8)/(1 - 1/2) = 1.25$ for large n.

With Theorem 4.6 in hand, we can establish a corresponding approximation for $b(n)$. This is the content of Theorem 1.5, which we restate for convenience.

Theorem 1.5. *Define* $N := \lfloor \log_2 \log_5 n \rfloor + 1$, *and define*

$$g(n) := n^2 - \left(\sum_{j=1}^{N} \frac{n^{1+\frac{1}{2^j}}}{2^{j-1}} \right) + \left(2 - \frac{1}{2^{N-1}} \right) n.$$

For all $n \geqslant 3$, $|b(n) - g(n)| < 2n$.

Proof. First, note that $g(n) = n^2 - r(n) \cdot n + n$. The theorem can be verified computationally for $3 \leqslant n \leqslant 8888$. For $n \geqslant 8889$, we use Theorems 1.4 and 4.6 and Lemma 4.3. We have

$$\begin{aligned} |b(n) - g(n)| &= \left| \left(n^2 - z_1 n + b(z_1)\right) - \left(n^2 - r(n) \cdot n + n\right) \right| \\ &\leqslant n|z_1 - r(n)| + |b(z_1) - n| \\ &< 1.985n + (z_1 - 3) \\ &< 1.985n + \sqrt{2n} \\ &< 2n, \end{aligned}$$

as desired. □

Theorem 1.5 provides an upper bound for the difference between $b(n)$ and $g(n)$, but, as noted previously, $r(n)$ will often be a much better estimate for $\zeta(n)$. Hence, $g(n)$ will often be more accurate than the bound in Theorem 1.5 indicates.

5 Estimating $|W(n)|$

We close the paper by using the theory we have built up so far to provide an estimate for $|W(n)|$. We begin with a technical result that strengthens Lemma 4.3.

Lemma 5.1.

(1) For all $n \geqslant 3$, $z_1 - \lfloor \sqrt{n} \rfloor < n^{\frac{1}{4}}$. In particular, $z_1 < n^{\frac{1}{2}} + n^{\frac{1}{4}} = \left(1 + \frac{1}{n^{1/4}}\right) n^{\frac{1}{2}}$.

(2) For all $n \geqslant 9$, $n^2 - b(n) < \left(1 + \frac{1}{n^{1/4}}\right) n^{\frac{3}{2}}$.

Proof. Part (1) can be verified by inspection for $3 \leqslant n \leqslant 6560$; see Appendix A.7. So, assume that $n \geqslant 6561$; then, $N = \lfloor \log_2 \log_5 n \rfloor + 1 \geqslant 3$, $\log_2 \log_5 n < n^{\frac{1}{8}}$, and $2.25 < \frac{1}{4} n^{\frac{1}{4}}$. By Theorem 4.6, we have $z_1 < r(n) + 2$, so

$$\begin{aligned} z_1 - \lfloor \sqrt{n} \rfloor &< \sum_{j=2}^{N} \frac{n^{\frac{1}{2^j}}}{2^{j-1}} + \frac{1}{2^{N-1}} + 2 \\ &< \frac{1}{2} n^{\frac{1}{4}} + (N-2) \cdot \frac{1}{4} n^{\frac{1}{8}} + 2.25 \\ &< \frac{1}{2} n^{\frac{1}{4}} + \log_2 \log_5(n) \cdot \frac{1}{4} n^{\frac{1}{8}} + 2.25 \\ &< n^{\frac{1}{4}}. \end{aligned}$$

Part (2) is a consequence of Part (1), since $n^2 - b(n) = z_1 n - b(z_1) < z_1 n$. □

Next, we provide an estimate for $|W(n)|$ that is recursive in nature.

Definition 5.2. For each $n \in \mathbb{N}$, $n \geqslant 3$, we define

$$h(n) := b(n) - (n-1) + \sum_{k=1}^{z_1} |W(k)|.$$

Lemma 5.3. *For $n \geqslant 25$, $||W(n)| - h(n)| < 3n$.*

Proof. The result follows by inspection when $25 \leqslant n \leqslant 388$; see Appendix A.8. Thus, we may assume that $n \geqslant 389$. Note that by Corollary 3.8 any integer $m \in W(n)$ that is larger than $b(n)$ and less than n^2 can be realized by adjoining $n - d$ mother vertices to a transitive digraph on d vertices, where $1 \leqslant d \leqslant z_1$. Hence, $h(n)$ provides an upper bound on $|W(n)|$, and the difference between $h(n)$ and $|W(n)|$ comes from the integers in the interval $[n(n-z_1) + (z_1 - 1), n(n-z_1) + b(z_1)]$, which are counted twice; the integers $m \in W(j) \cap W(k)$, where $j \neq k$; and $\{n^2\}$. We will bound the number of such integers.

Taken together, the integers in the interval $[n(n-z_1)+(z_1-1), n(n-z_1)+b(z_1)]$ and the single integer $\{n^2\}$ account for $b(z_1) - z_1 + 2 = \ell(z_1) - 1 < n$ such integers.

For each $1 \leqslant d \leqslant z_1$, let $I_d = [n(n-d) + d, n(n-d) + d^2]$. A straightforward calculation shows that, when $d \leqslant \lfloor \sqrt{n} \rfloor$,

$$n(n-d) + d^2 < n(n-(d-1)) + (d-1),$$

and so $I_d \cap I_{d-1} = \varnothing$. On the other hand, when $d \leqslant z_1$, we have $n \geqslant \ell(z_1) > b(d) - (d-1)$, which implies

$$n(n-d) + b(d) < n(n-(d-1)) + (d-1).$$

This means the only overlap between I_d and I_{d-1} comes from digraphs on d vertices with weights in $[b(d)+1, d^2]$. Using Lemma 5.1(2), the number of integers in the overlap is at most

$$d^2 - b(d) < \left(1 + \frac{1}{d^{1/4}}\right) d^{\frac{3}{2}} \leqslant \left(1 + \frac{1}{z_1^{1/4}}\right) z_1^{\frac{3}{2}}.$$

Applying Lemmas 5.1(1) and 4.3 and putting this all together, we see that the difference between $h(n)$ and $|W(n)|$ is bounded by

$$n + \left(z_1 - \lfloor\sqrt{n}\rfloor\right)\left(1 + \frac{1}{z_1^{1/4}}\right) z_1^{\frac{3}{2}} < n + n^{\frac{1}{4}} \cdot \left(1 + \frac{1}{n^{1/8}}\right)\left(n^{\frac{1}{2}} + n^{\frac{1}{4}}\right)^{\frac{3}{2}} < 3n,$$

since $n \geqslant 389$, which completes the proof. □

Our goal now is to find an explicit estimate that differs from $h(n)$ by at most a constant times n. Suppose that $\mathfrak{r}(n)$ is our approximation for $h(n)$. Since

$$h(n) = b(n) - (n-1) + \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} |W(k)| + \sum_{\lfloor\sqrt{n}\rfloor+1}^{z_1} |W(k)|, \tag{8}$$

we can build $\mathfrak{r}(n)$ by estimating each part of (8). Our function $g(n)$ is an approximation for $b(n)$, so $b(n) - (n-1) \approx g(n) - n$. For the remaining two terms, the estimates we will use are

$$\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} |W(k)| \approx \int_0^{\sqrt{n}} \mathfrak{r}(x)\, dx, \qquad \sum_{\lfloor\sqrt{n}\rfloor+1}^{z_1} |W(k)| \approx (r(n) - \sqrt{n})n.$$

Putting this together, we believe that a good estimate for $h(n)$ should satisfy

$$\mathfrak{r}(n) = g(n) - n + \int_0^{\sqrt{n}} \mathfrak{r}(x)\, dx + (r(n) - \sqrt{n})n = n^2 - n^{\frac{3}{2}} + \int_0^{\sqrt{n}} \mathfrak{r}(x)\, dx. \tag{9}$$

We show how to construct such a function $\mathfrak{r}(n)$ as an infinite series. Suppose that $\mathfrak{r}(n) = \sum_{j=0}^{\infty} c_j n^{1+\frac{1}{2^j}}$, where each c_j is a constant. Assuming that $\mathfrak{r}(0) = 0$ and that $\mathfrak{r}(n)$ admits an interchange of summations and integration, we have

$$\int_0^{\sqrt{n}} \mathfrak{r}(x)\, dx = \sum_{j=0}^{\infty} \frac{c_j 2^j}{2^{j+1}+1} n^{1+\frac{1}{2^{j+1}}},$$

Comparing the left-hand and right-hand sides of (9) gives

$$c_0 n^2 + c_1 n^{\frac{3}{2}} + \sum_{j=1}^{\infty} c_{j+1} n^{1+\frac{1}{2^{j+1}}} = n^2 + (\tfrac{1}{3}c_0 - 1)n^{\frac{3}{2}} + \sum_{j=1}^{\infty} \frac{c_j 2^j}{2^{j+1}+1} n^{1+\frac{1}{2^{j+1}}}.$$

Equating coefficients, this implies that $c_0 = 1$, $c_1 = -2/3$, and, for all $j \geqslant 1$,

$$c_{j+1} = \frac{c_j 2^j}{2^{j+1}+1}.$$

It follows by induction that, for all $k \geqslant 1$,

$$c_k = \frac{-2^k}{\prod_{i=1}^k (2^i+1)}.$$

We now define our function $\mathfrak{r}(n)$.

Definition 5.4. For each real number $x \geqslant 0$, define

$$\mathfrak{r}(x) := x^2 - \sum_{k=1}^{\infty} \left(\frac{2^k}{\prod_{i=1}^k (2^i+1)} \right) x^{1+\frac{1}{2^k}}.$$

When $x \geqslant 3$, let $N := \lfloor \log_2 \log_5 x \rfloor + 1$, and define the truncated summation

$$\overline{\mathfrak{r}}(x) := x^2 - \sum_{k=1}^{N} \left(\frac{2^k}{\prod_{i=1}^k (2^i+1)} \right) x^{1+\frac{1}{2^k}}.$$

Lemma 5.5.

(1) For each $x \geqslant 0$, $\mathfrak{r}(x)$ is well-defined.

(2) $\mathfrak{r}(x)$ satisfies (9).

(3) For $x \geqslant 25$, $|\overline{\mathfrak{r}}(x) - \mathfrak{r}(x)| < \frac{1}{3}x$.

Proof. For each k, let $c_k = -2^k / \prod_{i=1}^k (2^i+1)$. Clearly, $\mathfrak{r}(0) = 0$, and when $x > 0$, the ratio of consecutive terms in $\sum_{k=1}^{\infty} c_k x^{1+\frac{1}{2^k}}$ is equal to

$$\frac{2}{(2^{k+1}+1)x^{\frac{1}{2^{k+1}}}},$$

which converges to 0 as $k \to \infty$. Hence, $\sum_{k=1}^{\infty} c_k x^{1+\frac{1}{2^k}}$ is absolutely convergent for all $x \geqslant 0$, and $\mathfrak{r}(x)$ is well defined. It is now clear from the discussion prior to Definition 5.4 that $\mathfrak{r}(n) = n^2 - n^{\frac{3}{2}} + \int_0^{\sqrt{n}} \mathfrak{r}(x)\, dx$.

For (3), a straightforward induction shows that $\sum_{k=1}^{m} (-c_k) = 1 - 1/\prod_{i=1}^{m} (2^i+1)$, and so $\sum_{k=1}^{\infty} (-c_k) = 1$. By construction, x and N satisfy $x < 5^{2^N}$, and $N+1 \geqslant 3$ because $x \geqslant 25$. Hence,

$$\overline{\mathfrak{r}}(x) - \mathfrak{r}(x) = \sum_{k=N+1}^{\infty} -c_k x^{1+\frac{1}{2^k}} < \sum_{k=3}^{\infty} -c_k (5x) < \frac{1}{3}x,$$

as required. □

As we will show, one can use either $\mathfrak{r}(n)$ or $\overline{\mathfrak{r}}(n)$ to estimate $|W(n)|$. The truncated series $\overline{\mathfrak{r}}(n)$ is more amenable to calculation, but $\mathfrak{r}(n)$ is easier to work with in proofs. In the lemmas below, we will bound the difference between each term in (8) and its respective term in (9). These will later be used to bound $||W(n)| - \overline{\mathfrak{r}}(n)|$. The bounds we establish are not optimal; we are satisfied as long as our final bound for $||W(n)| - \overline{\mathfrak{r}}(n)|$ is a constant multiple of n.

Lemma 5.6. *For $n \geqslant 3$,*

$$\left|\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} |W(k)| - \int_0^{\sqrt{n}} \mathfrak{r}(x)\,dx\right| < \left(\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} ||W(k)| - \mathfrak{r}(k)|\right) + 2n.$$

Proof. First, by the triangle inequality,

$$\begin{aligned}\left|\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} |W(k)| - \int_0^{\sqrt{n}} \mathfrak{r}(x)\,dx\right| &\leqslant \left(\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} ||W(k)| - \mathfrak{r}(k)|\right)\\ &\quad + \left|\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \mathfrak{r}(k) - \int_0^{\lfloor\sqrt{n}\rfloor} \mathfrak{r}(x)\,dx\right| + \int_{\lfloor\sqrt{n}\rfloor}^{\sqrt{n}} \mathfrak{r}(x)\,dx.\end{aligned}$$

We will show that each of the last two summands on the right-hand side can be bounded by n.

One may verify that $\mathfrak{r}(x)$ is negative for $0 < x < 1$, and $\left|\int_0^1 \mathfrak{r}(x)\,dx\right| \leqslant 1$. Moreover, $\mathfrak{r}(1) = 0$ and $\mathfrak{r}(x)$ is increasing for $x > 1$, so

$$\sum_{k=1}^{\lfloor\sqrt{n}\rfloor-1} \mathfrak{r}(k) \leqslant \int_1^{\lfloor\sqrt{n}\rfloor} \mathfrak{r}(x)\,dx \leqslant \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \mathfrak{r}(k).$$

It follows that

$$\left|\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} \mathfrak{r}(k) - \int_0^{\lfloor\sqrt{n}\rfloor} \mathfrak{r}(x)\,dx\right| < \mathfrak{r}(\lfloor\sqrt{n}\rfloor) + 1 \leqslant n - 1 + 1 = n.$$

Finally, since $\mathfrak{r}(x)$ is increasing for on $\lfloor\sqrt{n}\rfloor \leqslant x \leqslant \sqrt{n}$, we get

$$\int_{\lfloor\sqrt{n}\rfloor}^{\sqrt{n}} \mathfrak{r}(x)\,dx \leqslant \mathfrak{r}(\sqrt{n})\left(\sqrt{n} - \lfloor\sqrt{n}\rfloor\right) < \mathfrak{r}(\sqrt{n}) < n,$$

which implies the stated result. □

Lemma 5.7. *For $n \geqslant 389$, we have*

$$\left|\sum_{k=\lfloor\sqrt{n}\rfloor+1}^{z_1} |W(k)| - (r(n) - \sqrt{n})n\right| < 7n.$$

Proof. For readability, let $m = \lfloor\sqrt{n}\rfloor + 1$. By the triangle inequality, we have

$$\left|\sum_{k=m}^{z_1} |W(k)| - (r(n) - \sqrt{n})n\right| \leqslant \left|\sum_{k=m}^{z_1} (|W(k)| - b(k))\right| + \left|\sum_{k=m}^{z_1} b(k) - (z_1 - \lfloor\sqrt{n}\rfloor)n\right| + \left|(z_1 - \lfloor\sqrt{n}\rfloor)n - (r(n) - \sqrt{n})n\right|.$$

Since by Lemma 5.1(2) we have

$$||W(k)| - b(k)| < k^2 - b(k) < \left(1 + \frac{1}{k^{1/4}}\right) k^{\frac{3}{2}} \leqslant \left(1 + \frac{1}{z_1^{1/4}}\right) z_1^{\frac{3}{2}},$$

we can use Lemma 5.1(1) to bound the first summand on the right-hand side by

$$\left|\sum_{k=m}^{z_1} (|W(k)| - b(k))\right| < (z_1 - \lfloor\sqrt{n}\rfloor)(z_1^2 - b(z_1)) < n^{\frac{1}{4}} \cdot \left(1 + \frac{1}{z_1^{1/4}}\right) z_1^{\frac{3}{2}} < 2n$$

in a similar fashion to the bound obtained in the proof of Lemma 5.3.

Next, for the second summand notice that $b(z_1 - 1) < \ell(z_1) \leqslant n$ and, by Lemma 5.1(2),

$$b(m) > m^2 - m^{\frac{3}{2}} - m^{\frac{5}{4}} > n - m^{\frac{3}{2}} - m^{\frac{5}{4}}.$$

Hence, $n - m^{\frac{3}{2}} - m^{\frac{5}{4}} < b(k) < n$ for all $m \leqslant k \leqslant z_1 - 1$, and so the second summand is bounded by

$$|b(z_1) - n| + ((z_1 - 1) - m + 1)(n - (n - m^{\frac{3}{2}} - m^{\frac{5}{4}})),$$

which is in turn bounded by

$$(z_1 - 3) + n^{\frac{1}{4}}(m^{\frac{3}{2}} + m^{\frac{5}{4}}) < (n^{\frac{1}{2}} + n^{\frac{1}{4}}) + n^{\frac{1}{4}}\left((\sqrt{n}+1)^{\frac{3}{2}} + (\sqrt{n}+1)^{\frac{5}{4}}\right) < 2n.$$

Finally, we use Theorem 4.6 to bound the last summand on the right-hand side by

$$\left|(z_1 - \lfloor\sqrt{n}\rfloor)n - (r(n) - \sqrt{n})n\right| \leqslant n\left|z_1 - r(n)\right| + n\left|\sqrt{n} - \lfloor\sqrt{n}\rfloor\right| < 3n.$$

The result follows. □

We are now ready to prove Theorem 1.6.

Proof of Theorem 1.6. By Lemma 5.5(3), it suffices to show that $||W(n) - \mathfrak{r}(n)| < \frac{89}{3}n$ for all $n \geqslant 3$. This claim holds by inspection for $3 \leqslant n \leqslant 509$; see Appendix A.8. Now assume that $n \geqslant 510$ and the result holds for all natural numbers between 3 and $n-1$. From Lemma 5.3, we obtain

$$||W(n)| - \mathfrak{r}(n)| \leqslant ||W(n)| - h(n)| + |h(n) - \mathfrak{r}(n)| < 3n + |h(n) - \mathfrak{r}(n)|.$$

To bound $|h(n) - \mathfrak{r}(n)|$, we use (8), (9), Theorem 1.5, and Lemmas 5.7 and 5.6 to get

$$\begin{aligned} |h(n) - \mathfrak{r}(n)| &\leqslant \big|b(n) - (n-1) - g(n) + n\big| + \left|\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} |W(k)| - \int_0^{\sqrt{n}} \mathfrak{r}(x)\,dx\right| \\ &\quad + \left|\sum_{k=\lfloor\sqrt{n}\rfloor+1}^{z_1} |W(k)| - (r(n) - \sqrt{n} - 1)n\right| \\ &< 1 + 11n + \sum_{k=1}^{\lfloor\sqrt{n}\rfloor} ||W(k)| - \mathfrak{r}(k)|. \end{aligned}$$

Applying the inductive hypothesis to $||W(k)| - \mathfrak{r}(k)|$ and recalling that $n \geqslant 510$ yields

$$\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} ||W(k)| - \mathfrak{r}(k)| < 30\sum_{k=1}^{\lfloor\sqrt{n}\rfloor} k = 15\lfloor\sqrt{n}\rfloor\left(\lfloor\sqrt{n}\rfloor + 1\right) < 15(n + \sqrt{n}) < \frac{47}{3}n - 1.$$

Combining all the bounds shows that $||W(n)| - \mathfrak{r}(n)| < \frac{89}{3}n$, and the theorem follows. □

A Data and Mathematica code

A.1 Data for Proposition 2.2

If we have already determined the weight sets $W(1), W(2), \ldots, W(n-1)$, then Lemma 2.1 can be used to construct a subset of $W(n)$. It is easily seen that $W(1) = \{1\}$, $W(2) = [2, 4]$, and $W(3) = [3, 7] \cup \{9\}$. Using these as a starting point, we can produce subsets of $W(n)$ for $n \geqslant 4$, and these subsets can be used to find lower bounds on $b(n)$. Table 2 summarizes these results for $1 \leqslant n \leqslant 18$ and lists the corresponding values of $(3/4)n^2$.

n	$b(n)$ lower bound	$(3/4)n^2$	n	$b(n)$ lower bound	$(3/4)n^2$
1	1	0.75	10	77	75
2	4	3	11	95	90.75
3	7	6.75	12	109	108
4	13	12	13	130	126.75
5	19	18.75	14	153	147
6	28	27	15	178	168.75
7	35	36.75	16	205	192
8	52	48	17	223	216.75
9	61	60.75	18	253	243

Table 2: Lower bounds for $b(n)$, using Lem. 2.1

A.2 Data for Theorem 1.4

Define functions and variables as follows.

- For each $n \geqslant 1$, $b(n)$ is the largest positive integer such that $[n, b(n)] \subseteq W(n)$.
- For each $z \geqslant 1$, $\ell(z) = b(z) - z + 3$.
- For each $n \geqslant 3$, $\zeta(n)$ is the unique positive integer such that $\ell(\zeta(n)) \leqslant n < \ell(\zeta(n) + 1)$. To save space we often let $z := \zeta(n)$ or $z_1 := \zeta(n)$.

Note that $b(n)$ can be found as soon as we know $W(n)$; and $W(n)$ can be computed by using the recursion proved in Corollary 2.6.

n	$b(n)$	$\ell(n)$	$z := \zeta(n)$	$n^2 - zn + b(z)$	n	$b(n)$	$\ell(n)$	$z := \zeta(n)$	$n^2 - zn + b(z)$
1	1	3	–	–	13	130	120	4	130
2	4	5	–	–	14	153	142	4	153
3	7	7	1	7	15	178	166	4	178
4	13	12	1	13	16	205	192	4	205
5	19	17	2	19	17	223	209	5	223
6	28	25	2	28	18	253	238	5	253
7	35	31	3	35	19	285	269	5	285
8	52	47	3	47	20	319	302	5	319
9	61	55	3	61	21	355	337	5	355
10	77	70	3	77	22	393	374	5	393
11	95	87	3	95	23	433	413	5	433
12	109	100	4	109	24	475	454	5	475

Table 3: Values of $b(n)$, $\ell(n)$, and $\zeta(n)$ for $1 \leqslant n \leqslant 24$

A.2.1 Algorithm and Code for $b(n)$

Having established Theorem 1.4, we can compute $b(n)$ without first knowing $W(n)$. Assume that we know the values of $b(k)$ for all $1 \leqslant k \leqslant n$. The value of $b(n+1)$ can be found by implementing the following algorithm:

1. Compute $\ell(k)$ for each $1 \leqslant k \leqslant n$.
2. Find $x := \zeta(n+1)$, which is the integer such that $\ell(x) \leqslant n+1 < \ell(x+1)$. Note that $1 \leqslant x \leqslant n$.
3. Compute $b(n+1) = (n+1)^2 - x(n+1) + b(x)$.

The code below will generated $b(n)$, $\ell(n)$, and $\zeta(n)$ in Mathematica. First, execute this snippet to initialize the variables:

```
Clear[n, b, z, L];
bTable = {1, 4, 7, 13, 19, 28, 35, 52};
L[z_] := bTable[[z]] - z + 3;
```

```
b[n_] := n^2 - z[n]*n + bTable[[z[n]]];
LTable = Table[L[z], {z, 1, Length[bTable]}];
FindLval[n_] := Max[Select[LTable, # <= n &]];
z[n_] := Position[LTable, FindLval[n]][[1]][[1]];
zTable =
   Prepend[Prepend[Table[z[n], {n, 3, Length[LTable]}], 0], 0];
```

Then, execute the following commands to calculate $b(n)$, $\ell(n)$, and $\zeta(n)$ for all n less than or equal to the value of `bound`.

```
bound = 1000;
For[i = Length[bTable] + 1, i <= bound, i++,
 zTable = Append[zTable, z[i]]; bTable = Append[bTable, b[i]];
 LTable = Append[LTable, L[i]]]
```

The `1000` above should be replaced with the desired value of n. The values of b, ℓ, and ζ are stored, respectively, in `bTable`, `LTable`, and `zTable`.

A.3 Data for Lemma 4.3

n	$\sqrt{n}$	$\zeta(n)$	$\sqrt{2n}$	n	$\sqrt{n}$	$\zeta(n)$	$\sqrt{2n}$	n	$\sqrt{n}$	$\zeta(n)$	$\sqrt{2n}$
11	3.32	3	4.690	41	6.40	7	9.055	71	8.43	10	11.92
12	3.46	4	4.899	42	6.48	7	9.165	72	8.49	10	12
13	3.61	4	5.099	43	6.56	7	9.274	73	8.54	10	12.08
14	3.74	4	5.292	44	6.63	7	9.381	74	8.60	10	12.17
15	3.87	4	5.477	45	6.71	7	9.487	75	8.66	10	12.25
16	4	4	5.657	46	6.78	7	9.592	76	8.72	10	12.33
17	4.12	5	5.831	47	6.86	8	9.695	77	8.77	10	12.41
18	4.24	5	6	48	6.93	8	9.798	78	8.83	10	12.49
19	4.36	5	6.164	49	7	8	9.899	79	8.89	10	12.57
20	4.47	5	6.325	50	7.07	8	10	80	8.94	10	12.65
21	4.58	5	6.481	51	7.14	8	10.10	81	9	10	12.73
22	4.69	5	6.633	52	7.21	8	10.20	82	9.06	10	12.81
23	4.80	5	6.782	53	7.28	8	10.30	83	9.11	10	12.88
24	4.90	5	6.928	54	7.35	8	10.39	84	9.17	10	12.96
25	5	6	7.071	55	7.42	9	10.49	85	9.22	10	13.04
26	5.10	6	7.211	56	7.48	9	10.58	86	9.27	10	13.11
27	5.20	6	7.348	57	7.55	9	10.68	87	9.33	11	13.19
28	5.29	6	7.483	58	7.62	9	10.77	88	9.38	11	13.27
29	5.39	6	7.616	59	7.68	9	10.86	89	9.43	11	13.34
30	5.48	6	7.746	60	7.75	9	10.95	90	9.49	11	13.42
31	5.57	7	7.874	61	7.81	9	11.05	91	9.54	11	13.49
32	5.66	7	8	62	7.87	9	11.14	92	9.59	11	13.56
33	5.74	7	8.124	63	7.94	9	11.22	93	9.64	11	13.64
34	5.83	7	8.246	64	8	9	11.31	94	9.70	11	13.71
35	5.92	7	8.367	65	8.06	9	11.40	95	9.75	11	13.78
36	6	7	8.485	66	8.12	9	11.49	96	9.80	11	13.86
37	6.08	7	8.602	67	8.19	9	11.58	97	9.85	11	13.93
38	6.16	7	8.718	68	8.25	9	11.66	98	9.90	11	14
39	6.24	7	8.832	69	8.31	9	11.75	99	9.95	11	14.07
40	6.32	7	8.944	70	8.37	10	11.83	100	10	12	14.14

Table 4: Values of $\sqrt{n}$, $\zeta(n)$, and $\sqrt{2n}$ for $11 \leqslant n \leqslant 100$

A.4 Data for Lemma 4.4(1)

In Table 5, Δb refers to $b(n+1) - b(n)$.

n	$b(n)$	$n+1$	Δb	$2n+1$
8	52	9	9	17
9	61	10	16	19
10	77	11	18	21
11	95	12	14	23
12	109	13	21	25
13	130	14	23	27
14	153	15	25	29
15	178	16	27	31
16	205	17	18	33
17	223	18	30	35
18	253	19	32	37
19	285	20	34	39
20	319	21	36	41
21	355	22	38	43
22	393	23	40	45
23	433	24	42	47
24	475	25	28	49
25	503	26	45	51
26	548	27	47	53
27	595	28	49	55

n	$b(n)$	$n+1$	Δb	$2n+1$
28	644	29	51	57
29	695	30	53	59
30	748	31	31	61
31	779	32	56	63
32	835	33	58	65
33	893	34	60	67
34	953	35	62	69
35	1015	36	64	71
36	1079	37	66	73
37	1145	38	68	75
38	1213	39	70	77
39	1283	40	72	79
40	1355	41	74	81
41	1429	42	76	83
42	1505	43	78	85
43	1583	44	80	87
44	1663	45	82	89
45	1745	46	84	91
46	1829	47	56	93
47	1885	48	87	95

Table 5: Values of $n+1$, $\Delta b := b(n+1) - b(n)$, and $2n+1$ for $8 \leqslant n \leqslant 47$

A.5 Data for Lemma 4.4(3)

In Tables 6, 7, 8, and 9, $z_1 := \zeta(n)$.

n	z_1	$b(z_1)$	$\|n - b(z_1)\|$	$z_1 - 3$
191	15	178	13	12
192	16	205	13	13
193	16	205	12	13
194	16	205	11	13
195	16	205	10	13
196	16	205	9	13
197	16	205	8	13
198	16	205	7	13
199	16	205	6	13
200	16	205	5	13
201	16	205	4	13
202	16	205	3	13
203	16	205	2	13
204	16	205	1	13

n	z_1	$b(z_1)$	$\|n - b(z_1)\|$	$z_1 - 3$
205	16	205	0	13
206	16	205	1	13
207	16	205	2	13
208	16	205	3	13
209	17	223	14	14
210	17	223	13	14
211	17	223	12	14
212	17	223	11	14
213	17	223	10	14
214	17	223	9	14
215	17	223	8	14
216	17	223	7	14
217	17	223	6	14
218	17	223	5	14

Table 6: Values of z_1, $|n - b(z_1)|$, and $z_1 - 3$ for $191 \leqslant n \leqslant 218$

n	z_1	$b(z_1)$	$\lvert n-b(z_1)\rvert$	z_1-3
219	17	223	4	14
220	17	223	3	14
221	17	223	2	14
222	17	223	1	14
223	17	223	0	14
224	17	223	1	14
225	17	223	2	14
226	17	223	3	14
227	17	223	4	14
228	17	223	5	14
229	17	223	6	14
230	17	223	7	14
231	17	223	8	14
232	17	223	9	14
233	17	223	10	14
234	17	223	11	14
235	17	223	12	14
236	17	223	13	14
237	17	223	14	14
238	18	253	15	15
239	18	253	14	15
240	18	253	13	15
241	18	253	12	15
242	18	253	11	15
243	18	253	10	15
244	18	253	9	15
245	18	253	8	15
246	18	253	7	15
247	18	253	6	15
248	18	253	5	15
249	18	253	4	15
250	18	253	3	15
251	18	253	2	15
252	18	253	1	15
253	18	253	0	15
254	18	253	1	15
255	18	253	2	15
256	18	253	3	15
257	18	253	4	15
258	18	253	5	15
259	18	253	6	15
260	18	253	7	15
261	18	253	8	15

n	z_1	$b(z_1)$	$\lvert n-b(z_1)\rvert$	z_1-3
262	18	253	9	15
263	18	253	10	15
264	18	253	11	15
265	18	253	12	15
266	18	253	13	15
267	18	253	14	15
268	18	253	15	15
269	19	285	16	16
270	19	285	15	16
271	19	285	14	16
272	19	285	13	16
273	19	285	12	16
274	19	285	11	16
275	19	285	10	16
276	19	285	9	16
277	19	285	8	16
278	19	285	7	16
279	19	285	6	16
280	19	285	5	16
281	19	285	4	16
282	19	285	3	16
283	19	285	2	16
284	19	285	1	16
285	19	285	0	16
286	19	285	1	16
287	19	285	2	16
288	19	285	3	16
289	19	285	4	16
290	19	285	5	16
291	19	285	6	16
292	19	285	7	16
293	19	285	8	16
294	19	285	9	16
295	19	285	10	16
296	19	285	11	16
297	19	285	12	16
298	19	285	13	16
299	19	285	14	16
300	19	285	15	16
301	19	285	16	16
302	20	319	17	17
303	20	319	16	17
304	20	319	15	17

Table 7: Values of z_1, $|n-b(z_1)|$, and z_1-3 for $219 \leqslant n \leqslant 304$

n	z_1	$b(z_1)$	$\lvert n-b(z_1)\rvert$	z_1-3
305	20	319	14	17
306	20	319	13	17
307	20	319	12	17
308	20	319	11	17
309	20	319	10	17
310	20	319	9	17
311	20	319	8	17
312	20	319	7	17
313	20	319	6	17
314	20	319	5	17
315	20	319	4	17
316	20	319	3	17
317	20	319	2	17
318	20	319	1	17
319	20	319	0	17
320	20	319	1	17
321	20	319	2	17
322	20	319	3	17
323	20	319	4	17
324	20	319	5	17
325	20	319	6	17
326	20	319	7	17
327	20	319	8	17
328	20	319	9	17
329	20	319	10	17
330	20	319	11	17
331	20	319	12	17
332	20	319	13	17
333	20	319	14	17
334	20	319	15	17
335	20	319	16	17
336	20	319	17	17
337	21	355	18	18
338	21	355	17	18
339	21	355	16	18
340	21	355	15	18
341	21	355	14	18
342	21	355	13	18
343	21	355	12	18
344	21	355	11	18
345	21	355	10	18
346	21	355	9	18
347	21	355	8	18
348	21	355	7	18
349	21	355	6	18
350	21	355	5	18
351	21	355	4	18
352	21	355	3	18
353	21	355	2	18
354	21	355	1	18
355	21	355	0	18
356	21	355	1	18
357	21	355	2	18
358	21	355	3	18
359	21	355	4	18
360	21	355	5	18
361	21	355	6	18
362	21	355	7	18
363	21	355	8	18
364	21	355	9	18
365	21	355	10	18
366	21	355	11	18
367	21	355	12	18
368	21	355	13	18
369	21	355	14	18
370	21	355	15	18
371	21	355	16	18
372	21	355	17	18
373	21	355	18	18
374	22	393	19	19
375	22	393	18	19
376	22	393	17	19
377	22	393	16	19
378	22	393	15	19
379	22	393	14	19
380	22	393	13	19
381	22	393	12	19
382	22	393	11	19
383	22	393	10	19
384	22	393	9	19
385	22	393	8	19
386	22	393	7	19
387	22	393	6	19
388	22	393	5	19
389	22	393	4	19
390	22	393	3	19

Table 8: Values of z_1, $|n-b(z_1)|$, and z_1-3 for $305 \leqslant n \leqslant 390$

n	z_1	$b(z_1)$	$\|n-b(z_1)\|$	z_1-3	n	z_1	$b(z_1)$	$\|n-b(z_1)\|$	z_1-3
391	22	393	2	19	436	23	433	3	20
392	22	393	1	19	437	23	433	4	20
393	22	393	0	19	438	23	433	5	20
394	22	393	1	19	439	23	433	6	20
395	22	393	2	19	440	23	433	7	20
396	22	393	3	19	441	23	433	8	20
397	22	393	4	19	442	23	433	9	20
398	22	393	5	19	443	23	433	10	20
399	22	393	6	19	444	23	433	11	20
400	22	393	7	19	445	23	433	12	20
401	22	393	8	19	446	23	433	13	20
402	22	393	9	19	447	23	433	14	20
403	22	393	10	19	448	23	433	15	20
404	22	393	11	19	449	23	433	16	20
405	22	393	12	19	450	23	433	17	20
406	22	393	13	19	451	23	433	18	20
407	22	393	14	19	452	23	433	19	20
408	22	393	15	19	453	23	433	20	20
409	22	393	16	19	454	24	475	21	21
410	22	393	17	19	455	24	475	20	21
411	22	393	18	19	456	24	475	19	21
412	22	393	19	19	457	24	475	18	21
413	23	433	20	20	458	24	475	17	21
414	23	433	19	20	459	24	475	16	21
415	23	433	18	20	460	24	475	15	21
416	23	433	17	20	461	24	475	14	21
417	23	433	16	20	462	24	475	13	21
418	23	433	15	20	463	24	475	12	21
419	23	433	14	20	464	24	475	11	21
420	23	433	13	20	465	24	475	10	21
421	23	433	12	20	466	24	475	9	21
422	23	433	11	20	467	24	475	8	21
423	23	433	10	20	468	24	475	7	21
424	23	433	9	20	469	24	475	6	21
425	23	433	8	20	470	24	475	5	21
426	23	433	7	20	471	24	475	4	21
427	23	433	6	20	472	24	475	3	21
428	23	433	5	20	473	24	475	2	21
429	23	433	4	20	474	24	475	1	21
430	23	433	3	20	475	24	475	0	21
431	23	433	2	20	476	24	475	1	21
432	23	433	1	20	477	24	475	2	21
433	23	433	0	20	478	24	475	3	21
434	23	433	1	20	479	24	475	4	21
435	23	433	2	20	480	24	475	5	21

Table 9: Values of z_1, $|n-b(z_1)|$, and z_1-3 for $391 \leqslant n \leqslant 480$

A.6 Code for Theorem 4.6

For all $n \geqslant 3$, define $N := \lfloor \log_2 \log_5 n \rfloor + 1$ and

$$r(n) = \left(\sum_{j=1}^{N} \frac{1}{2^{j-1}} n^{\frac{1}{2^j}} \right) - \frac{2^{N-1} - 1}{2^{N-1}}$$

Theorem 4.6 asserts that $|\zeta(n) - r(n)| < 1.985$ for all $n \geqslant 3$. The first portion of the proof requires verifying this directly for all $3 \leqslant n < 3350194786$. Table 10 summarizes the relevant data for the exceptional cases $n = 5^{2^d}$ and $n = 5^{2^d} - 1$ for $1 \leqslant d \leqslant 3$.

n	$\zeta(n)$	$r(n)$	$\lvert\zeta(n) - r(n)\rvert$
$5^2 - 1$	5	4.89898	0.101021
5^2	6	5.61803	0.381966
$5^4 - 1$	28	26.979	1.02101

n	$\zeta(n)$	$r(n)$	$\lvert\zeta(n) - r(n)\rvert$
5^4	28	27.309	0.690983
$5^8 - 1$	639	637.999	1.00081
5^8	639	638.155	0.845492

Table 10: Values of $\zeta(n)$, $r(n)$, and errors for $n = 5^{2^d} - 1$, 5^{2^d} with $1 \leqslant d \leqslant 3$

For the remaining values of n, recall that if $n \in [\ell(m), \ell(m+1)-1]$, then $\zeta(n) = m$. Now, one may compute that $\ell(58006) = 3350194786$. This means that, rather than check every value of n less than 3350194786, it suffices to work with $n = \ell(k)$ and $n = \ell(k+1) - 1$ for $1 \leqslant k \leqslant 58005$.

The Mathematica code below computes the required values of $\ell(k)$ and $\ell(k+1)-1$, and also evaluates $r(\ell(k))$ and $r(\ell(k+1) - 1)$. The values of ℓ come from `LTable`, which can be produced by the code given for computing $b(n)$ in Appendix A.2.1.

```
BigN[n_] := Floor[Log[2, Log[5, n]]] + 1;
r[n_] := Sum[1/(2^(j - 1))*n^(1/(2^j)), {j, 1, BigN[n]}]
         - (2^(BigN[n] - 1) - 1)/2^(BigN[n] - 1);
nTestVals =
 Flatten[Table[{LTable[[k]], LTable[[k + 1]] - 1}, {k, 1, 58005}]];
rVals = Table[N[r[nTestVals[[k]]]], {k, 1, 2*58005}];
```

With the above code, it is not necessary to compute $\zeta(n)$ again in order to evaluate $|\zeta(n) - r(n)|$. This is because the test values for n are stored in `nTestVals` in the following order:

$$\ell(1), \ell(2) - 1, \ell(2), \ell(3) - 1, \ldots, \ell(m), \ell(m+1) - 1, \ldots, \ell(58005), \ell(58006) - 1$$

Applying ζ to each element of this list produces, in order:

$$1, 1, 2, 2, \ldots, m, m, \ldots, 58005, 58005.$$

In this list, the number at index i is equal to $\lceil \frac{i}{2} \rceil$. Thus, the error between $r(n)$ and $\zeta(n)$ can be found by comparing `rVals[[i]]` and `Ceiling[i/2]`. The command below calculates the maximum error that occurs in this range.

```
maxError =
   Max[Table[Abs[rVals[[i]] - Ceiling[i/2]], {i, 1, 2*58005}]]
```

This returns a (rounded) value of `1.45175`.

A.7 Data for Lemma 5.1

Note that if $n \in [\ell(m), \ell(m+1)-1]$, then $\zeta(n) = m$. So, to establish upper bounds on $z_1 := \zeta(n)$, it suffices to compare z_1 to $u(z_1) := \lfloor \sqrt{\ell(z_1)} \rfloor + \ell(z_1)^{\frac{1}{4}}$.

n	z_1	$\ell(z_1)$	$u(z_1)$	n	z_1	$\ell(z_1)$	$u(z_1)$
$[3,4]$	1	3	2.316	$[1622,1702]$	44	1622	46.35
$[5,6]$	2	5	3.495	$[1703,1785]$	45	1703	47.42
$[7,11]$	3	7	3.627	$[1786,1840]$	46	1786	48.50
$[12,16]$	4	12	4.861	$[1841,1926]$	47	1841	48.55
$[17,24]$	5	17	6.031	$[1927,2014]$	48	1927	49.63
$[25,30]$	6	25	7.236	$[2015,2104]$	49	2015	50.70
$[31,46]$	7	31	7.360	$[2105,2196]$	50	2105	51.77
$[47,54]$	8	47	8.618	$[2197,2290]$	51	2197	52.85
$[55,69]$	9	55	9.723	$[2291,2386]$	52	2291	53.92
$[70,86]$	10	70	10.89	$[2387,2484]$	53	2387	54.99
$[87,99]$	11	87	12.05	$[2485,2538]$	54	2485	56.06
$[100,119]$	12	100	13.16	$[2539,2639]$	55	2539	57.10
$[120,141]$	13	120	13.31	$[2640,2742]$	56	2640	58.17
$[142,165]$	14	142	14.45	$[2743,2847]$	57	2743	59.24
$[166,191]$	15	166	15.59	$[2848,2954]$	58	2848	60.31
$[192,208]$	16	192	16.72	$[2955,3063]$	59	2955	61.37
$[209,237]$	17	209	17.80	$[3064,3174]$	60	3064	62.44
$[238,268]$	18	238	18.93	$[3175,3287]$	61	3175	63.51
$[269,301]$	19	269	20.05	$[3288,3402]$	62	3288	64.57
$[302,336]$	20	302	21.17	$[3403,3519]$	63	3403	65.64
$[337,373]$	21	337	22.28	$[3520,3638]$	64	3520	66.70
$[374,412]$	22	374	23.40	$[3639,3759]$	65	3639	67.77
$[413,453]$	23	413	24.51	$[3760,3882]$	66	3760	68.83
$[454,480]$	24	454	25.62	$[3883,4007]$	67	3883	69.89
$[481,524]$	25	481	25.68	$[4008,4134]$	68	4008	70.96
$[525,570]$	26	525	26.79	$[4135,4209]$	69	4135	72.02
$[571,618]$	27	571	27.89	$[4210,4339]$	70	4210	72.06
$[619,668]$	28	619	28.99	$[4340,4471]$	71	4340	73.12
$[669,720]$	29	669	30.09	$[4472,4605]$	72	4472	74.18
$[721,750]$	30	721	31.18	$[4606,4741]$	73	4606	75.24
$[751,805]$	31	751	32.23	$[4742,4879]$	74	4742	76.30
$[806,862]$	32	806	33.33	$[4880,5019]$	75	4880	77.36
$[863,921]$	33	863	34.42	$[5020,5161]$	76	5020	78.42
$[922,982]$	34	922	35.51	$[5162,5305]$	77	5162	79.48
$[983,1045]$	35	983	36.60	$[5306,5451]$	78	5306	80.53
$[1046,1110]$	36	1046	37.69	$[5452,5599]$	79	5452	81.59
$[1111,1177]$	37	1111	38.77	$[5600,5749]$	80	5600	82.65
$[1178,1246]$	38	1178	39.86	$[5750,5901]$	81	5750	83.71
$[1247,1317]$	39	1247	40.94	$[5902,6055]$	82	5902	84.76
$[1318,1390]$	40	1318	42.03	$[6056,6211]$	83	6056	85.82
$[1391,1465]$	41	1391	43.11	$[6212,6369]$	84	6212	86.88
$[1466,1542]$	42	1466	44.19	$[6370,6529]$	85	6370	87.93
$[1543,1621]$	43	1543	45.27	$[6530,6622]$	86	6530	88.99

Table 11: Values of z_1, $\ell(z_1)$, and $u(z_1) := \lfloor \sqrt{\ell(z_1)} \rfloor + \ell(z_1)^{\frac{1}{4}}$ for $3 \leqslant n \leqslant 6622$

A.8 Data for Lemma 5.5 and Theorem 1.6

Define the following functions.

- For each $n \geqslant 3$, $h(n) = b(n) - (n-1) + \sum_{k=1}^{z_1} |W(k)|$
- For each real number $x \geqslant 0$, $\mathfrak{r}(x) = x^2 - \sum_{k=1}^{\infty} \left(\frac{2^k}{\prod_{i=1}^{k}(2^i+1)} \right) x^{1+\frac{1}{2^k}}$

In Tables 12, 13, 14, 15, and 16, for each value of n, e_1 represents $||W(n)| - h(n)|$ and e_2 represents $||W(n)| - \mathfrak{r}(n)|$.

n	$\|W(n)\|$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2	n	$\|W(n)\|$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2
3	6	6	0	4.26	1.74	53	2484	2529	45	2507.98	23.98
4	11	11	0	8.84	2.16	54	2582	2627	45	2606.64	24.64
5	17	19	2	15.15	1.85	55	2681	2740	59	2707.23	26.23
6	25	27	2	23.20	1.80	56	2783	2841	58	2809.75	26.75
7	34	39	5	33.03	0.97	57	2886	2944	58	2914.20	28.20
8	47	55	8	44.65	2.35	58	2993	3049	56	3020.58	27.58
9	59	63	4	58.06	0.94	59	3101	3156	55	3128.89	27.89
10	74	78	4	73.30	0.70	60	3211	3265	54	3239.13	28.13
11	91	95	4	90.36	0.64	61	3322	3376	54	3351.30	29.30
12	109	119	10	109.26	0.26	62	3435	3489	54	3465.40	30.40
13	129	139	10	130.00	1.0	63	3550	3604	54	3581.44	31.44
14	152	161	9	152.58	0.58	64	3667	3721	54	3699.41	32.41
15	176	185	9	177.02	1.02	65	3786	3840	54	3819.31	33.31
16	202	211	9	203.32	1.32	66	3908	3961	53	3941.14	33.14
17	229	245	16	231.48	2.48	67	4031	4084	53	4064.91	33.91
18	259	274	15	261.51	2.51	68	4156	4209	53	4190.62	34.62
19	290	305	15	293.41	3.41	69	4283	4336	53	4318.26	35.26
20	323	338	15	327.19	4.19	70	4411	4485	74	4447.84	36.84
21	358	373	15	362.85	4.85	71	4542	4615	73	4579.35	37.35
22	396	410	14	400.39	4.39	72	4674	4747	73	4712.80	38.80
23	435	449	14	439.81	4.81	73	4808	4881	73	4848.19	40.19
24	476	490	14	481.12	5.12	74	4946	5017	71	4985.52	39.52
25	518	542	24	524.32	6.32	75	5085	5155	70	5124.78	39.78
26	562	586	24	569.41	7.41	76	5226	5295	69	5265.99	39.99
27	609	632	23	616.40	7.40	77	5368	5437	69	5409.13	41.13
28	657	680	23	665.28	8.28	78	5512	5581	69	5554.21	42.21
29	707	730	23	716.06	9.06	79	5658	5727	69	5701.24	43.24
30	759	782	23	768.74	9.74	80	5806	5875	69	5850.20	44.20
31	812	846	34	823.32	11.32	81	5956	6025	69	6001.10	45.10
32	869	901	32	879.81	10.81	82	6108	6177	69	6153.95	45.95
33	927	958	31	938.20	11.20	83	6263	6331	68	6308.73	45.73
34	987	1017	30	998.50	11.50	84	6419	6487	68	6465.46	46.46
35	1048	1078	30	1060.70	12.70	85	6577	6645	68	6624.13	47.13
36	1111	1141	30	1124.81	13.81	86	6737	6805	68	6784.75	47.75
37	1176	1206	30	1190.84	14.84	87	6898	6989	91	6947.30	49.30
38	1244	1273	29	1258.78	14.78	88	7062	7152	90	7111.80	49.80
39	1313	1342	29	1328.63	15.63	89	7227	7317	90	7278.24	51.24
40	1384	1413	29	1400.39	16.39	90	7394	7484	90	7446.63	52.63
41	1457	1486	29	1474.07	17.07	91	7563	7653	90	7616.96	53.96
42	1532	1561	29	1549.67	17.67	92	7736	7824	88	7789.24	53.24
43	1609	1638	29	1627.18	18.18	93	7910	7997	87	7963.46	53.46
44	1689	1717	28	1706.61	17.61	94	8086	8172	86	8139.63	53.63
45	1770	1798	28	1787.96	17.96	95	8263	8349	86	8317.74	54.74
46	1853	1881	28	1871.23	18.23	96	8442	8528	86	8497.79	55.79
47	1937	1983	46	1956.43	19.43	97	8623	8709	86	8679.80	56.80
48	2023	2069	46	2043.54	20.54	98	8806	8892	86	8863.75	57.75
49	2111	2157	46	2132.58	21.58	99	8991	9077	86	9049.64	58.64
50	2201	2247	46	2223.54	22.54	100	9177	9287	110	9237.49	60.49
51	2294	2339	45	2316.43	22.43	101	9366	9475	109	9427.28	61.28
52	2388	2433	45	2411.24	23.24	102	9558	9665	107	9619.02	61.02

Table 12: Values of $|W(n)|$, $h(n)$, $\mathfrak{r}(n)$, and errors for $3 \leqslant n \leqslant 102$

n	$\|W(n)\|$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2	n	$\|W(n)\|$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2
103	9751	9857	106	9812.70	61.70	167	26153	26330	177	26269.71	116.71
104	9946	10051	105	10008.34	62.34	168	26473	26649	176	26590.44	117.44
105	10143	10247	104	10205.92	62.92	169	26794	26970	176	26913.12	119.12
106	10341	10445	104	10405.45	64.45	170	27118	27293	175	27237.76	119.76
107	10542	10645	103	10606.93	64.93	171	27445	27618	173	27564.36	119.36
108	10744	10847	103	10810.36	66.36	172	27773	27945	172	27892.92	119.92
109	10948	11051	103	11015.73	67.73	173	28103	28274	171	28223.44	120.44
110	11154	11257	103	11223.06	69.06	174	28435	28605	170	28555.92	120.92
111	11362	11465	103	11432.34	70.34	175	28768	28938	170	28890.36	122.36
112	11574	11675	101	11643.57	69.57	176	29104	29273	169	29226.76	122.76
113	11787	11887	100	11856.74	69.74	177	29441	29610	169	29565.12	124.12
114	12002	12101	99	12071.87	69.87	178	29780	29949	169	29905.45	125.45
115	12218	12317	99	12288.95	70.95	179	30121	30290	169	30247.73	126.73
116	12436	12535	99	12507.98	71.98	180	30464	30633	169	30591.97	127.97
117	12656	12755	99	12728.96	72.96	181	30809	30978	169	30938.18	129.18
118	12878	12977	99	12951.89	73.89	182	31156	31325	169	31286.34	130.34
119	13102	13201	99	13176.77	74.77	183	31505	31674	169	31636.47	131.47
120	13327	13457	130	13403.61	76.61	184	31858	32025	167	31988.56	130.56
121	13554	13684	130	13632.39	78.39	185	32212	32378	166	32342.61	130.61
122	13784	13913	129	13863.13	79.13	186	32568	32733	165	32698.62	130.62
123	14017	14144	127	14095.82	78.82	187	32925	33090	165	33056.59	131.59
124	14251	14377	126	14330.47	79.47	188	33284	33449	165	33416.53	132.53
125	14487	14612	125	14567.06	80.06	189	33645	33810	165	33778.42	133.42
126	14725	14849	124	14805.61	80.61	190	34008	34173	165	34142.28	134.28
127	14964	15088	124	15046.11	82.11	191	34373	34538	165	34508.10	135.10
128	15206	15329	123	15288.57	82.57	192	34739	34942	203	34875.88	136.88
129	15449	15572	123	15532.98	83.98	193	35107	35310	203	35245.63	138.63
130	15694	15817	123	15779.34	85.34	194	35478	35680	202	35617.33	139.33
131	15941	16064	123	16027.66	86.66	195	35850	36052	202	35991.00	141.00
132	16190	16313	123	16277.93	87.93	196	36224	36426	202	36366.63	142.63
133	16441	16564	123	16530.15	89.15	197	36601	36802	201	36744.23	143.23
134	16696	16817	121	16784.33	88.33	198	36981	37180	199	37123.79	142.79
135	16952	17072	120	17040.47	88.47	199	37362	37560	198	37505.31	143.31
136	17210	17329	119	17298.55	88.55	200	37745	37942	197	37888.79	143.79
137	17469	17588	119	17558.60	89.60	201	38130	38326	196	38274.23	144.23
138	17730	17849	119	17820.60	90.60	202	38516	38712	196	38661.64	145.64
139	17993	18112	119	18084.55	91.55	203	38905	39100	195	39051.02	146.02
140	18258	18377	119	18350.46	92.46	204	39295	39490	195	39442.35	147.35
141	18525	18644	119	18618.32	93.32	205	39687	39882	195	39835.65	148.65
142	18793	18946	153	18888.14	95.14	206	40081	40276	195	40230.91	149.91
143	19063	19216	153	19159.92	96.92	207	40477	40672	195	40628.14	151.14
144	19336	19488	152	19433.65	97.65	208	40875	41070	195	41027.33	152.33
145	19611	19762	151	19709.34	98.34	209	41274	41508	234	41428.48	154.48
146	19889	20038	149	19986.99	97.99	210	41676	41909	233	41831.60	155.60
147	20168	20316	148	20266.59	98.59	211	42080	42312	232	42236.68	156.68
148	20449	20596	147	20548.15	99.15	212	42488	42717	229	42643.72	155.72
149	20732	20878	146	20831.66	99.66	213	42897	43124	227	43052.73	155.73
150	21016	21162	146	21117.13	101.13	214	43308	43533	225	43463.70	155.70
151	21303	21448	145	21404.56	101.56	215	43720	43944	224	43876.64	156.64
152	21591	21736	145	21693.95	102.95	216	44134	44357	223	44291.54	157.54
153	21881	22026	145	21985.29	104.29	217	44550	44772	222	44708.41	158.41
154	22173	22318	145	22278.60	105.60	218	44968	45189	221	45127.24	159.24
155	22467	22612	145	22573.85	106.85	219	45388	45608	220	45548.03	160.03
156	22763	22908	145	22871.07	108.07	220	45809	46029	220	45970.79	161.79
157	23061	23206	145	23170.25	109.25	221	46232	46452	220	46395.52	163.52
158	23363	23506	143	23471.38	108.38	222	46658	46877	219	46822.21	164.21
159	23666	23808	142	23774.47	108.47	223	47085	47304	219	47250.86	165.86
160	23971	24112	141	24079.52	108.52	224	47514	47733	219	47681.48	167.48
161	24277	24418	141	24386.53	109.53	225	47945	48164	219	48114.06	169.06
162	24585	24726	141	24695.50	110.50	226	48379	48597	218	48548.61	169.61
163	24895	25036	141	25006.42	111.42	227	48816	49032	216	48985.13	169.13
164	25207	25348	141	25319.31	112.31	228	49254	49469	215	49423.61	169.61
165	25521	25662	141	25634.15	113.15	229	49694	49908	214	49864.05	170.05
166	25836	26013	177	25950.95	114.95	230	50136	50349	213	50306.47	170.47

Table 13: Values of $|W(n)|$, $h(n)$, $\mathfrak{r}(n)$, and errors for $103 \leqslant n \leqslant 230$

n	$\lvert W(n)\rvert$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2	n	$\lvert W(n)\rvert$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2
231	50579	50792	213	50750.84	171.84	295	83052	83325	273	83282.44	230.44
232	51025	51237	212	51197.18	172.18	296	83624	83896	272	83854.72	230.72
233	51472	51684	212	51645.49	173.49	297	84197	84469	272	84428.97	231.97
234	51921	52133	212	52095.77	174.77	298	84772	85044	272	85005.19	233.19
235	52372	52584	212	52548.00	176.00	299	85349	85621	272	85583.38	234.38
236	52825	53037	212	53002.21	177.21	300	85928	86200	272	86163.54	235.54
237	53280	53492	212	53458.38	178.38	301	86509	86781	272	86745.67	236.67
238	53736	54000	264	53916.52	180.52	302	87091	87419	328	87329.77	238.77
239	54195	54458	263	54376.62	181.62	303	87676	88003	327	87915.84	239.84
240	54656	54918	262	54838.69	182.69	304	88262	88589	327	88503.88	241.88
241	55119	55380	261	55302.73	183.73	305	88850	89177	327	89093.90	243.90
242	55586	55844	258	55768.73	182.73	306	89441	89767	326	89685.88	244.88
243	56054	56310	256	56236.70	182.70	307	90034	90359	325	90279.83	245.83
244	56524	56778	254	56706.63	182.63	308	90631	90953	322	90875.75	244.75
245	56995	57248	253	57178.54	183.54	309	91229	91549	320	91473.64	244.64
246	57468	57720	252	57652.40	184.40	310	91829	92147	318	92073.50	244.50
247	57943	58194	251	58128.24	185.24	311	92430	92747	317	92675.33	245.33
248	58420	58670	250	58606.04	186.04	312	93033	93349	316	93279.13	246.13
249	58899	59148	249	59085.81	186.81	313	93638	93953	315	93884.91	246.91
250	59379	59628	249	59567.55	188.55	314	94245	94559	314	94492.65	247.65
251	59861	60110	249	60051.25	190.25	315	94854	95167	313	95102.37	248.37
252	60346	60594	248	60536.92	190.92	316	95464	95777	313	95714.05	250.05
253	60832	61080	248	61024.56	192.56	317	96076	96389	313	96327.71	251.71
254	61320	61568	248	61514.16	194.16	318	96691	97003	312	96943.33	252.33
255	61810	62058	248	62005.73	195.73	319	97307	97619	312	97560.93	253.93
256	62302	62550	248	62499.27	197.27	320	97925	98237	312	98180.50	255.50
257	62797	63044	247	62994.78	197.78	321	98545	98857	312	98802.04	257.04
258	63295	63540	245	63492.25	197.25	322	99167	99479	312	99425.55	258.55
259	63794	64038	244	63991.69	197.69	323	99791	100103	312	100051.03	260.03
260	64295	64538	243	64493.10	198.10	324	100417	100729	312	100678.48	261.48
261	64798	65040	242	64996.47	198.47	325	101046	101357	311	101307.90	261.90
262	65302	65544	242	65501.82	199.82	326	101678	101987	309	101939.29	261.29
263	65809	66050	241	66009.13	200.13	327	102311	102619	308	102572.66	261.66
264	66317	66558	241	66518.41	201.41	328	102946	103253	307	103208.00	262.00
265	66827	67068	241	67029.66	202.66	329	103583	103889	306	103845.30	262.30
266	67339	67580	241	67542.87	203.87	330	104221	104527	306	104484.58	263.58
267	67853	68094	241	68058.06	205.06	331	104862	105167	305	105125.83	263.83
268	68369	68610	241	68575.21	206.21	332	105504	105809	305	105769.05	265.05
269	68886	69181	295	69094.33	208.33	333	106148	106453	305	106414.25	266.25
270	69406	69700	294	69615.41	209.41	334	106794	107099	305	107061.41	267.41
271	69927	70221	294	70138.47	211.47	335	107442	107747	305	107710.55	268.55
272	70451	70744	293	70663.49	212.49	336	108092	108397	305	108361.66	269.66
273	70977	71269	292	71190.49	213.49	337	108743	109106	363	109014.74	271.74
274	71507	71796	289	71719.45	212.45	338	109397	109759	362	109669.79	272.79
275	72038	72325	287	72250.38	212.38	339	110052	110414	362	110326.81	274.81
276	72571	72856	285	72783.28	212.28	340	110709	111071	362	110985.80	276.80
277	73105	73389	284	73318.14	213.14	341	111368	111730	362	111646.77	278.77
278	73641	73924	283	73854.98	213.98	342	112030	112391	361	112309.71	279.71
279	74179	74461	282	74393.78	214.78	343	112694	113054	360	112974.62	280.62
280	74719	75000	281	74934.56	215.56	344	113362	113719	357	113641.50	279.50
281	75261	75541	280	75477.30	216.30	345	114031	114386	355	114310.35	279.35
282	75804	76084	280	76022.01	218.01	346	114702	115055	353	114981.18	279.18
283	76349	76629	280	76568.69	219.69	347	115374	115726	352	115653.98	279.98
284	76897	77176	279	77117.34	220.34	348	116048	116399	351	116328.74	280.74
285	77446	77725	279	77667.95	221.95	349	116724	117074	350	117005.49	281.49
286	77997	78276	279	78220.54	223.54	350	117402	117751	349	117684.20	282.20
287	78550	78829	279	78775.10	225.10	351	118082	118430	348	118364.89	282.89
288	79105	79384	279	79331.62	226.62	352	118763	119111	348	119047.55	284.55
289	79662	79941	279	79890.12	228.12	353	119446	119794	348	119732.18	286.18
290	80222	80500	278	80450.58	228.58	354	120132	120479	347	120418.78	286.78
291	80785	81061	276	81013.01	228.01	355	120819	121166	347	121107.35	288.35
292	81349	81624	275	81577.42	228.42	356	121508	121855	347	121797.90	289.90
293	81915	82189	274	82143.79	228.79	357	122199	122546	347	122490.42	291.42
294	82483	82756	273	82712.13	229.13	358	122892	123239	347	123184.91	292.91

Table 14: Values of $|W(n)|$, $h(n)$, $\mathfrak{r}(n)$, and errors for $231 \leqslant n \leqslant 358$

n	$\|W(n)\|$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2	n	$\|W(n)\|$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2
359	123587	123934	347	123881.38	294.38	423	172206	172637	431	172559.65	353.65
360	124284	124631	347	124579.82	295.82	424	173031	173460	429	173384.41	353.41
361	124983	125330	347	125280.23	297.23	425	173857	174285	428	174211.14	354.14
362	125685	126031	346	125982.61	297.61	426	174685	175112	427	175039.84	354.84
363	126390	126734	344	126686.97	296.97	427	175515	175941	426	175870.52	355.52
364	127096	127439	343	127393.29	297.29	428	176347	176772	425	176703.18	356.18
365	127804	128146	342	128101.60	297.60	429	177181	177605	424	177537.81	356.81
366	128514	128855	341	128811.87	297.87	430	178016	178440	424	178374.41	358.41
367	129225	129566	341	129524.12	299.12	431	178853	179277	424	179212.99	359.99
368	129939	130279	340	130238.34	299.34	432	179693	180116	423	180053.55	360.55
369	130654	130994	340	130954.53	300.53	433	180534	180957	423	180896.07	362.07
370	131371	131711	340	131672.69	301.69	434	181377	181800	423	181740.58	363.58
371	132090	132430	340	132392.83	302.83	435	182222	182645	423	182587.06	365.06
372	132811	133151	340	133114.94	303.94	436	183069	183492	423	183435.51	366.51
373	133534	133874	340	133839.03	305.03	437	183918	184341	423	184285.94	367.94
374	134258	134659	401	134565.09	307.09	438	184769	185192	423	185138.35	369.35
375	134985	135385	400	135293.12	308.12	439	185622	186045	423	185992.73	370.73
376	135713	136113	400	136023.12	310.12	440	186477	186900	423	186849.08	372.08
377	136443	136843	400	136755.10	312.10	441	187334	187757	423	187707.42	373.42
378	137175	137575	400	137489.05	314.05	442	188194	188616	422	188567.72	373.72
379	137910	138309	399	138224.97	314.97	443	189057	189477	420	189430.00	373.00
380	138647	139045	398	138962.87	315.87	444	189921	190340	419	190294.26	373.26
381	139386	139783	397	139702.74	316.74	445	190787	191205	418	191160.49	373.49
382	140129	140523	394	140444.58	315.58	446	191655	192072	417	192028.70	373.70
383	140873	141265	392	141188.40	315.40	447	192524	192941	417	192898.88	374.88
384	141619	142009	390	141934.19	315.19	448	193396	193812	416	193771.04	375.04
385	142366	142755	389	142681.95	315.95	449	194269	194685	416	194645.17	376.17
386	143115	143503	388	143431.69	316.69	450	195144	195560	416	195521.28	377.28
387	143866	144253	387	144183.40	317.40	451	196021	196437	416	196399.37	378.37
388	144619	145005	386	144937.08	318.08	452	196900	197316	416	197279.43	379.43
389	145374	145759	385	145692.74	318.74	453	197781	198197	416	198161.46	380.46
390	146130	146515	385	146450.37	320.37	454	198663	199144	481	199045.48	382.48
391	146888	147273	385	147209.98	321.98	455	199548	200028	480	199931.46	383.46
392	147649	148033	384	147971.56	322.56	456	200434	200914	480	200819.43	385.43
393	148411	148795	384	148735.11	324.11	457	201322	201802	480	201709.36	387.36
394	149175	149559	384	149500.64	325.64	458	202212	202692	480	202601.28	389.28
395	149941	150325	384	150268.14	327.14	459	203105	203584	479	203495.17	390.17
396	150709	151093	384	151037.61	328.61	460	203999	204478	479	204391.04	392.04
397	151479	151863	384	151809.06	330.06	461	204895	205374	479	205288.88	393.88
398	152251	152635	384	152582.48	331.48	462	205794	206272	478	206188.69	394.69
399	153025	153409	384	153357.88	332.88	463	206695	207172	477	207090.49	395.49
400	153801	154185	384	154135.25	334.25	464	207600	208074	474	207994.26	394.26
401	154580	154963	383	154914.59	334.59	465	208506	208978	472	208900.00	394.00
402	155362	155743	381	155695.91	333.91	466	209414	209884	470	209807.72	393.72
403	156145	156525	380	156479.20	334.20	467	210323	210792	469	210717.42	394.42
404	156930	157309	379	157264.47	334.47	468	211234	211702	468	211629.09	395.09
405	157717	158095	378	158051.71	334.71	469	212147	212614	467	212542.74	395.74
406	158505	158883	378	158840.92	335.92	470	213062	213528	466	213458.37	396.37
407	159296	159673	377	159632.11	336.11	471	213979	214444	465	214375.97	396.97
408	160088	160465	377	160425.27	337.27	472	214897	215362	465	215295.55	398.55
409	160882	161259	377	161220.41	338.41	473	215817	216282	465	216217.10	400.10
410	161678	162055	377	162017.52	339.52	474	216740	217204	464	217140.63	400.63
411	162476	162853	377	162816.61	340.61	475	217664	218128	464	218066.14	402.14
412	163276	163653	377	163617.67	341.67	476	218590	219054	464	218993.62	403.62
413	164077	164517	440	164420.70	343.70	477	219518	219982	464	219923.08	405.08
414	164881	165320	439	165225.71	344.71	478	220448	220912	464	220854.51	406.51
415	165686	166125	439	166032.70	346.70	479	221380	221844	464	221787.92	407.92
416	166493	166932	439	166841.66	348.66	480	222314	222778	464	222723.31	409.31
417	167302	167741	439	167652.59	350.59	481	223249	223779	530	223660.67	411.67
418	168114	168552	438	168465.50	351.50	482	224187	224716	529	224600.01	413.01
419	168927	169365	438	169280.38	353.38	483	225127	225655	528	225541.33	414.33
420	169743	170180	437	170097.23	354.23	484	226069	226596	527	226484.62	415.62
421	170561	170997	436	170916.07	355.07	485	227014	227539	525	227429.89	415.89
422	171383	171816	433	171736.87	353.87	486	227962	228484	522	228377.13	415.13

Table 15: Values of $|W(n)|$, $h(n)$, $\mathfrak{r}(n)$, and errors for $359 \leqslant n \leqslant 486$

n	$\lvert W(n)\rvert$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2	n	$\lvert W(n)\rvert$	$h(n)$	e_1	$\mathfrak{r}(n)$	e_2
487	228911	229431	520	229326.36	415.36	499	240444	240951	507	240871.19	427.19
488	229862	230380	518	230277.55	415.55	500	241417	241924	507	241846.11	429.11
489	230815	231331	516	231230.73	415.73	501	242392	242899	507	242823.01	431.01
490	231769	232284	515	232185.88	416.88	502	243370	243876	506	243801.88	431.88
491	232726	233239	513	233143.01	417.01	503	244349	244855	506	244782.73	433.73
492	233684	234196	512	234102.11	418.11	504	245330	245836	506	245765.55	435.55
493	234644	235155	511	235063.19	419.19	505	246313	246819	506	246750.35	437.35
494	235606	236116	510	236026.25	420.25	506	247299	247804	505	247737.13	438.13
495	236570	237079	509	236991.29	421.29	507	248287	248791	504	248725.89	438.89
496	237536	238044	508	237958.30	422.30	508	249279	249780	501	249716.62	437.62
497	238503	239011	508	238927.29	424.29	509	250272	250771	499	250709.33	437.33
498	239473	239980	507	239898.25	425.25	510	251267	251764	497	251704.02	437.02

Table 16: Values of $|W(n)|$, $h(n)$, $\mathfrak{r}(n)$, and errors for $487 \leqslant n \leqslant 510$

References

[1] P. Alexandroff, Diskrete Räume, *Rec. Math. [Mat. Sbornik] (N.S.)* 2 (1937), 501–519.

[2] J. Bang-Jensen and G. Gutin, *Digraphs. Theory, algorithms, and applications* (Second edition). Springer Monographs in Mathematics. Springer-Verlag London, Ltd., 2009.

[3] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, American Elsevier Publishing Co., Inc., New York (1976).

[4] M. Erné, On the cardinalities of finite topologies and the number of antichains in partially ordered sets, *Discrete Math.* 35 (1981), 119–133.

[5] R. Parchmann, On the cardinalities of finite topologies, *Discrete Math.* 11 (1975), 161–172.

[6] A. R. Rao, The number of reachable pairs in a digraph, *Discrete Math.* 306 (2006), no. 14, 1595–1600.

[7] K. Ragnarsson and B. E. Tenner, Obtainable sizes of topologies of finite sets, *J. Combin. Theory Ser. A* 117 (2010), 138–151.

[8] H. Sharp, Jr., Quasi-orderings and topologies on finite sets, *Proc. Amer. Math. Soc.* 17 (1966), 1344–1349.

[9] S. S. Skiena, *The Algorithm Design Manual* (Second edition), Springer (2011).

[10] R. P. Stanley, On the number of open sets of finite topologies, *J. Combin. Theory* 10 (1971), 74–79.

[11] Wolfram Research, Inc., Mathematica, Version 12.0, Champaign, IL (2019).

(Received 26 July 2019; revised 19 Mar 2020)