



## ON THE COMBINATORICS OF PLACING BALLS INTO ORDERED BINS

**Vedant Bonde**

*Cluster Innovation Centre, University of Delhi, New Delhi*  
vedantbonde19@gmail.com

**Joshua M. Siktar**

*227 Ayres Hall, 1403 Circle Drive, Knoxville, Tennessee*  
jsiktar@vols.utk.edu

*Received: 10/18/20, Accepted: 7/22/21, Published: 8/3/21*

### Abstract

In this paper, we use techniques of enumerative combinatorics to study the following problem: we count the number of ways to split  $n$  balls into nonempty, ordered bins so that the most crowded bin has exactly  $k$  balls. We find closed forms for three of the different cases that can arise:  $k > \frac{n}{2}$ ,  $k = \frac{n}{2}$ , and when there exists  $j < k$  such that  $n = 2k + j$ . As an immediate result of our proofs, we find a closed form for the number of positive integer solutions to  $x_1 + x_2 + \cdots + x_\ell = n$  with the maximum of  $\{x_1, x_2, \dots, x_\ell\}$  being equal to  $k$ , when  $n$  and  $k$  have one of the aforementioned algebraic relationships. The problem is generalized to find a formula that enumerates the total number of ways without specific conditions on  $n, \ell, k$ . Subsequently, various additional identities and estimates related to this enumeration are proven and interpreted.

### 1. Introduction

One of the most elementary, well-known enumerative combinatorics problems asks the following basic question.

**Question 1** (Balls into Bins). How many ways can we split  $n$  balls into  $\ell$  nonempty ordered bins?

This is better known as the “stars-and-bars” problem and frequents itself in combinatorics textbooks such as [4, 27, 43]. The well-known formula that answers this question is  $\binom{n-1}{\ell-1}$ . However, there are many different restrictions on the contents of the bins that can increase the difficulty of the problem at hand. Here is one example of such a question.

**Question 2** (Balls into Bins with Minimum Capacity). How many ways can we split  $n$  balls into  $\ell$  nonempty ordered bins so that each bin has at least  $t$  balls?

It turns out this is merely a generalization of Question 1 where  $t$  can have a value other than 1. The closed form for this problem is also well-known and is  $\binom{n-(t-1)\ell-1}{\ell-1}$ . In this paper, we will address a question that sounds very similar to Question 2 but is actually far more complicated.

**Question 3** (Balls into Bins with Maximum Capacity). How many ways can we split  $n$  balls into any number of nonempty ordered bins where the most crowded bin has exactly  $k$  balls?

A natural variation of Question 3 simply involves fixing the total number of bins used.

**Question 4** (Generalized Balls into Bins with restrictions problem). Let  $n, \ell, k \in \mathbb{N}^+$ . This problem asks how many ways we can split  $n$  balls into  $\ell$  nonempty bins such that the most crowded bin has exactly  $k$  balls.

Question 4 aims to enumerate all the possible combinations of balls into bins with the restriction on the maximum number of balls in a bin. The condition of the most crowded bin having exactly  $k$  balls still holds, but with the added fact that we are also given the exact number of nonempty bins to be filled.

**Definition 1** (Generalized Bins Restriction Problem). Let  $n, k, \ell \in \mathbb{N}^+$ . The *Generalized Bins Restriction Problem* aims to find the number of ways to split  $n$  balls into  $\ell$  nonempty bins such that the maximum number of balls in each bin is at most  $k$ .

Throughout the paper we will let  $B_{n,k}$  denote the answer to Question 3 for chosen values of  $n$  and  $k$ . Along with this,  $M_{n,\ell,k}$  will denote the number of ways to split  $n$  balls into exactly  $\ell$  nonempty bins where the most crowded bin has exactly  $k$  balls (the answer to Question 4). We will also denote the quantity in Definition 1 as  $R_{n,\ell,k}$ . The Principle of Inclusion and Exclusion will be crucial for interpreting this class of problems for the following reason: if the maximum number of balls in a bin is *exactly*  $k$ , then *at least* one bin will have  $k$  balls in it.

Our approach to answering Question 3 revolves around considering two cases separately: the case where most of the balls are in a single bin, and the case where the balls are, loosely speaking, “more spread out.” We formalize this notion of “spread out” with the following definition, which is utilized in many lemma and theorem statements throughout the paper.

**Definition 2.** A configuration of  $n$  balls in bins has a *dominant bin* if the most crowded bin has exactly  $k$  balls, where  $\frac{n}{2} < k < n$ . Otherwise we say the configuration has no dominant bin.

Now we proceed to survey the literature to review related problems. Binomial coefficients have surfaced in many problems of an enumerative nature, including combinatorial inequalities, lattice walks, and Stirling numbers [1, 3, 5, 6, 10, 11, 12, 13, 14, 16, 17, 20, 24, 29, 32, 35, 39, 41, 38, 42]. There are also numerous applications of binomial coefficient identities to number theory and computer science. One prominent application is that of the Bernoulli trial, where we flip a [weighted] coin in succession many times and track the number of consecutive heads. Long-time asymptotic behavior of Bernoulli trials is explored in [8, 9, 18, 19, 21, 22, 23, 26, 33]. Furthermore, Lucas polynomials have played a role in calculating the number of ways to place the numbers  $\{1, 2, \dots, n\}$  on a circle and find  $r$  adjacent numbers on the circle where no  $k$  of them are consecutive, and this problem is closely related to the aforementioned Bernoulli trial problem (see [8, 9]). The main difference with our problem is that the bins we place balls into are in a line, rather than a circular formation.

We now state the four main results of this paper. The first three are all closed forms for  $B_{n,k}$  for different algebraic relationships between  $n$  and  $k$ . The fourth one is a summation formula for  $M_{n,\ell,k}$ . One remarkable attribute of our proofs is that we do not utilize generating functions in any way, and instead resort to more elementary binomial coefficient manipulations.

**Theorem 1** (Closed form: dominant bin). *If  $n, k \in \mathbb{N}^+$  with  $\frac{n}{2} < k < n$  then*

$$B_{n,k} = (n - k + 3)2^{n-k-2}. \tag{1.1}$$

**Theorem 2** (Closed form for  $B_{2k,k}$ ). *If  $k \in \mathbb{N}^+$  then*

$$B_{2k,k} = (k + 3)2^{k-2} - 1. \tag{1.2}$$

**Theorem 3** (Closed form for  $B_{2k+j,k}$ ). *Let  $j, k \in \mathbb{N}^+$  with  $k > j$ . Then the number of ways to split  $2k + j$  balls into nonempty bins so the most crowded bin has exactly  $k$  balls is*

$$B_{2k+j,k} = (k + j + 3)2^{k+j-2} - (3j^2 + 19j + 18)2^{j-4}. \tag{1.3}$$

**Theorem 4** (Formula for Generalized Balls into Bins with restrictions problem). *Suppose  $n, k, \ell \in \mathbb{N}^+$  such that  $\ell + k - 1 \leq n \leq \ell k$ . Then the following identity for  $M_{n,\ell,k}$  holds:*

$$M_{n,\ell,k} = \sum_{t=0}^{\ell} (-1)^t \binom{\ell}{t} \left[ \binom{n - tk - 1}{\ell - 1} - \binom{n - t(k - 1) - 1}{\ell - 1} \right]. \tag{1.4}$$

Upon inspection of the conditions for these results, one may notice they are not exhaustive of all possible values of  $n, k, \ell \in \mathbb{N}^+$ . Aside from the trivial case where  $n \leq k$ , these formulas do not give a general formula for  $B_{mk+j,k}$  when  $m \geq 3$ . We focus on the  $m = 2$  case (Equation (1.3)) because it most succinctly showcases the

strategy for counting the configurations of bins when there is no dominant bin. This formula is derived from doing casework on the number of bins that contain  $k$  balls, and the calculations are far more tractable when  $m = 2$ .

The remainder of this paper is organized as follows: in Section 2 we will prove (1.1), and in Section 3 we will prove (1.2) and (1.3). In both of these sections we prove the given closed forms first by fixing the number of bins used and then later removing that restriction. Section 4 provides generalizations of the aforementioned cases where we can find suitable summation formulas but not closed forms. This strategy enables us to establish a connection to a well-studied counting problem for solutions to certain integer equations in Section 5. The conclusion, Section 6, provides some finishing remarks and possible directions for future research.

## 2. Balls into Bins with a Dominant Bin

This section will be devoted to studying the balls in bins problem (Question 3) when there is a dominant cluster; that is, we count how many ways can we sort  $n$  balls into bins when one bin has more than half the total number of balls. It turns out this is the most straightforward of the three cases, for the following reason: if the most crowded bin, henceforth called the *dominant bin*, has  $k$  balls, then no other bin can have  $k$  or more balls. This simplifies the combinatorial analysis to come.

The first step will be to find a formula for  $M_{n,\ell,k}$  when  $\frac{n}{2} < k < n$ . From here we will find a summation formula for  $B_{n,k}$ , and then find a closed form for the sum. We quickly remark that the smallest possible number of bins to be used is 2 (one with  $k$  balls, and the other with  $n - k$  balls), and the largest possible number of bins is  $n - k + 1$  (one with  $k$  balls, and  $n - k$  bins each with 1 ball). Recall that  $B_{n,k}$  denotes the number of ways to split  $n$  balls into any number of ordered nonempty bins where the most crowded bin has  $k$  balls. Furthermore,  $M_{n,\ell,k}$  denotes the number of ways to split  $n$  balls into exactly  $\ell$  nonempty bins where the largest group is of size  $k$ .

**Lemma 1.** *Let  $n, \ell, k \in \mathbb{N}^+$  such that  $\frac{n}{2} < k < n$  and  $2 \leq \ell \leq n - k + 1$ . Then the number of ways to split  $n$  balls into  $\ell$  nonempty bins where the most crowded bin has exactly  $k$  balls is*

$$M_{n,\ell,k} = \ell \binom{n - k - 1}{\ell - 2}. \tag{2.1}$$

*Proof.* If the largest group is of size  $k$ , then our task reduces to splitting  $n - k$  balls into nonempty groups of size at most  $n - k$ . Since  $\frac{n}{2} < k$ , this means  $n - k < k$ , and that all of the remaining groups have size smaller than  $k$ . In other words, we need not impose further restrictions when breaking the  $n - k$  remaining balls into groups. This reasoning also illustrates that our proof does not generalize to the case where  $0 < k \leq \frac{n}{2}$ .

We evaluate the given quantity by splitting the  $n - k$  remaining balls into  $\ell - 1$  bins. By the classical stars-and-bars argument depicted in [28], this step can be done in  $\binom{n-k-1}{\ell-2}$  ways. Finally, we must insert the bin of size  $k$  in between the  $\ell - 1$  groups already established, and this can be done in  $\ell$  ways; the bin of size  $k$  can go between two of the other bins, or it can be put at either end of the line of bins. Multiplying the number of ways to perform these two steps together yields the desired result.  $\square$

Now we derive a closed form for  $B_{n,k}$  when there is a dominant bin.

**Theorem 5** (Closed form: dominant bin). *If  $n, k \in \mathbb{N}^+$  with  $\frac{n}{2} < k < n$  then*

$$B_{n,k} = (n - k + 3)2^{n-k-2}. \tag{2.2}$$

*Proof.* We first obtain a summation formula for  $B_{n,k}$  by summing Equation (2.1) over  $2 \leq \ell \leq n - k + 1$ :

$$B_{n,k} = \sum_{\ell=2}^{n-k+1} \ell \binom{n-k-1}{\ell-2}. \tag{2.3}$$

Now we can verify the proposed closed form for  $B_{n,k}$ . Let  $t := n - k$ , because we will be using  $t$  as a parameter. Then

$$B_{n,k} = \sum_{\ell=1}^t (\ell + 1) \binom{t-1}{\ell-1}. \tag{2.4}$$

It remains to evaluate the sum in Equation (2.4), which fortunately is relatively easy once we realize that the sum resembles the formula (A.1) with different variable labels. In particular, that formula gives us

$$\sum_{\ell=0}^{t-1} \ell \binom{t-1}{\ell} = (t-1)2^{t-2}. \tag{2.5}$$

Shifting the index of the sum in Equation (2.5) and adding  $2 \sum_{\ell=1}^t \binom{t-1}{\ell-1}$  to both sides of the equation gives

$$\sum_{\ell=1}^t (\ell + 1) \binom{t-1}{\ell-1} = (t-1)2^{t-2} + 2 \sum_{\ell=1}^t \binom{t-1}{\ell-1}. \tag{2.6}$$

Upon shifting the index of the sum on the right-hand side of Equation (2.6) by 1 we notice it equals  $2^{t-1}$ , and so

$$B_{n,k} = (t + 3)2^{t-2}. \tag{2.7}$$

The desired result follows upon substituting  $n - k$  back in place of  $t$ .  $\square$

**Remark 1.** This lemma does not hold for  $n = k$  because the second step becomes degenerate if all of the balls are in a single group of size  $k$ . It is obvious that  $B_{k,k} = 1$  and we will henceforth ignore this case.

**Remark 2.** It makes heuristic sense that this formula depends on  $n - k$  but not on  $n$  nor  $k$  individually. This is because the first step in our proof reduces the group of balls to a group of size  $n - k$ .

### 3. Balls in Bins without a Dominant Bin

In Section 2, our focus was on the case where most (more than half) of all the balls were placed in the same bin. In this section we instead assume the sunk pool balls are more spread out, and it turns out most of the theory developed for the dominant cluster case is no longer valid.

We will begin with a subsection devoted to the special case where  $n = 2k$ . This situation highlights the main distinction between the behavior of the case where there is a dominant bin and when there is not, while still having a closed form very similar to that of the dominant bin case described in Section 2. Afterward, we will derive a closed form for  $B_{2k+j,k}$  when  $0 < j < k$ , which will involve a combination of several sums that, while manageable with purely elementary methods, is non-trivial. The general approach in both subsections will be to derive a formula for  $M_{2k+j,\ell,k}$  and use that to derive a formula for  $B_{2k+j,k}$ ; in the first subsection  $j$  will equal zero, and in the second subsection  $j$  will be positive but less than  $k$ .

#### 3.1. Formula for $B_{2k,k}$

**Lemma 2.** *If  $k \in \mathbb{N}^+$  then the number of ways to split  $2k$  balls into  $\ell$  nonempty bins where  $2 \leq \ell \leq k + 1$  and the most crowded bin has exactly  $k$  balls is*

$$M_{2k,\ell,k} = \begin{cases} 1, & \ell = 2 \\ \ell \binom{k-1}{\ell-2}, & 3 \leq \ell \leq k + 1 \end{cases} \quad (3.1)$$

*Proof.* In the special case that  $\ell = 2$  there is only one possible configuration: two bins, each containing exactly  $k$  balls. The existence of this special case is contrary to our derivation of a formula for  $B_{2k+j,k}$  for  $0 < j < k$  as we will need at least 3 bins when  $j > 0$ . On the other hand, if there are at least three bins we cannot possibly have two bins each containing  $k$  balls.

Now we assume that  $\ell \geq 3$ . Since the largest possible number of balls in a bin is  $k$ , and we only have one bin with  $k$  balls, each of these additional bins can have at most  $k - 1$  balls. Moreover, we cannot have more than  $k$  bins besides the one containing exactly  $k$  balls (the “greedy” way to do this is to put  $k$  balls in the first bin and a single ball in each bin thereafter). Thus once we fix the number of bins,

we want to count the number of ways to split  $k$  balls between  $\ell - 1$  bins. By the classical stars-and-bars argument detailed in [28], this can be done in  $\binom{k-1}{\ell-2}$  ways.

Finally, we must choose where to insert the bin with  $k$  balls amongst the other  $\ell$  bins. Since the bins are ordered we can do this in  $\ell$  ways; multiplying the number of ways to perform each of these two steps yields the desired result.  $\square$

The closed form for  $B_{2k,k}$  now follows readily.

**Theorem 6** (Closed Form for  $B_{2k,k}$ ). *If  $k \in \mathbb{N}^+$  then*

$$B_{2k,k} = (k + 3)2^{k-2} - 1. \tag{3.2}$$

*Proof.* We can easily deduce a summation formula for  $B_{2k,k}$  by summing the formula (3.1) over all  $2 \leq \ell \leq k + 1$ . We obtain

$$B_{2k,k} = 1 + \sum_{\ell=3}^{k+1} \ell \binom{k-1}{\ell-2} = 1 + \sum_{\ell=1}^{k-1} (\ell + 2) \binom{k-1}{\ell}. \tag{3.3}$$

The evaluation of this sum follows from (2.5) and (2.6); then the desired result is immediate.<sup>1</sup>  $\square$

### 3.2. Formula for $B_{2k+j,k}$

In this subsection, we will derive a closed form for  $B_{2k+j,k}$  when  $k > j$ . The strategy is much more involved than the one used in Section 3.1, and proceeds as follows. We will see that a valid configuration of balls will have either one or two bins with exactly  $k$  balls; thus we calculate the number of ways to split  $2k + j$  balls into a fixed number of bins so that there is at least one bin with  $k$  balls, and then when there are two bins with  $k$  balls each. However, it is possible to have a configuration where one bin has  $k$  balls and another bin has  $k + i$  balls, for some  $1 \leq i \leq j$ ; we do not want to include these configurations in our final formula because the largest bin has more than  $k$  balls in this case. The process of adding and subtracting these formulas from each other will closely resemble the Principle of Inclusion and Exclusion.

Before proceeding to the lemmas and their proofs, we introduce some notation that will be used only in this subsection.

1. First,  $T_{2k+j,[k+i,k]}$  denotes the number of ways to sort  $2k + j$  balls into nonempty bins so that one bin has  $k$  balls and another has  $k + i$  balls. Every time this notation is used, we will have  $1 \leq i \leq j$ .
2. If we want the aforementioned quantity where exactly  $\ell$  nonempty bins are used, we denote it as  $U_{2k+j,\ell,[k+i,k]}$ .

---

<sup>1</sup>The sequence  $\{B_{2k,k}\}_{k=1}^\infty$  appears in the Online Encyclopedia of Integer Sequences [40].

3. Next,  $F_{2k+j,k,t}$  denotes the total number of ways to split  $2k + j$  balls into nonempty bins so that at least  $t$  bins have exactly  $k$  balls. Every time this notation is used in this paper,  $t$  will equal 1 or 2.
4. Finally, if we want the aforementioned quantity where exactly  $\ell$  nonempty bins are used, we denote it as  $G_{2k+j,\ell,k,t}$ .

**Lemma 3.** *Let  $i, j, k \in \mathbb{N}^+$  with  $k > j > i$ . Then the number of ways to sort  $2k + j$  balls into nonempty bins so that one bin has  $k$  balls and another has  $k + i$  balls is*

$$T_{2k+j,[k+i,k]} = \sum_{\ell=1}^{j-i} (\ell^2 + 3\ell + 2) \binom{j-i-1}{\ell-1}. \tag{3.4}$$

*Proof.* Let  $1 \leq i < j$  be arbitrary. Since one bin has  $k$  balls and another bin has  $k + i$  balls, we have  $j - i$  balls to split between nonempty bins. Say that this number of bins is denoted by  $\ell$ , and then  $1 \leq \ell \leq j - i$ . By the classical stars-and-bars method,  $j - i$  balls can be split into  $\ell$  nonempty bins in  $\binom{j-i-1}{\ell-1}$  ways. Now, we insert the remaining two bins in between these  $\ell$  bins. There are  $\ell + 1$  possible slots in which to insert the bin with  $k$  balls, and then there are  $\ell + 2$  slots to insert the bin with  $k + i$  balls (these last two bins are distinguishable since  $i > 0$ ). So, there are a total of  $(\ell^2 + 3\ell + 2) \binom{j-i-1}{\ell-1}$  arrangements of these bins. Summing over all possible values of  $\ell$  yields the desired result.  $\square$

**Remark 3.** The other quantity related to those calculated in Lemma 3 that we will need is  $T_{2k+j,[k+j,k]}$ . However, it is clear that

$$T_{2k+j,[k+j,k]} = 2, \tag{3.5}$$

because the only valid configurations are those with a bin having  $k$  balls and the other bin having  $k + j$  balls, in either order; there must be exactly two bins.

**Lemma 4.** *Let  $j, k \in \mathbb{N}^+$  with  $k > j$ . Then the total number of ways to split  $2k + j$  balls into nonempty bins so that at least two bins have exactly  $k$  balls is*

$$F_{2k+j,k,2} = \sum_{\ell=1}^j \frac{\ell^2 + 3\ell + 2}{2} \binom{j-1}{\ell-1}. \tag{3.6}$$

*Proof.* First notice that we cannot have more than two bins with  $k$  balls each while the total number of balls is  $2k + j$ , because  $k > j$ . Thus the quantity we want to calculate is the total number of ways to split  $2k + j$  balls into nonempty bins so that exactly two bins have exactly  $k$  balls. In this case, the remaining  $j$  balls can be split into  $\ell$  nonempty bins, where  $1 \leq \ell \leq j$ . By the classical stars-and-bars method, this step can be done in  $\binom{j-1}{\ell-1}$  ways. Now we insert the two bins with  $k$  balls in between the other bins, and there are  $\ell + 1$  slots in which to do this. If both



bins with  $k$  balls are in the same slot, there are  $\ell + 1$  ways to do this, and if the two bins with  $k$  balls are in different slots, there are  $\frac{\ell(\ell+1)}{2}$  ways to do this. Thus the total number of ways to insert the two bins with  $k$  balls is

$$\ell + 1 + \frac{\ell(\ell + 1)}{2} = \frac{\ell^2 + 3\ell + 2}{2}. \tag{3.7}$$

Since the value of  $\ell$  ranges from 1 to  $j$ , the desired result follows. □

The main difference between Lemma 4 and Lemma 5 is that only the former precludes the possibility of having a bin with more than  $k$  balls. If two of the bins each have at least  $k$  balls, and there are a total of  $2k + j$  balls, then no other bin can have  $k$  or more balls because  $k > j$ .

**Lemma 5.** *Let  $j, k \in \mathbb{N}^+$  with  $k > j$ . Then the total number of ways to split  $2k + j$  balls into nonempty bins so that at least one bin has exactly  $k$  balls is*

$$F_{2k+j,k,1} = \sum_{\ell=1}^{k+j} (\ell + 1) \binom{k + j - 1}{\ell - 1} - \sum_{\ell=1}^j \frac{\ell^2 + 3\ell + 2}{2} \binom{j - 1}{\ell - 1}. \tag{3.8}$$

*Proof.* Much as in the proof of (3.6), there will be either one or two bins with exactly  $k$  balls. We will find the number of ways in which to split  $2k + j$  balls into nonempty bins so that at least one bin has exactly  $k$  balls, that happens to count the configurations having two bins with  $k$  balls twice. Deliberately allowing for this over-counting gives us an easier way to count all of the configurations with exactly one bin with  $k$  balls; we will then subtract the over-counted amount, which is represented by the formula (3.6). Fix one bin to have  $k$  balls, and suppose we are splitting the remaining  $k + j$  balls into  $\ell$  nonempty bins, where  $1 \leq \ell \leq k + j$ . We can split  $k + j$  balls into  $\ell$  nonempty bins in  $\binom{k+j-1}{\ell-1}$  ways (by the classical stars-and-bars argument). Then, we insert the bin with  $k$  balls in between the other  $\ell$  bins, which can be done in  $\ell + 1$  ways. Finally, sum over the possible values of  $\ell$  and we obtain the first sum in (3.8).

However, we have double-counted configurations that have two bins with exactly  $k$  balls each. This is because when we insert the bin with  $k$  balls, there is the possibility of having one other bin with  $k$  balls already, and these two bins are not distinguishable (this reasoning is actually an implicit use of the Principle of Inclusion and Exclusion). Hence we subtract the formula (3.6) and obtain the desired result. □

In Lemma 5, the argument is only valid because the maximum number of bins with exactly  $k$  balls is two. The resulting summations become considerably more complicated if three or more bins can all have  $k$  balls apiece.

**Lemma 6.** *Let  $i, j, k, \ell \in \mathbb{N}^+$  with  $i < j < k$  and  $3 \leq \ell \leq j - i + 2$ . Then the number of ways to split  $2k + j$  balls into  $\ell$  nonempty bins so one bin has  $k$  balls and another has  $k + i$  balls is*

$$U_{2k+j, \ell, [k+i, k]} = (\ell^2 - \ell) \binom{j-i-1}{\ell-3}. \tag{3.9}$$

*Proof.* Let  $1 \leq i < j$  be arbitrary. Since one bin has  $k$  balls and another bin has  $k + i$  balls, we have  $j - i$  balls to split between  $\ell - 2$  nonempty bins. The range of possible values of  $\ell$  goes from 3 to  $j - i + 2$ , because one bin has  $k$  balls and another has  $k + i$  balls. However, we still have  $j - i$  balls left over, which can be split between as few as one additional bin and as many as  $j - i$  additional bins.

By the classical stars-and-bars method of [28],  $j - i$  balls can be split into  $\ell - 2$  nonempty bins in  $\binom{j-i-1}{\ell-3}$  ways. Now, we insert the remaining two bins in between these  $\ell$  bins. There are  $\ell - 1$  possible slots in which to insert the bin with  $k$  balls, and then there are  $\ell$  slots to insert the bin with  $k + i$  balls (these last two bins are distinguishable since  $i > 0$ ). So, there are a total of  $(\ell^2 - \ell) \binom{j-i-1}{\ell-3}$  arrangements of these bins, as desired.  $\square$

**Remark 4.** As a slight extension of the result of Lemma 6, we consider how to split  $2k + j$  balls into exactly  $\ell = 2$  nonempty bins when one bin has exactly  $k$  balls. In this case, the other bin must have  $k + j$  balls. Thus we cannot split  $2k + j$  balls into exactly 2 nonempty bins when the most crowded bin has exactly  $k$  balls.

**Lemma 7.** *Let  $j, k, \ell \in \mathbb{N}^+$  with  $j < k$  and  $3 \leq \ell \leq j + 2$ . Then the total number of ways to split  $2k + j$  balls into  $\ell$  nonempty bins so that at least two bins have exactly  $k$  balls is*

$$G_{2k+j, \ell, k, 2} = \frac{\ell^2 - \ell}{2} \binom{j-1}{\ell-3}. \tag{3.10}$$

*Proof.* First notice that we cannot have more than two bins with  $k$  balls each while the total number of balls is  $2k + j$ , because  $k > j$ . Thus the quantity we want to calculate is the total number of ways to split  $2k + j$  balls into  $\ell$  nonempty bins so that exactly two bins have exactly  $k$  balls. In this case, the remaining  $j$  balls can be split into  $\ell - 2$  nonempty bins, and now it is clear the total number of bins must range from 3 to  $j + 2$ .

By the classical stars-and-bars method used in [28], this step can be done in  $\binom{j-1}{\ell-3}$  ways. Now we insert the two bins with  $k$  balls in between the other bins, and there are  $\ell - 1$  slots in which to place the bins with exactly  $k$  balls. If both bins with  $k$  balls are in the same slot, there are  $\ell - 1$  ways to do this, and if the two bins with  $k$  balls are in different slots, there are  $\frac{(\ell-2)(\ell-1)}{2}$  ways to do this. Thus the total number of ways to insert the two bins with  $k$  balls is

$$\ell - 1 + \frac{(\ell - 2)(\ell - 1)}{2} = \frac{\ell^2 - \ell}{2}. \tag{3.11}$$

Multiplying the number of ways to perform the two steps yields the desired result.  $\square$

In this lemma we finally combine the previous lemmas in this subsection to obtain a closed form for  $M_{2k+j,\ell,k}$ .

**Lemma 8.** *Let  $j, k \in \mathbb{N}^+$  with  $j < k$  and  $2 \leq \ell \leq k + j + 1$ . Then the total number of ways to split  $2k + j$  balls into  $\ell$  nonempty bins so the most crowded bin has exactly  $k$  balls is*

$$M_{2k+j,\ell,k} =$$

$$\begin{cases} 0, & \ell = 2 \\ \ell \binom{k+j-1}{\ell-2} - \frac{\ell^2-\ell}{2} \binom{j-1}{\ell-3} - \sum_{i=1}^s (\ell^2-\ell) \binom{j-i-1}{\ell-3}, & \ell = j+2-s, 1 \leq s < j \\ \ell \binom{k+j-1}{\ell-2} - \frac{\ell^2-\ell}{2} \binom{j-1}{\ell-3}, & \ell = j+2 \\ \ell \binom{k+j-1}{\ell-2}, & j+3 \leq \ell \leq k+j+1. \end{cases} \tag{3.12}$$

*Proof.* We will handle each case separately. If  $\ell = 2$  and one bin has exactly  $k$  balls, then as discussed in the proof of Lemma 7, the other bin has  $k + j$  balls, so this case gives an answer of zero.

Now we will discuss the case where  $j + 3 \leq \ell \leq k + j + 1$ . As stated in (3.8), the number of ways to split  $2k + j$  balls into  $\ell$  nonempty bins so at least one bin has exactly  $k$  balls is  $\ell \binom{k+j-1}{\ell-2}$ . We do not want to include configurations containing a bin with  $k + i$  balls for some  $1 \leq i \leq j$ , but there are no such configurations when  $\ell$  is at least  $j + 3$  because we cannot split  $j$  balls into more than  $j$  nonempty bins.

Handling the case of  $3 \leq \ell \leq j + 1$  (the second case) is more subtle. Just as in the previous case, the number of ways to split  $2k + j$  balls into  $\ell$  nonempty bins so at least one bin has exactly  $k$  balls is  $\ell \binom{k+j-1}{\ell-2}$ , but there are now some configurations where one bin has  $k + i$  balls for some  $1 \leq i \leq j$ . There will exist  $1 \leq s < j$  such that  $\ell = j + 2 - s$ . The value of  $s$  imposes a further restriction on  $i$ , and we will determine what term to subtract from  $\ell \binom{k+j-1}{\ell-2}$  by repeatedly using Lemma 6. It is possible to have a configuration with both a bin having  $k$  balls and a bin having  $k + i$  balls for any  $1 \leq i \leq s$ , so we must subtract (3.9) for each  $1 \leq i \leq s$ , proving the proposed formula in this case. Notice that since  $j < k$  we cannot have multiple bins each having more than  $k$  balls and a bin having exactly  $k$  balls. This is why our procedure is only effective for studying  $B_{2k+j,k}$  and not  $B_{mk+j,k}$  for  $m \geq 3$ .

Finally, the case of  $\ell = j + 2$  serves as an edge case. There are no configurations having both a bin with  $k$  balls and a bin with  $k + i$  balls for some  $1 \leq i < j$ , but according to (3.10) we must subtract  $\frac{\ell^2-\ell}{2} \binom{j-1}{\ell-3}$  to get the desired result in this case.  $\square$

The formula (3.12) gives us a natural partition for the number of ways to split  $2k + j$  balls into nonempty bins where the most crowded bin has  $k$  balls: we partition on the total number of bins used.

**Lemma 9** (Summation Formula for  $B_{2k+j,k}$ ). *Let  $j, k \in \mathbb{N}^+$  with  $k > j$ . Then the total number of ways to split  $2k + j$  balls into nonempty bins so the most crowded bin has exactly  $k$  balls is*

$$B_{2k+j,k} = \sum_{\ell=2}^{k+j+1} \ell \binom{k+j-1}{\ell-2} - \sum_{\ell=3}^{j+2} \frac{\ell^2 - \ell}{2} \binom{j-1}{\ell-3} - \sum_{i=1}^{j-1} \sum_{\ell=3}^{j-i+2} (\ell^2 - \ell) \binom{j-i-1}{\ell-3} - 2. \tag{3.13}$$

*Proof.* This lemma follows from summing the result of (3.12) over  $3 \leq \ell \leq k + j + 1$ . The double sum arises from summing  $\binom{j-i-1}{\ell-3}$  over different values of  $i$  that represented over-counted configurations for each value of  $s$ . Here  $i$  ranges from 1 to  $j - 1$  because those were the possible values of  $s$  in (3.12), and we can interchange these two sums freely as they are both finite; it will be easier to evaluate this sum by summing over  $\ell$  first.

In this formula, we extend the range of the first sum to  $\ell = 2$ , and clearly that term will always have value 2; hence we subtract a 2 at the end of the equation. We do this because it will make the future calculation of that sum slightly less cumbersome.  $\square$

We now have a combination of sums which together represent the value of  $B_{2k+j,k}$ . These sums can all be converted to closed forms, but doing so is a lengthy process. We will treat the evaluation of the sums in (3.13) as distinct parts of the same lemma, and then the overall evaluation of (3.13) as a single closed form will be the main theorem for this section.

**Lemma 10.** *Let  $j, k \in \mathbb{N}^+$  with  $k > j$ . Then the sums appearing in Equation (3.13) are evaluated as follows:*

$$\sum_{\ell=2}^{k+j+1} \ell \binom{k+j-1}{\ell-2} = (k+j+3)2^{k+j-2} \tag{3.14}$$

$$\sum_{\ell=3}^{j+2} \frac{\ell^2 - \ell}{2} \binom{j-1}{\ell-3} = 2^{j-4}(j^2 + 9j + 14) \tag{3.15}$$

$$\sum_{i=1}^{j-1} \sum_{\ell=3}^{j-i+2} (\ell^2 - \ell) \binom{j-i-1}{\ell-3} = 2^{j-3}(j^2 + 5j + 2) - 2. \tag{3.16}$$

*Proof.* We will start by proving formula (3.14). In order to use the lemmas in Appendix A, we shift the index of the sum by 2:

$$\sum_{\ell=2}^{k+j+1} \ell \binom{k+j-1}{\ell-2} = \sum_{\ell=0}^{k+j-1} (\ell+2) \binom{k+j-1}{\ell}. \tag{3.17}$$

Now split this into two sums: Equation (3.17) equals

$$\sum_{\ell=0}^{k+j-1} \ell \binom{k+j-1}{\ell} + 2 \sum_{\ell=0}^{k+j-1} \binom{k+j-1}{\ell}. \tag{3.18}$$

To handle the first sum, we use Equation (A.1) from Appendix A, and the second is the binomial expansion of  $(1 + 1)^{k+j-1}$ , so Equation (3.18) equals

$$(k + j - 1)2^{k+j-2} + 2^{k+j} = (k + j + 3)2^{k+j-2}, \tag{3.19}$$

completing the proof of (3.14).

Now we will proceed with proving formula (3.15), starting by shifting the index of the sum by 3:

$$\sum_{\ell=3}^{j+2} \frac{\ell^2 - \ell}{2} \binom{j-1}{\ell-3} = \frac{1}{2} \sum_{\ell=0}^{j-1} (\ell^2 + 5\ell + 6) \binom{j-1}{\ell}. \tag{3.20}$$

We split this sum into three based on the power of  $\ell$  before the binomial coefficient:

$$\frac{1}{2} \sum_{\ell=0}^{j-1} \ell^2 \binom{j-1}{\ell} + \frac{5}{2} \sum_{\ell=0}^{j-1} \ell \binom{j-1}{\ell} + 3 \sum_{\ell=0}^{j-1} \binom{j-1}{\ell}. \tag{3.21}$$

The three sums can be evaluated respectively as follows: for the first, use (A.2); for the second use (A.1); and finally, the third is the binomial expansion of  $(1 + 1)^{j-1}$ , so Equation (3.21) equals

$$\frac{1}{2} j(j-1)2^{j-3} + \frac{5}{2} (j-1)2^{j-2} + 3 \cdot 2^{j-1}. \tag{3.22}$$

We can factor out a  $2^{j-4}$  from every term and simplify the resulting quadratic factor to obtain (3.15).

Finally, we will prove formula (3.16). Again we start by shifting the index of the (inner) sum by 3:

$$\sum_{i=1}^{j-1} \sum_{\ell=3}^{j-i+2} (\ell^2 - \ell) \binom{j-i-1}{\ell-3} = \sum_{i=1}^{j-1} \sum_{\ell=0}^{j-i-1} (\ell^2 + 5\ell + 6) \binom{j-i-1}{\ell}. \tag{3.23}$$

Break Equation (3.23) into three sums based on the power of  $\ell$  to allow us to use the identities in Appendix A. Then, (3.23) equals

$$\sum_{i=1}^{j-1} \sum_{\ell=0}^{j-i-1} \ell^2 \binom{j-i-1}{\ell} + 5 \sum_{i=1}^{j-1} \sum_{\ell=0}^{j-i-1} \ell \binom{j-i-1}{\ell} + 6 \sum_{i=1}^{j-1} \sum_{\ell=0}^{j-i-1} \binom{j-i-1}{\ell}. \tag{3.24}$$

Now, we use (A.2) on the first double sum and (A.1) on the second double sum, and notice the third term's inner sum is the binomial expansion of  $(1 + 1)^{j-i-1}$ ,

allowing us to conclude that Equation (3.24) equals

$$\sum_{i=1}^{j-1} ((j-i-1)(j-i)2^{j-i-3} + 5(j-i-1)2^{j-i-2} + 6 \cdot 2^{j-i-1}). \tag{3.25}$$

To greatly simplify the remainder of this calculation, we use a trick that somewhat resembles u-substitution for integrals. The current sum ranges from  $i = 1$  to  $i = j - 1$ , where  $j$  is fixed, so if we let  $t = j - i$  and index the sum with respect to  $t$ , the range of the sum is  $t = j - 1$  to  $t = 1$ . This greatly simplifies the summand, and by performing this substitution we realize that the sum (3.25) equals

$$\sum_{t=1}^{j-1} (t(t-1)2^{t-3} + 5(t-1)2^{t-2} + 6 \cdot 2^{t-1}). \tag{3.26}$$

We utilize a series of algebraic steps that are not particularly insightful but lead us to a closed form for Equation (3.26):

$$\begin{aligned} \sum_{t=1}^{j-1} (t(t-1)2^{t-3} + 5(t-1)2^{t-2} + 6 \cdot 2^{t-1}) &= \frac{1}{8} \sum_{t=1}^{j-1} 2^t (t^2 + 9t + 14) \\ &= \frac{1}{8} \left( \sum_{t=1}^{j-1} t^2 2^t + 9 \sum_{t=1}^{j-1} t 2^t + 14 \sum_{t=1}^{j-1} 2^t \right) \\ &= \frac{1}{8} (2^j (j^2 - 4j + 6) - 6 + 9(j2^j - 2^{j+1} + 2) + 14(2^j - 2)) \\ &= \frac{1}{8} (j^2 2^j + 5j 2^j + 2^{j+1} - 16) \\ &= 2^{j-3} (j^2 + 5j + 2) - 2, \end{aligned} \tag{3.27}$$

completing the proof. □

We can now prove the main result of this section: a closed form for  $B_{2k+j,k}$ .

**Theorem 7** (Closed Form for  $B_{2k+j,k}$ ). *Let  $j, k \in \mathbb{N}^+$  with  $k > j$ . Then the number of ways to split  $2k + j$  balls into nonempty bins so the most crowded bin has exactly  $k$  balls is*

$$B_{2k+j,k} = (k + j + 3)2^{k+j-2} - (3j^2 + 19j + 18)2^{j-4}. \tag{3.28}$$

*Proof.* This formula follows from directly substituting (3.14), (3.15), and (3.16) into (3.13). Notice that the constant terms cancel each other out. □

While we successfully found a closed form for  $B_{2k+j,k}$  with  $k > j$ , this method is not tractable as  $m$  gets larger. Section 4 serves in part to generalize our summation formulas to the case where the parameters are not within fixed multiple factors of each other.

**4. Generalized Solution for Balls into Bins with Restrictions**

The previous section dealt with the formulae for the Balls in bins problem without a dominant bin for the case when the number of bins was  $\ell$ , the most crowded bin had exactly  $k$  balls and the total number of balls was  $2k + j$  with  $j < k$ . However, one may wonder if it would be possible for the number of balls to be arbitrary and not dependent on other variables (like  $k$  in this case). Therefore, we aim to generalize the concept of putting Balls into Bins with restrictions, as defined in Question 4. Let us revisit Definition 1 and Question 4, which were posed in Section 1, to obtain a better understanding of the meaning of the problem and the notation which will be used for the rest of the paper.

**Definition 3** (Generalized Bins Restriction Problem). Let  $n, \ell, k \in \mathbb{N}^+$ . The *Generalized Bins Restriction Problem* aims to find the number of ways to split  $n$  balls into  $\ell$  bins such that the maximum number of balls in each bin is at most  $k$ .

**Question 5** (Generalized Balls into Bins with restrictions problem). Let  $n, \ell, k \in \mathbb{N}^+$ . This problem asks how many ways can we split  $n$  balls into  $\ell$  nonempty bins such that the most crowded bin has exactly  $k$  balls?

The notation associated with the quantity defined in the Generalized Bins Restriction Problem in Definition 3 will henceforth be denoted as  $R_{n,\ell,k}$ . Note that this is different from the problem we want to tackle, in the sense that the maximum number of balls in each bin is at most  $k$ , not exactly  $k$ , which was the case investigated in the earlier sections. Also, the notation associated with the quantity defined in the Generalized Balls in Bins with restriction problem in Question 5 will henceforth be denoted by  $M_{n,\ell,k}$ . The main result of this section, Theorem 10, will be a formula for  $R_{n,\ell,k}$ . An intermediate step will be to count the number of configurations described by Definition 3.

**4.1. Formula for Generalized Bins Restriction Problem:  $R_{n,\ell,k}$**

**Theorem 8.** Let  $n, \ell, k \in \mathbb{N}^+$ . Then

$$R_{n,\ell,k} = \sum_{t=0}^{\ell} (-1)^t \binom{\ell}{t} \binom{n - t(k + 1) + \ell - 1}{\ell - 1}. \tag{4.1}$$

*Proof.* We will prove this formula with the help of the Principle of Inclusion and Exclusion. First, we note that the total number of filling  $n$  balls into  $\ell$  bins with no upper restriction on the number of bins is simply

$$\binom{n + \ell - 1}{\ell - 1}. \tag{4.2}$$

Now, this would include all the configurations that satisfy the restriction condition that each bin contains at most  $k$  and all the configurations which do not satisfy

this condition. Thus the aim of this proof is to first calculate the total number of configurations that do not satisfy the restriction condition and then subtract that from Equation (4.2).

Let us assume that the bins have been labelled in some order, indexed from 1 to  $\ell$ . Let  $A_i$  denote the finite set of all configurations such that the  $i^{th}$  bin contains more than  $k$  balls. Using this definition, note that the total number of configurations that violate the restriction condition is equal to

$$\left| \bigcup_{i=1}^{\ell} A_i \right|. \tag{4.3}$$

Using the Principle of Inclusion and Exclusion, this is equivalent to the following formula:

$$\left| \bigcup_{i=1}^{\ell} A_i \right| = \sum_{i=1}^{\ell} |A_i| - \sum_{1 \leq i < j \leq \ell} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq \ell} |A_i \cap A_j \cap A_k| - \dots + (-1)^{\ell-1} |A_1 \cap A_2 \dots \cap A_{\ell}|. \tag{4.4}$$

Now, to find all such configurations, let us find the cardinalities appearing in (4.4) by breaking into cases.

Terms in the First Summation: This is the number of configurations such that bin number  $i$  contains more than  $k$  balls. Note that this statement means that bin  $i$  contains at least  $k + 1$  balls. We first put  $k + 1$  balls into bin  $i$ . Now, we just need to distribute the remaining  $n - k - 1$  balls into  $\ell$  bins (note that bin  $i$  is included because we have filled only the "minimum" number of balls in the bin, and it may contain more than  $k + 1$  balls). Therefore, the total number of ways to place  $n - 1(k + 1)$  balls into  $\ell$  bins is

$$\binom{n - 1(k + 1) + \ell - 1}{\ell - 1}. \tag{4.5}$$

We need to sum this formula over all possible values of  $i$ , i.e.,  $\sum_{i=1}^{\ell} |A_i|$ . This means that we just need to multiply the formula obtained in Equation (4.5) by  $\ell$  to sum over all values of  $i$ . Therefore, we have the following formula:

$$\sum_{i=1}^{\ell} |A_i| = \ell \binom{n - 1(k + 1) + \ell - 1}{\ell - 1} = \binom{\ell}{1} \binom{n - 1(k + 1) + \ell - 1}{\ell - 1}. \tag{4.6}$$

Terms in the Second Summation: This is the number of configurations such that the bins numbered  $i$  and  $j$ , where  $i < j$ , have more than  $k$  balls. This means both the  $i$ th and  $j$ th bins contain at least  $k + 1$  balls. So we again put  $k + 1$  balls in each bin. Therefore, we have  $n - 2(k + 1)$  balls remaining to put into  $\ell$  bins; the number of ways to place  $n - 2(k + 1)$  balls into  $\ell$  bins is

$$\binom{n - 2(k + 1) + \ell - 1}{\ell - 1}. \tag{4.7}$$



We need to sum this formula over all possible values of  $i, j$ , i.e.,  $\sum_{1 \leq i < j \leq \ell} |A_i \cap A_j|$ . The number of ways of choosing two such bins out of  $\ell$  bins is simply  $\binom{\ell}{2}$ ; we just need to multiply the formula obtained in Equation 4.7 by  $\binom{\ell}{2}$ . Therefore, we obtain the formula

$$\sum_{1 \leq i < j \leq \ell} |A_i \cap A_j| = \binom{\ell}{2} \binom{n - 2(k + 1) + \ell - 1}{\ell - 1}. \tag{4.8}$$

We proceed similarly for the other summations given in the formula of Inclusion and Exclusion, and substitute the formulae obtained into Equation (4.4) which gives us the following formula:

$$\left| \bigcup_{i=1}^{\ell} A_i \right| = \sum_{t=1}^{\ell} (-1)^{t-1} \binom{\ell}{t} \binom{n - t(k + 1) + \ell - 1}{\ell - 1}. \tag{4.9}$$

Note that this is the number of all configurations that do not satisfy the restriction condition. So, to obtain the number of configurations satisfying the restriction condition, we just subtract Equation (4.9) from Equation (4.2) to obtain the formula. Therefore, we obtain

$$R_{n,\ell,k} = \binom{n + \ell - 1}{\ell - 1} - \sum_{t=1}^{\ell} (-1)^{t-1} \binom{\ell}{t} \binom{n - t(k + 1) + \ell - 1}{\ell - 1}. \tag{4.10}$$

This equation can be further simplified as follows:

$$R_{n,\ell,k} = \sum_{t=0}^{\ell} (-1)^t \binom{\ell}{t} \binom{n - t(k + 1) + \ell - 1}{\ell - 1}. \tag{4.11}$$

□

This formula should actually have the restriction on  $t$  that  $0 \leq t \leq \frac{n}{k+1}$ , because for the cases where  $t \geq \frac{n}{k+1}$ , the binomial coefficient  $\binom{n - t(k+1) + \ell - 1}{\ell - 1}$  would return erroneous values, depending on how we extend the definition of the binomial coefficient for negative values. That said, we want all of these cases to be equated to 0 because they all violate the definition of our main problem. To make things easier, we can define the binomial coefficient  $\binom{n}{k}$  in a new (albeit equivalent) way to use the formula in Equation (4.11). The following definition of binomial coefficients will be used in the remainder of the paper.

$$\binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!}, & n \geq 0, k \geq 0, n \geq k \\ 0, & \text{otherwise} \end{cases}. \tag{4.12}$$

This definition will ease the computation of the binomial coefficients and the cases associated with these in the coming sections.

**4.2. Formula for  $M_{n,\ell,k}$**

We approached the problem of Generalized balls in bins with restriction in two different ways to obtain different formulae representing the same problem. One of the highlights of our work is that we are establishing new connections between the quantities  $R_{n,\ell,k}$  and  $M_{n,\ell,k}$  that have not been previously explored in the literature. The first method utilized in this section features the Principle of Inclusion and Exclusion to derive Equation (4).

**Theorem 9** (Formula for Generalized Balls into Bins with restrictions problem (I)). *Suppose  $n, \ell, k \in \mathbb{N}^+$  such that  $\ell + k - 1 \leq n \leq \ell k$ . Then the following identity holds:*

$$M_{n,\ell,k} = \sum_{t=1}^{\ell} (-1)^{t-1} \binom{\ell}{t} R_{n-t(k-1)-\ell, \ell-t, k-1}. \tag{4.13}$$

*Proof.* We begin the proof by taking cases for the number of bins which are full (i.e., how many bins contain exactly  $k$  balls) and then distributing the remaining balls into the bins which are not meant to be full with updated restrictions. After that, we use the Principle of Inclusion and Exclusion to sum up all the cases to prove the theorem. So, for case 1 (where at least one bin is full), the number of ways we can choose one bin out of  $\ell$  bins to be full is simply  $\binom{\ell}{1}$ . Now, let us fill the rest of the bins with the balls, while keeping the new restriction in mind. Let us call these variables new balls, new bins, and new restrict. Note that we have already given  $k$  balls to a bin, so the number of remaining bins is  $\ell - 1$ . Also note that none of the remaining  $\ell - 1$  bins can remain empty, so we put one ball in each of the remaining  $\ell - 1$  bins. The number of balls remaining is simply  $n - k - \ell + 1$ , but now each of the remaining  $\ell - 1$  bins contains only one ball. In order to apply Theorem 8 on the remaining bins, we need to reduce the restriction on the number of balls in a bin from  $k$  to  $k - 1$ . So, we have the values of the new variables as follows:

$$\text{new balls} = n - k - \ell + 1 \tag{4.14}$$

$$\text{new bins} = \ell - 1 \tag{4.15}$$

$$\text{new restrict} = k - 1. \tag{4.16}$$

In order to distribute the remaining balls into bins with the given restriction, we use Theorem 8 to identify  $R_{\text{new balls}, \text{new bins}, \text{new restrict}}$  as the total number of ways of distribution, given the above conditions. Finally, we multiply this by  $\binom{\ell}{1}$  and substitute the values of the variables to get the total number of ways such that at least 1 bin is full is

$$\binom{\ell}{1} R_{n-k-\ell+1, \ell-1, k-1}. \tag{4.17}$$

However, notice that we have over-counted in some cases, specifically the cases where at least 2 bins are full. To explain this in more detail, let  $(b_1, b_2, b_3, \dots, b_\ell)$  denote a particular valid configuration of balls in the bins numbered  $1, 2, \dots, \ell$  with the aforementioned conditions satisfied. Let us assume there are two indices  $i$  and

$j$  with  $i \neq j$  such that  $b_i = b_j = k$ ; that is, bins  $i$  and  $j$  are full with  $k$  balls each. In the explanation above, when we chose a bin that was full, we counted this configuration once when we chose  $b_i$  as the bin which is full. We also counted this configuration once when we chose  $b_j$  as the bin which is full. So, we need to subtract the case where there are at least two bins full from the above sum. The formula for the case where at least two bins are full and which satisfy the original conditions is

$$\binom{\ell}{2} R_{n-2(k-1)-\ell, \ell-2, k-1}. \tag{4.18}$$

However, if we subtract this from the case where at least one bin was full, there would be some cases, like the number of bins with at least 3 bins full, which would be under-counted, as these cases would appear in both of the above cases. So we need to add this to the original sum. By invoking the Principle of Inclusion and Exclusion, this would go on until the maximum possible number of bins contain  $k$  balls.

Considering the generalized case of putting  $n$  balls in  $\ell$  bins such that each bin has at most  $k$  balls and at least  $s$  bins are full, we have the formula

$$\binom{\ell}{s} R_{n-s(k-1)-\ell, \ell-s, k-1}. \tag{4.19}$$

Using this formula along with the Principle of Inclusion and Exclusion to sum up all the cases gives us the required formula (4.13). □

Note that in the above statement, we use the statement "maximum possible number of bins contain  $k$  balls", but have taken the summation of cases up to  $\ell$ . The above proof is valid due to the definition of  $\binom{n}{k}$ , which equates to 0 for the case where  $n \leq 0$  or  $n < k$ . So the above formula is still valid because the rest of the cases do not contribute anything to the sum.

This approach involving The Principle of Inclusion and Exclusion has given us a very complicated formula in terms of  $R_{n,\ell,k}$ , which itself has a complicated formula! Thus it makes sense to find alternate ways to represent solution to the same problem.

**Theorem 10** (Formula for Generalized Balls into Bins with restrictions problem (II)). *Suppose  $n, \ell, k \in \mathbb{N}^+$  such that  $\ell + k - 1 \leq n \leq \ell k$ . Then the following identity for  $M_{n,\ell,k}$  holds:*

$$M_{n,\ell,k} = R_{n-\ell,\ell,k-1} - R_{n-\ell,\ell,k-2} = \sum_{t=0}^{\ell} (-1)^t \binom{\ell}{t} \left[ \binom{n-tk-1}{\ell-1} - \binom{n-t(k-1)-1}{\ell-1} \right]. \tag{4.20}$$

*Proof.* We will prove this identity by counting all the configurations that satisfy only some parts of the definition of  $M_{n,\ell,k}$ , and then subtract all the configurations which do not satisfy the remaining criteria. Specifically, we temporarily relax the condition that a given bin can have at most  $k$  balls. Let us start by remembering

that the formula  $M_{n,\ell,k}$  holds true only for nonempty bins. We put one ball in each of the  $\ell$  bins. Now, the remaining number of balls is  $(n - \ell)$ , so we apply Equation (4.1) to fill the bins with the remaining  $n - \ell$  balls. However, note that the restriction must now be reduced from  $k$  to  $k - 1$  because each bin has a ball and we want at most  $k$  balls in each bin. So, the number of ways to distribute  $(n - \ell)$  balls into  $\ell$  bins such that each bin gets at most  $k - 1$  balls is simply  $R_{n-\ell,\ell,k-1}$ .

However, note that we have included some unwanted configurations in the above formula. The statement of Definition 4 also says that the most crowded bin must contain *exactly*  $k$  balls, so there are some cases where the most crowded bin might contain less than  $k$  balls, but other than that, it would satisfy all the conditions. In order to eliminate these cases, note that if the most crowded bin does not contain  $k$  balls, then it would contain at most  $k - 1$  balls. These configurations fall under Definition 1. Hence the number of ways we can distribute  $n$  balls into  $\ell$  nonempty bins such that each bin contains at most  $k - 1$  balls is  $R_{n-\ell,\ell,k-2}$ . Subtracting this from  $R_{n-\ell,\ell,k-1}$  gives the desired answer.  $\square$

**Remark 5.** The restriction  $\ell + k - 1 \leq n \leq \ell k$  was not explicitly used in the proof of (4.20). However, it is important because of the definition of  $M_{n,\ell,k}$ . Recall that  $M_{n,\ell,k}$  counts the number of ways to split  $n$  balls into  $\ell$  nonempty bins so that the most crowded bin has exactly  $k$  balls; the remaining  $\ell - 1$  bins will contain at least one ball each. If this were to be possible for some values of  $n$ ,  $\ell$ , and  $k$ , then necessarily  $\ell + k - 1 \leq n \leq \ell k$ . At least one bin must actually be filled to the maximum capacity, giving the lower bound. On the other hand, given  $\ell$  bins with maximum capacity  $k$ , one obtains the greatest possible number of balls by filling all bins with the maximum capacity. This gives the upper bound. Therefore we are justified using these bounds in the next lemma as well.

The natural follow-up question after proving (4.20) is to ask whether there is a closed form for that sum. We believe that the answer to this is negative, because the presence of  $tk$  inside the binomial coefficients makes the sum inaccessible by evaluation techniques such as Snake Oil and the WZ Method (see [43] for a detailed explanation of these techniques). As a result, we look to derive estimates instead. These estimates will involve exponential functions for two reasons. First, binomial coefficients are intimately related to exponential functions through Stirling’s formula (see for instance [4, 15, 34, 38, 42]). Second, as we will see, such estimates will produce factors with  $kt$  as a power, which is more manageable than  $kt$  inside a binomial coefficient. The statement and proof of this lemma are technical in nature and are comprised mostly of elementary manipulations of sums and inequalities. Therefore we defer the exact statement of the lemma and its proof to Appendix B.

### 4.3. Additional Identities Involving $M_{n,\ell,k}$ and $R_{n,\ell,k}$

There are various identities which involve the variables  $M_{n,\ell,k}$  and  $R_{n,\ell,k}$ . These identities help us understand the nature of the variables. The expression  $R_{n,\ell,k}$  used in this paper here is also known as **Polynomial coefficients** in various papers

like [12] and [13]. These papers provide a more rigorous definition and expression for the term  $R_{n,\ell,k}$ . The paper [12] has also explored the various combinatorial interpretations of the Polynomial coefficients in different areas of mathematics.

There are many identities of involving the Polynomial coefficients mentioned in these papers, but almost every paper has proved them using generating functions and their algebraic manipulations. Here, we present the proof of four identities using only the elementary principles of enumerative combinatorics.

**Lemma 11** (Identities involving  $R_{n,\ell,k}$ ). *Some of the prominent identities involving the variable  $R_{n,\ell,k}$  are highlighted below:*

1. 
$$R_{n,\ell,k} = R_{\ell k-n,\ell,k}; \tag{4.21}$$

2. 
$$R_{n,\ell,m+k} = \sum_{i=0}^n R_{i,\ell,m} R_{n-i,\ell,k}; \tag{4.22}$$

3. 
$$R_{n,\ell+1,k} = \sum_{i=0}^k R_{n-i,\ell,k}; \tag{4.23}$$

4. 
$$R_{n+1,\ell+1,k} - R_{n,\ell+1,k} = R_{n+1,\ell,k} - R_{n-k,\ell,k}. \tag{4.24}$$

*Proof.* Let us prove all the identities from (4.21) to (4.24) in the order listed.

1. For proving  $R_{n,\ell,k} = R_{\ell k-n,\ell,k}$ , we will use the principle of one-to-one correspondence between two sets to show that the cardinalities of both sets are equal, which would lead us to the above identity. We know that each configuration of balls in the bins with conditions expressed with the formula of  $R_{n,\ell,k}$  can be uniquely denoted as  $(b_1, b_2, b_3, \dots, b_\ell)$ , where each  $b_i$  denotes the number of balls in the bin numbered  $i$  from a selection of bins numbered from 1 to  $\ell$ . The sum of all  $b_i$ 's equals  $n$ . Furthermore, for each  $1 \leq i \leq \ell$ ,  $b_i$  is a non-negative integer which can have a maximum value of  $k$ . Let us denote  $\{b_i\}_{i=1}^\ell$  to be the same configuration  $(b_1, b_2, b_3, \dots, b_\ell)$ , but expressed in a more compact form.

Now, for each configuration of balls in bins denoted by  $\{b_i\}_{i=1}^\ell$ , we have another unique configuration of the form  $\{k-b_i\}_{i=1}^\ell$ . Note that this configuration has each term of the form  $k-b_i$ , which is always non-negative due to the fact that  $b_i \leq k$ . Also note that the maximum value of  $k-b_i$  is at most  $k$  due to

the fact that  $b_i \geq 0$ . Now, this configuration can be said to be a collection of bins with  $b_i$  balls in the  $i$ th bin; the  $\ell$  bins are indexed such that each bin is nonempty and has at most  $k$  balls. Also, the sum of the number of balls in each bin over all the bins is

$$(k - b_1) + (k - b_2) + \dots + (k - b_\ell) = \ell k - (b_1 + b_2 + \dots + b_\ell) = \ell k - n. \tag{4.25}$$

On the other hand, the total number of such configurations of this type can be expressed with the formula of  $R_{k\ell-n,\ell,k}$ . This means there is a one-to-one correspondence between the two sets of configurations satisfying their respective conditions. This implies that the cardinalities of both the sets are equal, which implies that  $R_{n,\ell,k} = R_{k\ell-n,\ell,k}$ .

2. For proving  $R_{n,\ell,m+k} = \sum_{i=0}^n R_{i,\ell,m} R_{n-i,\ell,k}$ , we proceed by counting in two ways. Suppose we have  $\ell$  bins and  $n$  balls such that each bin has at most  $(m + k)$  balls. The total number of such configurations, due to the simple definition of the Generalized Bins Restriction Problem, equals  $R_{n,\ell,m+k}$  from the formula (4.1).

Let us count the same problem from a different perspective. Let each bin have a partition which divides the bin into two parts (say  $A$  and  $B$ ) such that one component of the partition  $A$  can contain a maximum of  $m$  balls and the other component  $B$  can contain a maximum of  $k$  balls. Note that this problem is still equivalent to Question 4, because in both problems, each configuration has bins with at most  $m + k$  balls. Enumerating all the possible cases via partitions, let us fill all the partition components  $A$  for every bin. Let us suppose we require  $i$  balls ( $0 \leq i \leq n$ ) to fill each partition component  $A$  of every bin such that no bin's partition component labelled  $A$  has more than  $m$  balls. The number of such configurations is  $R_{i,\ell,m}$ . Now, we must fill the remaining  $n - i$  balls into all the partition components labelled  $B$  of every bin such that no bin has more than  $k$  balls. By a similar argument, the number of such configurations equals  $R_{n-i,\ell,k}$ .

Now, note that for any two bins  $B_i$  and  $B_j$ , the  $B_i$ s that partition component  $A$  can be interchanged with the partition component  $A$  of bin  $B_j$ , and both the bins will still satisfy the restrictions on the bins and their partition components. Hence, the total number of ways to fill  $\ell$  bins with  $n$  balls such that every bin has at most  $(m + k)$  balls is the product of the number of ways to fill the two partitions of the bins. That is, this describes the expression  $R_{i,\ell,m} R_{n-i,\ell,k}$ .

Now,  $i$  can range from 0 to  $n$  due to the number of balls available, so we should sum the total number of such configurations over all possible values of  $i$ . This is equal to  $\sum_{i=0}^n R_{i,\ell,m} R_{n-i,\ell,k}$ . So, we have just counted the same

problem in two ways, hence we have the identity

$$R_{n,\ell,k} = \sum_{i=0}^n R_{i,\ell,m} R_{n-i,\ell,k}. \tag{4.26}$$

3. To prove  $R_{n,\ell+1,k} = \sum_{i=0}^k R_{n-i,\ell,k}$ , we will use the concept of counting in two ways to prove this identity.

We first find a way to represent  $R_{n,\ell+1,k}$ . Note that for each bin  $b_i$  satisfying the restriction, we have that  $0 \leq b_i \leq k$  for any integer  $i$  such that  $1 \leq i \leq \ell+1$ . Now, let us fill the first bin first with some number of balls, and then fill the rest of the bins accordingly. Now, the minimum number of balls we can put in the first bin is  $b_1 = 0$ , and the maximum number of balls we can fill into a bin is  $b_1 = k$ . Let us find the number of ways to fill the rest of the balls into the remaining bins for each of the cases that lie within the given restrictions.

Term in the First Case ( $b_1 = 0$ ): This means the number of balls in the first bin is 0, so in order to fill the remaining  $n - 0 = n$  balls into  $\ell + 1 - 1 = \ell$  bins such that each bin contains at most  $k$  balls, the number of such configurations is  $R_{n,\ell,k}$  from Theorem 8.

Term in the Second Case ( $b_1 = 1$ ): This means the first bin has a single ball; therefore, in order to fill the remaining  $n - 1$  balls into  $\ell$  bins, such that each bin contains at most  $k$  balls, the number of such configurations is  $R_{n-1,\ell,k}$ .

For the remaining cases, i.e., where  $b_1$  ranges from 1 to  $k$ , we proceed similarly as shown above. We now discuss the last case.

Term in the ( $k + 1$ )<sup>th</sup> Case ( $b_1 = k$ ): This means the first bin has  $k$  balls; therefore, in order to distribute the remaining  $n - k$  balls amongst  $\ell$  bins, such that each bin contains at most  $k$  balls, the number of such configurations is  $R_{n-k,\ell,k}$ . Now, note that summing each of the cases above would account for every available configuration having  $n$  balls and  $\ell + 1$  bins where each bin contains at most  $k$  balls. However, the formula associated with the total number of such configurations is represented by  $R_{n,\ell+1,k}$ . We conclude that the following identity holds:

$$R_{n,\ell,k} + R_{n-1,\ell,k} + \dots + R_{n-k,\ell,k} = R_{n,\ell+1,k} \implies R_{n,\ell+1,k} = \sum_{i=0}^k R_{n-i,\ell,k}. \tag{4.27}$$

4. Finally, to prove,  $R_{n+1,\ell+1,k} - R_{n,\ell+1,k} = R_{n+1,\ell,k} - R_{n-k,\ell,k}$ , we will leverage formula (4.23) to prove this identity. Using that formula, we have the following results:

$$R_{n+1,\ell+1,k} = \sum_{i=0}^k R_{n+1-i,\ell,k}; \tag{4.28}$$

$$R_{n,\ell+1,k} = \sum_{i=0}^k R_{n-i,\ell,k}. \tag{4.29}$$

Subtracting Equation (4.29) from Equation (4.28) gives us

$$R_{n+1,\ell+1,k} - R_{n,\ell+1,k} = \sum_{i=0}^k R_{n+1-i,\ell,k} - \sum_{i=0}^k R_{n-i,\ell,k}, \tag{4.30}$$

Further, splitting the first term from Equation (4.28) and the last term from Equation (4.29) upon further simplification gives us the following:

$$R_{n+1,\ell+1,k} - R_{n,\ell+1,k} = R_{n+1,\ell,k} - R_{n-k,\ell,k} + \sum_{i=1}^k R_{n+1-i,\ell,k} - \sum_{i=0}^{k-1} R_{n-i,\ell,k}. \tag{4.31}$$

Now, notice that we may shift the index of the first sum in Equation (4.31)

$$\sum_{i=1}^k R_{n+1-i,\ell,k} = \sum_{i=0}^{k-1} R_{n-i,\ell,k} \tag{4.32}$$

Substituting Equation (4.32) into Equation (4.31) gives us the required identity (4.24).

□

These identities help develop intuition about the combinatorial properties of the Balls in bins with restriction formula  $R_{n,\ell,k}$ . We can further apply these properties to sum up the formula  $R_{n,\ell,k}$  while keeping two terms between  $n, \ell, k$  as a constant, and the third term as a variable.

#### 4.4. Sums Associated with $M_{n,\ell,k}$

In Section 1, Question 3 asked the number of ways to split  $n$  balls into any number of nonempty ordered bins where the most crowded bin has  $k$  balls. Note that in this question, the number of bins was not provided. A natural variant of this question would be to pose the same question when the other variables, such as the number of balls and the restriction on the bin, do not have fixed values. We have already discussed the case with a variable number of bins in Sections 2 and 3, which was denoted by  $B_{n,k}$ . Now, we try keeping the other factors as variables and find the appropriate sums associated with these terms.

**Definition 4.** Let  $n, \ell \in \mathbb{N}^+$ . Then,  $K_{n,\ell}$  represents the total number of ways to fill  $\ell$  nonempty bins with  $n$  balls without any restrictions.

Note that this is the same as enumerating all the possible configurations of balls and nonempty bins with the restriction on each bin ranging from 0 to  $n - \ell + 1$ .



Note that technically a restriction of 0 is not possible unless the number of balls is 0, but we have accounted for all such cases in the formula of  $R_{n,\ell,k}$ . Therefore we can write this sum as

$$K_{n,\ell} = \sum_{i=1}^{n-\ell+1} M_{n,\ell,i}. \tag{4.33}$$

Now, if we represent  $M_{n,\ell,k}$  in terms of  $R_{n,\ell,k}$ , it can be seen that the sum is actually a telescoping sum in which most of the terms cancel, and only  $R_{n-\ell,\ell,n-\ell+1}$  remains.

**Theorem 11.** *Let  $k, \ell \in \mathbb{N}^+$ . Then*

$$K_{n,\ell} = \binom{n-1}{\ell-1}. \tag{4.34}$$

*Proof.* Notice that there are no restrictions on how many balls can be in each bin. Also, Definition 4 simply represents the number of ways to distribute  $n$  balls amongst  $\ell$  nonempty bins, which is already a very well-known problem in combinatorics. To give a perspective for the formula associated with  $K_{n,\ell,k}$ , we start by filling each bin with one ball. This is because we do not have any empty bins. So, we have  $n - \ell$  balls remaining to be distributed into  $\ell$  bins, such that each bin gets 0 or more balls. The number of ways this can be done is

$$\binom{(n-\ell) + \ell - 1}{\ell - 1} = \binom{n-1}{\ell-1}$$

from the stars-and-bars argument. □

The next definition focuses on having a constant number of bins, and a restriction on the maximum number of balls that can be present in the bins. The main task is to enumerate all possibilities of filling these bins with balls while satisfying the given restrictions.

**Definition 5.** Let  $k, \ell \in \mathbb{N}^+$ . Then  $N_{\ell,k}$  represents the total number of ways to fill any number of balls into  $\ell$  nonempty bins such that the most crowded bin(s) contain exactly  $k$  balls.

In this case, we are varying the number of balls, whereas the bins and the restriction on the bin remain the same. The minimum number of balls required in the  $\ell$  bins such that all the bins are nonempty and at least 1 bin contains  $k$  balls would be equal to  $k + \ell - 1$ . The maximum number of balls would be utilized when all bins have  $k$  balls, which would be equal to  $k\ell$ . Therefore, we can write the sum of all such configurations as

$$N_{\ell,k} = \sum_{i=k+\ell-1}^{k\ell} M_{i,\ell,k}. \tag{4.35}$$

**Theorem 12.** *Let  $k, \ell \in \mathbb{N}^+$ . Then*

$$N_{\ell,k} = k^\ell - (k - 1)^\ell. \tag{4.36}$$

*Proof.* We will solve this problem by first counting a more general version of the problem, and then subtracting all the cases violating the original conditions. Recall that we have these restrictions:

1. The most crowded bin(s) have exactly  $k$  balls.
2. All of the bins are nonempty.

Let us partially fulfill these conditions, and fill all the bins such that all bins are nonempty and the maximum number of balls in the bins are at most  $k$ . We have  $\ell$  bins, and each bin could have anywhere between 1 to  $k$  balls; as there is no restriction on the total number of balls available, the total number of filling  $\ell$  bins with this condition is  $k^\ell$ . Note that in the above case, we have also counted the configurations where the most crowded bin has less than  $k$  balls. For example, consider the configuration  $(k - 1, k - 1, \dots, k - 1)$  where all  $\ell$  bins contain  $(k - 1)$  balls. Here, the most crowded bins contain  $(k - 1)$  balls. However, in the original definition, the most crowded bin has exactly  $k$  balls. Thus in order to count only those configurations which have the most crowded bin with exactly  $k$  balls, we subtract all such configurations whose most crowded bin contains less than  $k$  balls from  $k^\ell$  as shown above. The number of ways to fill  $\ell$  bins such that the most crowded bin(s) contain at most  $(k - 1)$  balls is simply  $(k - 1)^\ell$ . Subtracting the two derived expressions yields the desired result.  $\square$

### 5. Interpretation in Counting Solutions to Integer Equations

The problem studied in depth through earlier sections of this paper happens to be analogous with a class of problems that have a more robust literature. One elementary problem that frequents itself in mathematics textbooks such as [4, 27, 43] is to find the number of positive integer solutions to

$$x_1 + x_2 + \dots + x_\ell = n, \tag{5.1}$$

where  $\ell \leq n$ . This equation with various choices for the coefficients is closely studied in additive and enumerative combinatorics papers such as [7, 30, 36, 37]. Incidentally, this problem also motivated the development of the “stars-and-bars” argument we used numerous times throughout the paper. There are many restrictions that can be added to these solutions. One of particular relevance is where we require that the maximum attained value of  $\{x_1, x_2, \dots, x_n\}$  is some  $k > 0$ ; this is relevant because finding a formula for  $M_{n,\ell,k}$  is clearly equivalent to finding the number of positive integer solutions to (5.1) with this restriction, and the authors are not aware of any previously published resources that provide a closed form to

the solution of this problem. Among other things, this paper gives us such a closed form when  $n \leq 2k$  or when there exists  $j \in \mathbb{N}^+$  such that  $j < k$  and  $n = 2k + j$ . Here are two examples where we demonstrate that our formulas agree with the list of solutions to (5.1) that is generated by inspection.

**Example 1.** Using formula (3.1), we calculate

$$M_{8,5,4} = 5 \binom{4-1}{5-2} = 5. \tag{5.2}$$

On the other hand, the list of solutions to (5.1) when  $n = 8, \ell = 5, k = 4$  is

$$(x_1, x_2, x_3, x_4, x_5) \in \{(4, 1, 1, 1, 1), (1, 4, 1, 1, 1), (1, 1, 4, 1, 1), (1, 1, 1, 4, 1), (1, 1, 1, 1, 4)\}, \tag{5.3}$$

and there are 5 total solutions on this list.

**Example 2.** Using the third case of Equation (3.12), we calculate

$$M_{8,4,3} = 4 \binom{3+2-1}{4-2} - \frac{4^2-4}{2} \binom{2-1}{4-3} = 18. \tag{5.4}$$

On the other hand, the list of solutions to (5.1) when  $n = 8, \ell = 4, k = 3$  is

$$(x_1, x_2, x_3, x_4) \in \{(3, 3, 1, 1), (3, 1, 3, 1), (3, 1, 1, 3), (1, 3, 3, 1), (1, 3, 1, 3), (1, 1, 3, 3), (3, 2, 2, 1), (3, 2, 1, 2), (3, 1, 2, 2), (2, 3, 2, 1), (2, 3, 1, 2), (1, 3, 2, 2), (2, 2, 3, 1), (2, 1, 3, 2), (1, 2, 3, 2), (2, 2, 1, 3), (2, 1, 2, 3), (1, 2, 2, 3)\}, \tag{5.5}$$

and there are 18 total solutions on this list.

### 6. Conclusion and Future Work

We studied two variants of the same problem in this paper; we counted how many ways to split  $n$  balls into nonempty, ordered bins so that the most crowded bin has exactly  $k$  balls, for certain values of  $n$  and  $k$ . In the first variant, we fixed the total number of bins that were allowed; in the second variant, we permitted any number of bins to be used, as long as all the bins were nonempty. Clearly, the solution to the second problem follows very quickly from the first, but the first problem is of interest in its own right. As we demonstrated in Section 5, the first problem yields a previously unknown closed form for counting the number of solutions to certain integer equations.

That being said, we have yet to find a closed form for  $B_{n,k}$  for all  $n, k \in \mathbb{N}^+$ . The case where there exists  $j < k$  such that  $n = 2k + j$  may be most illustrative to

how we handle the case where  $n = mk + j$  for some  $m \geq 3$ . Thus, the natural open problem to consider is whether the technique used in Section 3.2 can be extended to this more general case. Finding a closed form for this general case will completely solve the problem at hand, because whenever  $n > k$ , there exists  $m \in \mathbb{N}^+$  and  $0 \leq j < k$  such that  $n = mk + j$ . The case where  $m = 1$  was the dominant bin case considered in Section 2, and the  $m = 2$  case was handled in Section 3.

In a previous discussion on public forums<sup>2</sup>, people have approximated the distribution of  $M_{n,\ell,k}$  as a normal distribution. Here the parameter is  $\ell$  (i.e., the number of bins); thus it may be worthwhile to solidify this theory.

There are also other aspects of this problem which could be explored in the future. One of them is exploring the behaviour of the average number of bins, given the number of balls  $n$ , the restriction on each bin  $k$ , and the fact that at least 1 bin contains exactly  $k$  balls. Basically, we are exploring the properties of  $M_{n,\ell,k}$  with  $\ell$  as a variable. The average number of bins required in this case would just be the weighted sum of the number of bins required and the number of such configurations divided by the total number of configurations possible for all cases of the number of bins. Using Python libraries such as NumPy and Matplotlib, we explored the nature of the average number of bins with respect to other factors (such as the number of balls, and the restriction on the bins). Figure 1 shows the number of configurations for a constant restriction on the bins and an increasing number of balls. Using many simulations of such test cases, it can be safely conjectured that the average number of bins increases if we keep the restriction on bins as a constant value and increase the number of balls.

Finally, there are some other problems to consider for future research. One variant of the problems studied in this paper is as follows: we can enumerate ways to split  $n$  balls into nonempty, ordered bins so that the most crowded bin has exactly  $k$  balls, and exactly  $t$  bins have this many balls. This is a sensible problem to study because the calculations in Section 3.2 implicitly broke the problem at hand into the cases  $t = 1$  and  $t = 2$ . There is also an asymptotic problem to consider: for fixed  $n$  and  $k$ , we can consider the list of all configurations where we split  $n$  balls into nonempty, ordered bins so the most crowded bin has exactly  $k$  balls as a probability space with the parameter being the number of bins used; then we ask what probability distribution is resembled as  $n \rightarrow \infty$  for a fixed  $k$ .

<sup>2</sup><https://math.stackexchange.com/questions/3548108/counting-solutions-to-2Dx-1-x-2-dots-x-k-n-with-x-i-leq-r-closed-for>

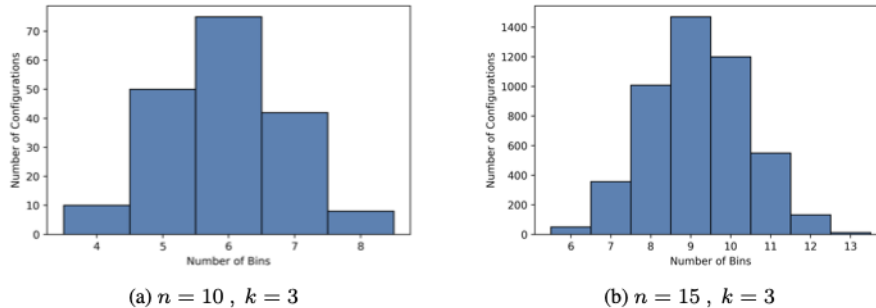


Figure 1: Distribution of the number of configurations with varying number of balls; our GitHub repository with the code is located at <https://github.com/vedantbonde/Balls-in-Bins-Analysis>

**Acknowledgements.** We thank our colleague Steven J. Miller and the Polymath Jr. Program for facilitating this collaboration.

**References**

- [1] H. Alzer and H. Prodinger, Identities and inequalities for sums involving binomial coefficients, *Integers* **20** (2020), 1–12.
- [2] S. Aki, Survey of discrete distributions of order k and related distributions, <https://web.archive.org/web/20081230061932/http://www2.ipcku.kansai-u.ac.jp/~aki/survey.pdf>.
- [3] N. Borade, D. Cai, D.Z. Chang, B. Fang, A. Liang, S.J. Miller, and W. Xu, Gaps of summands of the Zeckendorf lattice, <https://arxiv.org/pdf/1909.01935.pdf>.
- [4] M. Bóna, *A Walk Through Combinatorics: An Introduction to Enumeration and Graph Theory*, World Scientific, Singapore, 2012.
- [5] K. Boyadzhiev, Close encounters with Stirling numbers of the second kind, *Math. Mag.* **85** (2012), 252-266.
- [6] K. Boyadzhiev, *Notes on the Binomial Transform: Theory and Table*, World Scientific, Singapore, 2018.
- [7] B. Bukh, Non-trivial solutions to a linear equation in integers, *Acta Arith.* Vol. **131** (2008), no. 1, 51-55.
- [8] C.A. Charalambides, Lucas numbers and polynomials of order k and the length of the longest circular success run, *Fibonacci Quart.* **29** (1991), 290–297.
- [9] C.A. Charalambides, Success runs in a circular sequence of independent Bernoulli trials, A.P. Godbole (ed.) St.G. Papastavrides (ed.), *Runs and Patterns in Probability*, Kluwer Acad. Publ. (1994), 15–30.
- [10] E. Chen, R. Chen, L. Guo, C. Jiang, S. J. Miller, J. M. Siktar, and P. Yu, Gaussian behavior in Zeckendorf decompositions from lattices, *Fibonacci Quart.* **57** (2019), no. 3, 201-212.

- [11] R.W. Donley, Jr., Binomial arrays and generalized Vandermonde identities, <https://arxiv.org/pdf/1905.01525.pdf>.
- [12] N. Fahssi, Polynomial triangles revisited, <https://arxiv.org/pdf/1202.0228.pdf>.
- [13] N. Fahssi, Some identities involving polynomial coefficients, *Fibonacci Quart.* **54** (2016), 125-136.
- [14] E. Fang, J. Jenkins, Z. Lee, D. Li, E. Lu, S. J. Miller, D. Salgado, and J. M. Siktar, Central limit theorems for compound paths on the 2-dimensional lattice, *Fibonacci Quart.* **58** (2020), 208-225.
- [15] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd edition, New York, Wiley, 1968.
- [16] B. Guo and F. Qi, Some identities and an explicit formula for Bernoulli and Stirling numbers, *J. of Comput. Appl. Math.* **255** (2014), 568-579.
- [17] J. Hardtke, Higher derivatives of the inverse tangent function and a summation formula involving binomial coefficients, <https://arxiv.org/pdf/1809.08184v2.pdf>.
- [18] L. Holst and T. Konstantopoulos, Runs in coin tossing: a general approach for deriving distributions for functionals, *J. Appl. Probab.* **52** (2015), 752-770.
- [19] W. T. Huang and C.S. Tsai, On a modified binomial distribution of order  $k$ , *Statist. Prob. Lett.* **11** (1991), 125-131.
- [20] M. Kologlu, G. Kopp, S.J. Miller, and Y. Wang, On the number of summands in Zeckendorf decompositions, *J. Number Theory* **49** (2011), no. 2, 116-130.
- [21] T. Konstantopoulos, Z. Liu, and X. Yang, Laplace transform asymptotics and large deviation principles for longest success runs in Bernoulli trials, <https://arxiv.org/pdf/1507.02935.pdf>.
- [22] A.N. Kumar and N.S. Upadhye, On generalizations of  $(k_1, k_2)$ -runs, <https://arxiv.org/pdf/1707.08367.pdf>.
- [23] A.N. Kumar and N.S. Upadhye, Pseudo-binomial approximation to  $(k_1, k_2)$ -runs, <https://arxiv.org/pdf/1609.07847.pdf>.
- [24] M. Li and R. Goldman, Limits of sums for binomial and Eulerian numbers and their associated distributions, <https://arxiv.org/pdf/1903.06317.pdf>.
- [25] R. Li and S. J. Miller, Central limit theorems for gaps of generalized Zeckendorf decompositions, *Fibonacci Quart.* **57** (2019), no. 3, 213-230.
- [26] N. D. Megill, and M. Pavičić, Estimating Bernoulli trial probability from a small sample, <https://arxiv.org/pdf/1105.1486.pdf>
- [27] S. J. Miller, *The Probability Lifesaver*, Princeton University Press, Princeton, NJ, 2017.
- [28] S. J. Miller and Y. Wang, Gaussian behavior in generalized Zeckendorf decompositions, *Combinatorial and Additive Number Theory CANT 2011 and 2012* (Melvyn B. Nathanson, editor), Springer Proceedings in Mathematics & Statistics (2014), 159-173.
- [29] Balls in bins With limited capacity, MathPages <https://www.mathpages.com/home/kmath337/kmath337.htm>.
- [30] V. N. Murty, Counting the integer solutions of a linear equation with unit coefficients, *Math. Mag.* **54** (1981), 79-81.

- [31] M. Muselli, Simple expressions for success run distributions in Bernoulli trials, *Statistics & Probability Letters* **31** (1996), 121–128.
- [32] A. S. Nimbran, Sums of series involving central binomial coefficients and harmonic numbers, <https://arxiv.org/pdf/1806.03998v2.pdf>.
- [33] M. Riggle, Runs of Identical Outcomes in a Sequence of Bernoulli Trials, *Masters Theses and Specialist Projects* (2018), Paper 2451.
- [34] H. Robbins, A remark on Stirling’s formula, *Amer. Math. Monthly* **62** (1995), 26-29.
- [35] R. Roy, Binomial identities and hypergeometric series, *Amer. Math. Monthly* **94** (1987), 36-46.
- [36] I.Z. Ruzsa, Solving a linear equation in a set of integers. I., *Acta Arith.* **65** (1993), 259–282.
- [37] I.Z. Ruzsa, Solving a linear equation in a set of integers. II., *Acta Arith.* **72** (1995), 385–397.
- [38] P. Stanica, Good lower and upper bounds on binomial coefficients, *J. Inequal. Pure and Appl. Math.* **3** (2001), Art.30, 5pp.
- [39] M. Shattuck and T. Waldhauser, Proofs of some binomial identities using the method of last squares, *Fibonacci Quart.* **48** (2010), no. 4, 290-297.
- [40] N. J. A. Sloane, *The on-line encyclopedia of integer sequences*.
- [41] A. Sofo, Sums of derivatives of binomial coefficients, *Adv. Appl. Math.* **42** (2009), 123-134.
- [42] Z.H. Sun, Inequalities for binomial coefficients, <https://arxiv.org/pdf/1310.0353.pdf>.
- [43] H.S. Wilf, *Generatingfunctionology*, Academic Press, 1994, 226 pages.

## A. Some Useful Binomial Coefficient Identities

In this appendix we will provide proofs of some binomial coefficient identities that are used throughout the paper and have many uses outside the context of this paper. The proofs in this section are very similar in nature to those in the Appendix of [10]. They may be regarded as well-known, but we provide proofs for sake of completeness and for future reference.

**Lemma 12.** For any  $n \in \mathbb{N}^+$ ,

$$\sum_{k=0}^n k \binom{n}{k} = n2^{n-1}. \quad (\text{A.1})$$

*Proof.* This follows immediately from using the formula for the mean of a binomial random variable on page 330 of [27], using value of the parameter  $p = \frac{1}{2}$ .  $\square$

**Lemma 13.** For any  $n \in \mathbb{N}^+$ ,

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

*Proof.* We aim to rewrite the left-hand side of (A.2) so that we can invoke the formula (A.1). To do this we turn the binomial coefficient into  $\binom{n}{k-1}$ :

$$\sum_{k=0}^n k^2 \binom{n}{k} = \sum_{k=0}^n k^2 \frac{n!}{k!(n-k)!} = \sum_{k=0}^n k(n-k+1) \binom{n}{k-1}. \tag{A.3}$$

We notice this sum's first term vanishes and shift the index of it:

$$\sum_{k=0}^n k(n-k+1) \binom{n}{k-1} = \sum_{k=1}^n k(n-k+1) \binom{n}{k-1} = \sum_{k=0}^{n-1} (k+1)(n-k) \binom{n}{k}. \tag{A.4}$$

Now we decompose this into three sums:

$$\sum_{k=0}^{n-1} (k+1)(n-k) \binom{n}{k} = -\sum_{k=0}^{n-1} k^2 \binom{n}{k} + (n-1) \sum_{k=0}^{n-1} k \binom{n}{k} + n \sum_{k=0}^{n-1} \binom{n}{k}. \tag{A.5}$$

All three sums look familiar if we include the  $k = n$  terms: the first sum is the opposite of the left-hand side of (A.2), the second sum is the left-hand side of (A.1), and the right-hand side is the binomial expansion of  $(1 + 1)^n$ . Thus we decide to add the  $k = n$  terms and then subtract them:

$$\begin{aligned} & -\sum_{k=0}^{n-1} k^2 \binom{n}{k} + (n-1) \sum_{k=0}^{n-1} k \binom{n}{k} + n \sum_{k=0}^{n-1} \binom{n}{k} = \\ & -\sum_{k=0}^n k^2 \binom{n}{k} + (n-1) \sum_{k=0}^n k \binom{n}{k} + n \sum_{k=0}^n \binom{n}{k} + n^2 \binom{n}{n} - (n-1)n \binom{n}{n} + n \binom{n}{n}. \end{aligned} \tag{A.6}$$

It turns out the last three terms on the right-hand side of (A.6) cancel each other out. In addition, everything in (A.6) is equal to the left-hand side of (A.2), so we can treat those quantities like an equation and rearrange it to obtain

$$2 \sum_{k=0}^n k^2 \binom{n}{k} = (n-1) \sum_{k=0}^n k \binom{n}{k} + n \sum_{k=0}^n \binom{n}{k}. \tag{A.7}$$

The desired result follows from a direct application of (A.1) and the fact that the rightmost sum in (A.7) is equal to  $2^n$ .  $\square$

**Remark 6.** The process used in this proof can be repeated to evaluate sums of the form  $\sum_{k=0}^n k^m \binom{n}{k}$ , but in this paper we only need the results for  $m = 1$  and  $m = 2$ . There is no known closed form for this sum for general values of  $m$ . See [5, 6] for more information on past explorations with this family of sums, including a discussion on how Stirling's triangle was developed.



**B. Estimates on Enumerating Generalized Balls into Bins with Restrictions**

Before obtaining our estimate on Equation (4.20), we need one other preliminary lemma. This identity may be regarded as well-known, but we provide a quick proof for sake of completeness.

**Lemma 14.** *Let  $m \in \mathbb{N}^+$ . Then*

$$\sum_{\substack{0 \leq t \leq m \\ t \text{ even}}} \binom{m}{t} = \sum_{\substack{0 \leq t \leq m \\ t \text{ odd}}} \binom{m}{t} = 2^{m-1}. \tag{B.1}$$

*Proof.* The following two identities follow from applying the Binomial Theorem to expand  $(1 + 1)^m$  and  $(1 - 1)^m$ , respectively.

$$\sum_{t=0}^m \binom{m}{t} = 2^m \tag{B.2}$$

$$\sum_{t=0}^m (-1)^t \binom{m}{t} = 0. \tag{B.3}$$

We can add (B.2) and (B.3) together. In doing so, the terms with odd index  $t$  cancel out while the terms with even index  $t$  double in magnitude. This proves the first equality in (B.1). Subtracting this equality from (B.2) yields the second equality in (B.1), completing the proof.  $\square$

**Lemma 15.** *Let  $n, \ell, k \in \mathbb{N}^+$  such that  $k \leq n \leq \ell k$  and  $n \geq 2$ . Then we have the following estimates on  $M_{n,\ell,k}$ :*

$$\begin{aligned} M_{n,\ell,k} \leq & 2 \binom{\ell}{\alpha} \binom{n - \alpha k - 1}{\ell - 1} + 2 \binom{\ell}{\beta} \binom{n - \beta(k - 1) - 1}{\ell - 1} + \\ & \left( \frac{2^{\ell-1}}{(\ell - 1)!} \cdot \frac{(n - 1)^{n-\frac{1}{2}} e^{1-\ell+\ell k} e^{\frac{1}{12(n-\ell k-1)} - \frac{1}{12(n-\ell)+1}}}{(n - (\alpha - 1)k - \ell)^{n-\ell k-\ell+\frac{1}{2}}} \right) - \\ & \left( \frac{2^\ell}{(\ell - 1)!} \cdot (n - \ell)^{-\ell-1} e^{1-\ell} e^{\frac{1}{12n-11} - \frac{1}{12}} \right) + \\ & \left( \frac{2^{\ell-1}}{(\ell - 1)!} \cdot \frac{(n - 1)^{n-\frac{1}{2}} e^{1-\ell+\ell k} e^{\frac{1}{12(n-\ell(k-1)-1)} - \frac{1}{12(n-\ell)+1}}}{(n - (\beta - 1)(k - 1) - \ell)^{n-\ell(k-1)-\ell+\frac{1}{2}}} \right). \tag{B.4} \end{aligned}$$

$$\begin{aligned}
 M_{n,\ell,k} \geq & -2 \binom{\ell}{\alpha} \binom{n - \alpha k - 1}{\ell - 1} - 2 \binom{\ell}{\beta} \binom{n - \beta(k - 1) - 1}{\ell - 1} + \\
 & \left( \frac{2^\ell}{(\ell - 1)!} \cdot (n - \ell)^{-\ell - 1} e^{1 - \ell} e^{\frac{1}{12n - 11} - \frac{1}{12}} \right) - \\
 & \left( \frac{2^{\ell - 1}}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}} e^{1 - \ell + \ell k} e^{\frac{1}{12(n - \ell k - 1)} - \frac{1}{12(n - \ell) + 1}}}{(n - (\alpha - 1)k - \ell)^{n - \ell k - \ell + \frac{1}{2}}} \right) - \\
 & \left( \frac{2^{\ell - 1}}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}} e^{1 - \ell + \ell k} e^{\frac{1}{12(n - \ell(k - 1) - 1)} - \frac{1}{12(n - \ell) + 1}}}{(n - (\beta - 1)(k - 1) - \ell)^{n - \ell(k - 1) - \ell + \frac{1}{2}}} \right). \tag{B.5}
 \end{aligned}$$

The constants  $\alpha$  and  $\beta$  are defined as follows:

$$\alpha := \max\{t \in \{0, 1, \dots, \ell\}, n - tk - 1 \geq \ell - 1\} \tag{B.6}$$

$$\beta := \max\{t \in \{0, 1, \dots, \ell\}, n - t(k - 1) - 1 \geq \ell - 1\}. \tag{B.7}$$

*Proof.* The idea behind the proof is as follows: recall that Equation (4.20)'s summation form is a sum over  $t$  ranging from 0 to  $\ell$ . Each term in the sum has binomial coefficients  $\binom{n - tk - 1}{\ell - 1}$  and  $\binom{n - t(k - 1) - 1}{\ell - 1}$ . We will estimate these binomial coefficients from above and below, and then use the alternating nature of the sum (4.20) to bound each term of the sum appropriately, based on the parity of  $t$ . An application of (B.1) to the resulting bounds will complete the proof.

We will also need to rule out some trivial cases for how  $n$ ,  $\ell$ , and  $k$  relate in order to validate the forthcoming calculations. Recall that  $M_{n,\ell,k}$  denotes the number of ways to split  $n$  balls into  $\ell$  nonempty bins where the most crowded bin has exactly  $k$  balls. In this setup, the total number of balls is at most  $\ell k$ , where all  $\ell$  bins are filled to maximum capacity. That is,  $n \leq \ell k$ . On the other hand, at least one bin must be filled to maximum capacity, so  $k \leq n$ . See Remark 5.

If  $n - k\ell - \ell < 0$  then  $n - k\ell - 1 < \ell - 1$ , which means the binomial coefficient  $\binom{n - tk - 1}{\ell - 1}$  vanishes when  $t = \ell$ . The binomial coefficient estimates we will prove are

$$\binom{n - tk - 1}{\ell - 1} \leq \frac{1}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}}}{(n - (\alpha - 1)k - \ell)^{n - \ell k - \ell + \frac{1}{2}}} e^{1 - \ell + \ell k} e^{\frac{1}{12(n - \ell k - 1)} - \frac{1}{12(n - \ell) + 1}}; \tag{B.8}$$

$$\binom{n - tk - 1}{\ell - 1} \geq \frac{1}{(\ell - 1)!} \cdot (n - \ell)^{-\ell - 1} e^{1 - \ell} e^{\frac{1}{12n - 11} - \frac{1}{12}}; \tag{B.9}$$

$$\begin{aligned}
 \binom{n - t(k - 1) - 1}{\ell - 1} \leq & \frac{1}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}}}{(n - (\beta - 1)(k - 1) - \ell)^{n - \ell(k - 1) - \ell + \frac{1}{2}}} \\
 & e^{1 - \ell + \ell k} e^{\frac{1}{12(n - \ell(k - 1) - 1)} - \frac{1}{12(n - \ell) + 1}}; \tag{B.10}
 \end{aligned}$$

$$\binom{n-t(k-1)-1}{\ell-1} \geq \frac{1}{(\ell-1)!} \cdot (n-\ell)^{-\ell-1} e^{1-\ell} e^{\frac{1}{12n-11} - \frac{1}{12}}, \tag{B.11}$$

where  $0 \leq t < \alpha$  in (B.8) and (B.9), and  $0 \leq t < \beta$  in (B.12) and (B.14). Notice that if  $t > \alpha$ , then the binomial coefficient  $\binom{n-tk-1}{\ell-1}$  vanishes; if  $t > \beta$  then the binomial coefficient  $\binom{n-t(k-1)-1}{\ell-1}$  vanishes. With that in mind, (4.20) is rewritten by truncating the sum:

$$M_{n,\ell,k} = (-1)^\alpha \binom{\ell}{\alpha} \binom{n-\alpha k-1}{\ell-1} - (-1)^\beta \binom{\ell}{\beta} \binom{n-\beta(k-1)-1}{\ell-1} + \sum_{t=0}^{\alpha-1} (-1)^t \binom{\ell}{t} \binom{n-tk-1}{\ell-1} - \sum_{t=0}^{\beta} (-1)^t \binom{\ell}{t} \binom{n-t(k-1)-1}{\ell-1}. \tag{B.12}$$

The central Stirling-type estimates we will invoke are

$$\sqrt{2\pi m}^{m+\frac{1}{2}} e^{-m} e^{\frac{1}{12m+1}} \leq m! \leq \sqrt{2\pi m}^{m+\frac{1}{2}} e^{-m} e^{\frac{1}{12m}}. \tag{B.13}$$

These are proven in [15, 34]. There are also similar estimates proven in [42], but those are only relevant when  $n$  is large. To obtain the upper bound in (B.8), we use the definition of the binomial coefficient along with both estimates in (B.13):

$$\begin{aligned} \binom{n-tk-1}{\ell-1} &= \frac{(n-tk-1)!}{(\ell-1)!(n-tk-\ell)!} \\ &\leq \frac{1}{(\ell-1)!} \cdot \frac{(n-tk-1)^{n-tk-\frac{1}{2}} e^{-n+tk+1} e^{\frac{1}{12(n-tk-1)}}}{(n-tk-\ell)^{n-tk-\ell+\frac{1}{2}} e^{-n+tk+\ell} e^{\frac{1}{12(n-tk-\ell)+1}}}. \end{aligned} \tag{B.14}$$

However, the above is only guaranteed to hold if  $t < \alpha$ . It can be seen that for any  $t < \alpha$ , we have  $n-tk-1 > \ell-1$ . However, if  $n-\alpha k-1 = \ell-1$ , then the binomial coefficient in question is actually equal to 1. Rather than break into cases depending on whether  $\alpha$  has this property or not, we simply separate the term where  $t = \alpha$  within our summation formula (B.12). We use analogous reasoning to justify separating the  $t = \beta$  term from the second sum in (B.12). Henceforth in our further estimation of (B.14), we will have  $0 \leq t \leq \alpha-1$ . As a consequence,  $n-tk-\ell > 0$ , and we will no longer need to be concerned with inadvertent division by 0. The goal is to estimate (B.14) as tightly as possible while removing all dependencies on  $t$ . Our next step in this vein will be to estimate  $n-tk-1 \leq n-1$  in the numerator and  $n-tk-\ell \geq n-(\alpha-1)k-\ell$  in the denominator:

$$\binom{n-tk-1}{\ell-1} \leq \frac{1}{(\ell-1)!} \cdot \frac{(n-1)^{n-tk-\frac{1}{2}} e^{-n+tk+1} e^{\frac{1}{12(n-tk-1)}}}{(n-(\alpha-1)k-\ell)^{n-tk-\ell+\frac{1}{2}} e^{-n+tk+\ell} e^{\frac{1}{12(n-tk-\ell)+1}}}. \tag{B.15}$$

Now we estimate exponents in the numerator from above, and exponents in the denominator from below, to obtain

$$\binom{n-tk-1}{\ell-1} \leq \frac{1}{(\ell-1)!} \cdot \frac{(n-1)^{n-\frac{1}{2}} e^{-n+tk+1} e^{\frac{1}{12(n-tk-1)}}}{(n-(\alpha-1)k-\ell)^{n-\ell k-\ell+\frac{1}{2}} e^{-n+\ell} e^{\frac{1}{12(n-\ell)+1}}}. \tag{B.16}$$

Notice that  $n - 1 > 0$  since we assumed  $n \geq 2$ . Finally, this inequality can be consolidated slightly to obtain (B.8). This concludes the proof of the upper bound for  $\binom{n-tk-1}{\ell-1}$ . Now we will prove the lower bound (B.9) in a similar manner. Use the definition of the binomial coefficient along with both estimates in (B.13):

$$\begin{aligned} \binom{n-tk-1}{\ell-1} &= \frac{(n-tk-1)!}{(\ell-1)!(n-tk-\ell)!} \\ &\geq \frac{1}{(\ell-1)!} \cdot \frac{(n-tk-1)^{n-tk-\frac{1}{2}} e^{-n+tk+1} e^{\frac{1}{12(n-tk-1)+1}}}{(n-tk-\ell)^{n-tk-\ell+\frac{1}{2}} e^{-n+tk+\ell} e^{\frac{1}{12(n-tk-\ell)}}}. \end{aligned} \tag{B.17}$$

Just as in the proof of the upper bound, we assume  $n-tk-\ell > 0$  and use  $n-tk-1 > n-tk-\ell$  to simplify (B.17) into

$$\binom{n-tk-1}{\ell-1} \geq \frac{1}{(\ell-1)!} \cdot (n-tk-\ell)^{-\ell-1} \cdot \frac{e^{-n+tk+1} e^{\frac{1}{12(n-tk-1)+1}}}{e^{-n+tk+\ell} e^{\frac{1}{12(n-tk-\ell)}}}, \tag{B.18}$$

which in turn is easily consolidated into

$$\binom{n-tk-1}{\ell-1} \geq \frac{1}{(\ell-1)!} \cdot (n-tk-\ell)^{-\ell-1} e^{1-\ell} \cdot \frac{e^{\frac{1}{12(n-tk-1)+1}}}{e^{\frac{1}{12(n-tk-\ell)}}}. \tag{B.19}$$

Next, we utilize the inequalities  $\frac{1}{12(n-tk-1)+1} \geq \frac{1}{12n-11}$  and  $\frac{1}{12(n-tk-\ell)} \leq \frac{1}{12}$  to conclude

$$\binom{n-tk-1}{\ell-1} \geq \frac{1}{(\ell-1)!} \cdot (n-tk-\ell)^{-\ell-1} e^{1-\ell} e^{\frac{1}{12n-11} - \frac{1}{12}}. \tag{B.20}$$

There is one remaining appearance of  $t$ . Since the exponent  $-\ell - 1$  of  $n - tk - \ell$  is negative, we will acquire another bound from below by using  $n - tk - \ell < n - \ell$ . The result (B.9) follows from this. Moreover, the results (B.12) and (B.14) follow from replicating the proofs of (B.8) and (B.9), but with replacing  $k$  with  $k - 1$  and  $\alpha$  with  $\beta$ . Now, we prove the estimate (B.4). We rewrite (B.12) by breaking the sums into two sums each: one over even indices, and the other over odd indices. Precisely,

$$\begin{aligned} M_{n,\ell,k} &= (-1)^\alpha \binom{\ell}{\alpha} \binom{n-\alpha k-1}{\ell-1} - (-1)^\beta \binom{\ell}{\beta} \binom{n-\beta(k-1)-1}{\ell-1} + \\ &\quad \sum_{\substack{0 \leq t \leq \alpha-1 \\ t \text{ even}}} \binom{\ell}{t} \binom{n-tk-1}{\ell-1} - \sum_{\substack{0 \leq t \leq \alpha-1 \\ t \text{ odd}}} \binom{\ell}{t} \binom{n-tk-1}{\ell-1} - \\ &\quad \sum_{\substack{0 \leq t \leq \beta-1 \\ t \text{ even}}} \binom{\ell}{t} \binom{n-t(k-1)-1}{\ell-1} + \sum_{\substack{0 \leq t \leq \beta-1 \\ t \text{ odd}}} \binom{\ell}{t} \binom{n-t(k-1)-1}{\ell-1}. \end{aligned} \tag{B.21}$$

By our choice of  $\alpha$ ,  $\binom{n-\alpha k-1}{\ell-1}$  equals either 1 or 0. Similarly,  $\binom{n-\beta(k-1)-1}{\ell-1}$  also equals either 1 or 0. Since we are handling inequalities, we will assume those terms

possess the largest possible magnitudes. Then, we will use the bounds from above (B.8) and (B.12) to estimate the positive terms in the sums from above; we will use the bounds from below (B.9) and (B.14) to estimate the negative terms in the sums from above:

$$\begin{aligned}
 M_{n,\ell,k} \leq & 2 \binom{\ell}{\alpha} \binom{n - \alpha k - 1}{\ell - 1} + 2 \binom{\ell}{\beta} \binom{n - \beta(k - 1) - 1}{\ell - 1} + \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ even}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}} e^{1 - \ell} e^{\frac{1}{12(n - \ell k - 1)} - \frac{1}{12(n - \ell) + 1}}}{(n - (\alpha - 1)k - \ell)^{n - \ell k - \ell + \frac{1}{2}}} \right) - \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ odd}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot (n - \ell)^{-\ell - 1} e^{1 - \ell + \ell k} e^{\frac{1}{12n - 11} - \frac{1}{12}} \right) - \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ even}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot (n - \ell)^{-\ell - 1} e^{1 - \ell + \ell k} e^{\frac{1}{12n - 11} - \frac{1}{12}} \right) + \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ odd}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}} e^{1 - \ell} e^{\frac{1}{12(n - \ell(k - 1) - 1)} - \frac{1}{12(n - \ell) + 1}}}{(n - (\beta - 1)(k - 1) - \ell)^{n - \ell(k - 1) - \ell + \frac{1}{2}}} \right). \tag{B.22}
 \end{aligned}$$

Finally, one can interchange the summation with the factors that are independent of  $t$ , and invoke Lemma 15 to conclude the desired upper bound (B.4).

Similarly, we will use the bounds from below (B.9) and (B.14) to estimate the positive terms in the sums from below; we will use the bounds from above (B.8) and (B.12) to estimate the negative terms in the sums from below:

$$\begin{aligned}
 M_{n,\ell,k} \geq & -2 \binom{\ell}{\alpha} \binom{n - \alpha k - 1}{\ell - 1} - 2 \binom{\ell}{\beta} \binom{n - \beta(k - 1) - 1}{\ell - 1} + \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ even}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot (n - \ell)^{-\ell - 1} e e^{\frac{1}{12n - 11} - \frac{1}{12}} \right) - \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ odd}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}}}{(n - (\alpha - 1)k - \ell)^{n - \ell k - \ell + \frac{1}{2}}} e^{1 - \ell + \ell k} e^{\frac{1}{12(n - \ell k - 1)} - \frac{1}{12(n - \ell) + 1}} \right) - \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ even}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot \frac{(n - 1)^{n - \frac{1}{2}} e^{1 - \ell + \ell k} e^{\frac{1}{12(n - \ell(k - 1) - 1)} - \frac{1}{12(n - \ell) + 1}}}{(n - (\beta - 1)(k - 1) - \ell)^{n - \ell(k - 1) - \ell + \frac{1}{2}}} \right) + \\
 & \sum_{\substack{0 \leq t \leq \ell \\ t \text{ odd}}} \binom{\ell}{t} \left( \frac{1}{(\ell - 1)!} \cdot (n - \ell)^{-\ell - 1} e^{1 - \ell} e^{\frac{1}{12n - 11} - \frac{1}{12}} \right). \tag{B.23}
 \end{aligned}$$

Just as before, interchange the summation with the factors that are independent of  $t$ , and invoke Lemma 14. This procedure will prove (B.5), completing the entire proof.  $\square$

**Remark 7.** The last step in the proof of Lemma 15 could arguably be made more precise if we broke the lemma into cases based on the parity of  $\alpha$  and  $\beta$ . However, doing so would further obscure the desired result while providing virtually no additional insight.