COMpletely Effective Error Bounds For Stirling Numbers of the First And Second Kinds Via Poisson Approximation

Richard Arratia and Stephen DeSalvo

Abstract. We provide completely effective error estimates for Stirling numbers of the first and second kind, denoted by \( s(n, m) \) and \( S(n, m) \), respectively, for values of \( m \geq n - O(\sqrt{n}) \). An application of our Theorem 3 yields, for example,

\[
s(10^{12}, 10^{12} - 2 \times 10^6)/10^{35664464} \in [1.87567, 1.87801],
S(10^{12}, 10^{12} - 2 \times 10^6)/10^{35664463} \in [1.28068, 1.32750].
\]

The bounds are obtained via Poisson approximation, using an interpretation of Stirling numbers as the number of ways of placing non-attacking rooks on a chess board. As a corollary, we obtain a simple and explicit asymptotic formula, one for each of \( s(n, m) \) and \( S(n, m) \), which agrees with a recent expansion by Louchard [11, 10] for the parametrization \( m = n - t n^a, \frac{1}{2} < a < 1 \), namely,

\[
|s(n, m)| \sim \binom{n}{t n^a} e^{-\frac{3}{4} t^2 n^{2a-1}},
S(n, m) \sim \binom{n}{t n^a} e^{-\frac{4}{3} t^2 n^{2a-1}}.
\]

We thus conclude in Proposition 4 that these asymptotic formulas hold for all \( 0 \leq a < 1 \).

Finally, we generalize to rook and file numbers.

1. Introduction

The asymptotic enumeration of Stirling numbers of the first kind and of the second kind has a history dating back at least to Jordan [8], who found asymptotic formulas for \( S(n, m) \) and \( s(n, m) \) when \( m \) is fixed and \( n \) tends to infinity. Since then, a number of asymptotic formulas covering a range of parameter values when \( m \to \infty \) has been provided by numerous authors. In particular, in a series of two papers, Moser and Wyman [12, 13] treated the case when \( 0 < m < n \) and \( \lim n - m = \infty \) by a saddle point analysis. More recently, Chelluri, Richmond, and Temme [2] succinctly summarized the asymptotic enumeration of Stirling numbers of both kinds for the full range of parameter values. Their approach is the same as Moser and Wyman’s, but generalized to include real values of the parameters.

For Stirling numbers of the second kind, Moser and Wyman [13] obtained a complete asymptotic expansion with an explicit hard error term valid for parameter values \( m = \).
\[ n - o(\sqrt{n}), \text{ namely,} \]

\[ S(n, m) = \binom{n}{m} q^{-(n-m)} \sum_{k=0}^{n-m-1} A_k^{n-m} q^k, \]

where \( q = \frac{2}{n-m} \) and the \( A_k^{n-m} \)'s are polynomials in \( m \), with

\[ \sum_{k=0}^{n-m-1} A_k^{n-m} q^k = \left[ 1 + \frac{(n-m)^2}{12} q + \frac{(n-m)q^2}{288} + \left( \frac{(n-m)_6}{10368} - \frac{(n-m)_4}{1440} \right) q^3 + \cdots \right]. \]

In particular, using only the first \( s \) powers of \( q \) in the summation above, they obtained

\[ (1) \quad S(n, m) = \binom{n}{m} q^{-(n-m)} \left( \sum_{k=0}^{s} A_k^{n-m} q^k + E_s \right), \quad |E_s| \leq \frac{(2(n-m)^2/5m)^{s+1}}{1 - 2(n-m)^2/5m}. \]

This error tends to zero for values of \( m = n - o(\sqrt{n}) \), and does provide completely effective error estimates in that range.

Their analysis continued in general for all \( 0 < m < n \) such that \( \lim_{n \to \infty} n - m = \infty \), although their resulting formula is defined in terms of implicit parameters, and no hard error bounds were obtained. The main result is the first order asymptotic formula

\[ (2) \quad S(n, m) = \frac{n! (\exp(R) - 1)^m}{2R^m m! \sqrt{\pi mRH}} (1 + O(m^{-1/3})), \]

where \( R \) is the solution to

\[ R(1 - e^{-R})^{-1} = n/m, \]

and

\[ H = e^R(e^R - 1)/(e^R - 1)^2. \]

Quoting from [13, Page 40],

In order to complete our discussion it would be valuable to give an accurate estimate of the accuracy of our formulae. Such an estimate would involve the order of magnitude of the terms that have been dropped in the derivation of our asymptotic formula (2). These terms are of a very complicated nature, and such an estimate would be difficult to obtain.

They go on to provide a heuristic analysis of the error and demonstrate its accuracy for particular values of the parameters. While giving the reader a general sense of the accuracy for large values of \( n \), there are no hard error bounds akin to Equation (1).

For Stirling numbers of the first kind, similarly, we have from [12]

\[ (3) \quad |s(n, m)| = \frac{\Gamma(n + R)}{R^n \Gamma(R) \sqrt{2\pi H}} (1 + O(m^{-1})), \]

\[ ^1 \text{but using our equation numbers} \]
where $R$ is the unique solution to

$$\sum_{k=0}^{n-1} \frac{R}{R+k} = m,$$

and

$$H = m - \sum_{k=0}^{n-1} \frac{R^2}{(R+k)^2}.$$

Quoting from [12, Pages 142–143],

One of the defects of formula (3) is the fact that we have given no estimate of
the error involved in using only those terms shown in (3). ... Unfortunately,
we have been unable to give even such a crude estimate of the error involved
in the use of (3).

Our main result, Theorem 5, is an explicit upper and lower bound for each of the Stirling
numbers of the first and second kind that holds for all finite values of $n$ and $m$. The bounds
are non-trivial for values of $m \geq n - O(\sqrt{n})$. An application of this theorem implies, for
example, that

$$s(10^{12}, 10^{12} - 2 \times 10^6)/10^{35664464} \in [1.87567, 1.87801],$$

$$S(10^{12}, 10^{12} - 2 \times 10^6)/10^{35664463} \in [1.28068, 1.32750].$$

Our theorem is an application of Poisson convergence via the Chen-Stein method, which
provides completely effective error estimates by bounding the total variation distance be-
tween a sum of Bernoulli random variables and an appropriately chosen Poisson random
variable. We use an interpretation of the Stirling numbers as the number of ways of placing
$k$ non-attacking rooks on the lower triangular half of an $(n-1) \times (n-1)$ chess board
[9], see also [17, Page 75]. For Stirling numbers of the second kind, this corresponds to the
value $S(n,n-k)$. A similar interpretation for Stirling numbers of the first kind also holds,
with value $|s(n,n-k)|$, except one takes “non-attacking” to mean only column-wise.

In Section 2 we recall the relevant definitions and a summary of asymptotic formulas
for various ranges of the parameters given in [2]. A more complete historical treatment
is contained in [2, 10], and we refer the interested reader to the references therein. We
state in Proposition 1 a simple explicit asymptotic formula for the Stirling numbers, one
for each kind, valid for values $m = n - t n^a$, $0 \leq a < 1$, $t > 0$, $n \to \infty$. The validity of
the formulas for all $0 \leq a < 1$ follows from Louchard’s analysis [11, 10] for $\frac{1}{2} < a < 1$, and
our own for $0 \leq a \leq \frac{1}{2}$. The asymptotic formulas involving implicitly defined parameters
were originally derived by Moser and Wyman [12, 13] in the forms of Equation (2) and
Equation (3), without hard error bounds, and in Section 3 we state them in a compact,
explicit formula with completely effective error estimates that hold for all finite values of the
parameters.

2These values were computed using Mathematica’s arbitrary precision library.
3Moser and Wyman cover the case $0 \leq a < \frac{1}{2}$. 
In Section 4 we prove our main result, the concrete error bounds on the Stirling numbers, by defining an appropriate collection of indicator random variables and obtaining a bound in total variation distance of their sum to a Poisson random variable with appropriately chosen mean. Our method does not involve a single integral, but rather constructs a coupling between two random variables which consequently requires only a straightforward counting argument.

2. Stirling numbers

The Stirling numbers of the first and second kind, denoted by \( s(n,m) \) and \( S(n,m) \), respectively, where \( n \geq m \geq 1 \) are integers, are defined as follows: let

\[
(x)_n = x(x - 1) \ldots (x - n + 1)
\]
denote the falling factorial function, then we have

\[
\sum_{k=0}^{n} s(n,k) x^k = (x)_n, \tag{4}
\]

\[
\sum_{k=0}^{n} S(n,k)(x)_k = x^n. \tag{5}
\]

Alternatively, one can define the Stirling numbers by the recursions:

\[
s(n+1,m) = -ns(n,m) + s(n,m-1), \quad n \geq m \geq 1,
\]

\[
s(0,0) = 1, \quad s(n,0) = s(0,n) = 0,
\]

and

\[
S(n+1,m) = mS(n,m) + S(n,m-1), \quad n \geq m \geq 1,
\]

\[
S(0,0) = 1, \quad S(n,0) = S(0,n) = 0.
\]

The values for \( m = 1, 2 \) and all \( n \geq m \) have simple formulas:

\[
s(n,1) = (-1)^{n-1}(n-1)!, \quad s(n,2) = (-1)^n(n-1)! \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right).
\]

\[
S(n,1) = 1, \quad S(n,2) = 2^{n-1} - 1.
\]

The numbers \( s(n,m) \) are sometimes referred to as the signed Stirling numbers of the first kind, since they take both positive and negative values. If, however, one considers instead \( |s(n,m)| \), i.e., the unsigned Stirling numbers of the first kind, then these numbers count the number of permutations of \( n \) distinct elements into exactly \( m \) disjoint cycles. Hence,

\[
n! = \sum_{m=0}^{n} |s(n,m)|.
\]

The simple relation \( s(n,m) = (-1)^{n-m}|s(n,m)| \) allows us to state our results in terms of the more natural combinatorial quantity \( |s(n,m)| \).
The numbers $S(n, m)$ count the number of set partitions of size $n$ into exactly $m$ non-empty blocks. Hence

$$B_n = \sum_{m=0}^{n} S(n, m),$$

where $B_n$ is the $n$-th Bell number, which counts the number of partitions of a set of size $n$.

An alternative characterization of Stirling numbers, the one we utilize in the present paper, involves non-attacking rooks on a chess board.

**Theorem 1.** Let $B$ denote the set of lower triangular squares of an $(n - 1) \times (n - 1)$ chess board.

a. [9] The number $S(n, n - k)$ counts the number of ways of placing $k$ rooks on $B$ such that no pairs of rooks are in the same row or column.

b. [4] The number $|s(n + 1, n + 1 - k)|$ counts the number of ways of placing $k$ rooks on $B$ such that no pairs of rooks lie in the same column.

We briefly summarize a few of the known asymptotic results for Stirling numbers, just enough to characterize the asymptotic behavior for all ranges of the parameters values. For a more comprehensive treatment, see [10, 2]. References are given in the right-most column.

**Theorem 2.** In what follows, $\delta$ is a positive constant that stays fixed as $n \to \infty$. For Stirling numbers of the first kind, as $n \to \infty$,

$$|s(n, m)| = \frac{(n-1)!}{(m-1)!} (\log n + \gamma)^{m-1} (1 + O(\log^{-1} n)), \quad m = O(\log(n)), \quad [12, 8, 7],$$

$$|s(n, m)| = \frac{\Gamma(n+R)}{R^n \Gamma(R) \sqrt{2\pi H}} (1 + O(n^{-1})), \quad \sqrt{\log n} \leq m \leq n - n^{1/3}, \quad [12, 2],$$

$$|s(n, m)| = {n \choose m} \left(\frac{m}{2}\right)^{n-m} (1 + O(n^{-1/3})), \quad n - n^{1/3} \leq m \leq n, \quad [12, 2],$$

where $\gamma$ is Euler’s constant, $R$ is the unique solution to

$$\sum_{k=0}^{n-1} \frac{R}{R+k} = m,$$

and

$$H = m - \sum_{k=0}^{n-1} \frac{R^2}{(R+k)^2}.$$
For Stirling numbers of the second kind, as \( n \to \infty \),
\[
S(n, m) = \frac{n^n}{m!} \exp \left[ \left( \frac{n}{m} - m \right) e^{-n/m} \right] (1 + o(1)), \quad m < n / \ln n, \quad \text{[13 Page 144]},
\]
\[
S(n, m) = \frac{n!(\exp(R) - 1)^m}{2R^n m! \sqrt{\pi mR}} (1 + O(n^{-1})) \quad 0 < \delta \leq m \leq n - n^{1/3}, \quad \text{[12 2]},
\]
\[
S(n, m) = \frac{n^{2(n-m)}}{2^n m! (n-m)!} (1 + O(n^{-1/3})) \quad n - n^{1/3} \leq m \leq n, \quad \text{[6 2]},
\]
where \( R \) is the solution to
\[
R(1 - e^{-R})^{-1} = n/m,
\]
with
\[
H = e^R (e^R - 1 - R)/2(e^R - 1)^2.
\]

Those results in a sense completely describe the asymptotic behavior of the Stirling numbers of both kinds for various values of \( n \) and \( m \). Our contribution is the addition of error estimates, and an explicit formula in terms of \( n \) and \( m \). In fact, recently, in [11, 10], the implicit dependence on \( R \) in the formulas above was explicitly calculated for \( m = n - n^a \), \( a > 1/2 \), and an asymptotic expansion was given that depends explicitly on the value of \( a \). These results can be summarized as follows:

**Theorem 3** ([11 10]). Fix any \( a \in (\frac{1}{2}, 1) \). Then as \( n \) tends to infinity we have
\[
|s(n, n - n^a)| \sim e^{n^a((2-a) \ln n + 1 - \ln(2))} \left\{ \frac{e^{n^a((2-a) \ln n + 1 - \ln(2))}}{\sqrt{2\pi n^a/2}}, \right. \]
\[
S(n, n - n^a) \sim e^{n^a((2-a) \ln n + 1 - \ln(2))} \left\{ \frac{e^{n^a((2-a) \ln n + 1 - \ln(2))}}{\sqrt{2\pi n^a/2}}, \right. \]
where
\[
T_1 = x \left[ 1 - \ln(2) + 2 \ln(y) + \ln(x) - \frac{2}{3y} - \frac{2}{9y^2} + \ldots \right] + \ln(2) - 2 \ln(y) - \ln(x),
\]
\[
T_1' = x \left[ 1 - \ln(2) + 2 \ln(y) + \ln(x) - \frac{4}{3y} - \frac{5}{9y^2} + \ldots \right] + \ln(2) - 2 \ln(y) - \ln(x),
\]
and \( x = n^a, \ y = n^{1-a} \).

Theorem 3 above is from a multiseries expansion (see e.g. [10]) and requires \( n^{1-a} = o(n^a) \) and \( n^a = o(n) \), hence the case \( a = 1/2 \) does not apply directly. In addition, the explicit error term in Equation (1) does not go to zero for \( a = \frac{1}{2} \). Nevertheless, we obtain as a corollary to our main result that the coefficient of \( x/y \) in the asymptotic expansions of \( T_1 \) and \( T_1' \) is indeed the correct term when \( a = 1/2 \), and thus the above asymptotic formulas apply more generally for all \( 0 \leq a < 1 \).
Proposition 1. Suppose \( m = n - O(\sqrt{n}) \), then asymptotically as \( n \to \infty \),

\[
|s(n,m)| \sim \left( \binom{n}{m} \right) e^{-2\mu} \left( 1 + O\left( \frac{n-m}{n} \right) \right),
\]

\[
S(n,m) \sim \left( \binom{n}{m} \right) e^{-\mu} \left( 1 + O\left( \frac{n-m}{n} \right) \right).
\]

In particular, let \( t > 0 \) and \( 0 \leq a < 1 \) be fixed. Then for \( m = n - tn^a \), as \( n \to \infty \) we have

\[
|s(n,m)| \sim \left( \frac{n}{t n^a} \right) e^{-\frac{t^2 n^{2a}}{2}} \sim \frac{1}{\sqrt{2\pi t^3 n^{a}}} e^{(2-a)n^a \ln n - \frac{3}{4} t^2 n^{2a-1}},
\]

\[
S(n,m) \sim \left( \frac{n}{t n^a} \right) e^{-\frac{t^2 n^{2a}}{4}} \sim \frac{1}{\sqrt{2\pi t^3 n^{a}}} e^{(2-a)n^a \ln n - \frac{3}{4} t^2 n^{2a-1}}.
\]

Proof. Equations (6) and (7) follow from Theorem 5. Equations (8) and (9) follow from Proposition 3 for \( \frac{1}{2} < a < 1 \), and Theorem 5 for \( 0 \leq a \leq \frac{1}{2} \). \( \square \)

3. Main Result

We begin by defining total variation distance between the laws of two random variables as

\[
d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{A \subset \mathbb{R}} |P(X \in A) - P(Y \in A)|.
\]

In particular, when \( X \) and \( Y \) are discrete random variables we have

\[
d_{TV}(\mathcal{L}(X), \mathcal{L}(Y)) = \frac{1}{2} \sum_i |P(X = i) - P(Y = i)|.
\]

In order to simplify notation, we make the common abuse of notation by denoting the left hand side of the above equation as \( d_{TV}(X,Y) \). Finally, we use the case \( A = \{0\} \) of Equation (10), i.e.,

\[
|P(X = 0) - P(Y = 0)| \leq d_{TV}(X,Y).
\]

Now we recall a classical Stein’s method result on Poisson convergence due to Chen, see for example [5, Page 130].

Theorem 4 (Chen-Stein). Let \( P \) denote a Poisson random variable with expectation \( \lambda \). For an index set \( \Gamma \), let \( W := \sum_{\alpha \in \Gamma} X_\alpha \) denote the sum of indicator random variables, and for each \( \alpha \in \Gamma \), let \( V_\alpha \) be a random variable on the same probability space as \( W \), distributed with

\[
\mathcal{L}(V_\alpha) = \mathcal{L}(W - 1|X_\alpha = 1).
\]

\footnote{Note that our definition of total variation distance differs from another conventional version by a factor of 2. This is so that \( d_{TV} \) is a number between 0 and 1 rather than between 0 and 2.}
Let \( p_\alpha := \mathbb{E} X_\alpha \), then for \( \sum_\alpha \mathbb{E} p_\alpha = \lambda \in (0, \infty) \), we have

\[
d_{TV}(W, P) \leq \min(1, \lambda^{-1}) \sum_{\alpha \in \Gamma} p_\alpha \mathbb{E} |W - V_\alpha|.
\]

In our setting, we have the board \( B \) with \( |B| = \binom{n}{2} \), and we place \( k \) rooks in distinct squares, with all \( \binom{|B|}{k} \) placements equally likely. We index each of the \( k \) rooks by numbers in the set \( \{1, 2, \ldots, k\} \), and index each pair of rooks by \( \alpha = (a, b) \), where \( 1 \leq a < b \leq k \). Then \( X_\alpha \) is defined as the indicator random variable that takes the value 1 when the pair of rooks \( (a, b) \) are attacking. We let \( W = \sum_\alpha X_\alpha \) equal the total number of pairs of rooks that are in attacking configurations. The random variable \( W \) is approximately Poisson distributed with mean \( \mathbb{E} W = \sum_\alpha \mathbb{E} X_\alpha \), and the event of interest is \( \{W = 0\} \). We have

\[
P(W = 0) = \frac{S(n, n - k)}{\binom{n}{2}},
\]

using the standard definition of attacking rooks.

Let \( Y \) denote a Poisson random variable with expectation \( \sum_\alpha \mathbb{E} X_\alpha = \lambda \), then we have for each \( n \geq k > 0 \),

\[
|P(W = 0) - P(Y = 0)| = \left| \frac{S(n, n - k)}{\binom{n}{2}} - e^{-\lambda} \right| \leq d_{TV}(W, Y).
\]

One can then bound the right hand side of Equation (14) using Theorem 4. The calculation for Stirling numbers of the first kind is analogous.

**Theorem 5.** Suppose \( n \geq 3 \) and \( n \geq k \geq 2 \). Let \( N := \binom{n}{2} \),

\[
\mu \equiv \mu_{n,k} := \frac{\binom{k}{2} \binom{n}{3}}{N},
\]

\[
d_3 := \left[ \frac{4(k - 1) \binom{n}{3}}{N} + (k - 2) \frac{6 \binom{n}{3}}{\binom{n}{2} \binom{(n/2) - 2}{2}} + \frac{(k - 2)}{\binom{n}{2} - 2} \left( \frac{5n - 11}{4} \right) + \frac{k(k - 4) + 5}{N} \right],
\]

\[
d_4 := \mu d_3,
\]

and

\[
D_{n,k} := \min(d_3, d_4, 1).
\]

Then we have

\[
\binom{n}{k} e^{-2\mu} (1 - 2e^{2\mu} D_{n,k}) \leq S(n, n - k) \leq \binom{n}{k} e^{-2\mu} (1 + 2e^{2\mu} D_{n,k}).
\]

\[
\binom{n}{k} e^{-\mu} (1 - e^{\mu} D_{n,k}) \leq |s(n, n - k)| \leq \binom{n}{k} e^{-\mu} (1 + e^{\mu} D_{n,k}).
\]
Note that the error term is absolute for all finite values of \( n \geq k \geq 2 \), and goes to 0 for \( k = O(\sqrt{n}) \), which proves Proposition 1 for values of \( 0 \leq a \leq \frac{1}{2} \). We note that this range is strictly greater than the one originally obtained by Moser and Wyman for Stirling numbers of the second kind, and the first of its kind for Stirling numbers of the first kind.

4. Proof of Theorem 5

We begin with the setup described in Equations (13) and (14)\[ \] In what follows, pairs of rooks will be indexed by the ordered pair \( \alpha = (a,b) \), \( 1 \leq a < b \leq k \). We define indicator random variables \( Y_\alpha = 1(\text{rooks } a \text{ and } b \text{ are attacking}) \). Let \( q_\alpha = \mathbb{E} Y_\alpha \), and let \( R_\alpha \) (respectively, \( C_\alpha \)) denote the corresponding indicator random variable for the event that rooks \( a \) and \( b \) are attacking the same row (respectively, column), so that \( Y_\alpha = R_\alpha + C_\alpha \).

We denote by \( \binom{n}{k} \) the standard binomial coefficients, and we define

\[
M := \binom{n}{2}, \quad N := \frac{\binom{n}{2}}{2} = \frac{M(M-1)}{2}.
\]

The quantity \( M \) is the number of squares on the board \( B \), and \( N \) is the number of ways of selecting two unordered squares on \( B \). In addition, we shall also use the identity

\[
\sum_{m=k}^{n-1} \left( \begin{array}{l} m \\ k \end{array} \right) = \left( \begin{array}{l} n \\ k+1 \end{array} \right), \quad n > k \geq 1.
\]

First we establish the Poisson rate, starting with

\[
q_\alpha = \mathbb{E} (R_\alpha) + \mathbb{E} (C_\alpha) = \frac{1}{N} \left( \sum_{r=2}^{n-1} \binom{r}{2} + \sum_{c=1}^{n-2} \binom{n-c}{2} \right) = \frac{2 \binom{n}{3}}{N}.
\]

Summing over all unordered pairs of rooks, we have

\[
\mu = \sum_\alpha q_\alpha = \frac{2 \binom{k}{2} \binom{n}{3}}{N}.
\]

In order to proceed with the total variation bound implied by Theorem 5 we must first define a joint probability space for variables \( W \) and \( V_\alpha \). The total variation bound holds for all random variables \( V_\alpha \) with distribution given by (11), and we are free to choose any coupling; ideally, one that minimizes \( \mathbb{E} |W - V_\alpha| \).

To describe the coupling, we think of the board as containing \( k \) rooks with labels \( \{1,2,\ldots,k\} \) that are visible and \( |B| - k \) rooks with labels \( \{k+1,k+2,\ldots,|B|\} \) that are invisible. Let \( \pi \in S_B \), where \( S_B \) denotes the set of permutations of the numbers \( 1,2,\ldots,|B| \), which corresponds to the set of all possible arrangements of all rooks, visible and invisible. By symmetry in the rooks 1 to \( k \), in Equation (12) we know that the value of \( p_\alpha \) does not vary with \( \alpha \), and hence we can arrange couplings of the \( V_\alpha \) with \( W \) so that \( \mathbb{E} |W - V_\alpha| \) also does not vary with \( \alpha \). Therefore we consider, from now on, and without loss of generality, the specific case \( \alpha = (1,2) \).

\[5\] An alternative setup is where rooks can possibly lie on the same square. This makes their locations independent, but weakens the eventual bound.
The coupling is: independent of $\pi$, pick a pair of locations $(I, J)$ according to the distribution $P((I, J) = (i, j)) = p_{i,j}$ that assigns equal probability to every pair of locations on $B$ for two attacking rooks.\footnote{Note that $I$ and $J$ have the same marginal distribution, which is not uniform over $B$.}

Let $\pi_1$ and $\pi_2$ denote the locations of rooks 1 and 2, respectively, under the permutation $\pi$. Then swap the coordinates of rooks labelled 1 and 2 with the rooks in locations $I$ and $J$, respectively. There are several cases to consider when performing the swap, since there may be some overlap between the sets $\{\pi_1, \pi_2\}$ and $\{I, J\}$.

Case 1: Both rooks lie in the two squares (in any order), i.e., $\{\pi_1, \pi_2\} = \{I, J\}$.

Case 2: One rook is in $\{I, J\}$, the other is not.

Case 3: Neither of the two rooks is in $\{I, J\}$, i.e., $\{\pi_1, \pi_2\} \cap \{I, J\} = \emptyset$.

In case 1, the coupling does nothing to the original permutation $\pi$. In case 2, e.g., if $\pi_1 = I$, then it swaps rooks 2 and $\pi_j^{-1}$; if $\pi_1 = J$, then it swaps rooks 2 and $\pi_i^{-1}$, etc., so that after the swapping rooks 1 and 2 lie in spots $I$ and $J$, respectively. In case 3, we choose to swap rooks 1 and $\pi_j^{-1}$ and rooks 2 and $\pi_i^{-1}$. Call this new permutation under the coupling $\pi'$, then $\pi' \in S_B$ is uniform over all permutations conditioned on rooks 1 and 2 attacking.

A formal description of the coupling is as follows: Here we use cycle notation $(i \ j)$ to denote a 2-cycle if $i \neq j$, and $(i \ j) \circ \pi$ denotes the permutation $\pi$ composed with an additional swap between the entries at $i$ and $j$ when $i \neq j$, and simply $\pi$ when $i = j$. Then our coupling for $\alpha = (1, 2)$ has the form (see also Figure 2):

$$\pi' = \begin{cases} 
\pi & \pi_1 = I, \pi_2 = J \\
(I \ J) \circ \pi & \pi_1 = J, \pi_2 = I \\
(\pi_1 \ I) \circ \pi & \pi_2 = J, \pi_1 \neq I \\
(\pi_1 \pi_2 \ J) \circ \pi & \pi_1 \neq J, \pi_2 = I \\
(\pi_2 \ \pi_1 \ I) \circ \pi & \pi_1 = J, \pi_2 \neq I \\
(\pi_2 \ J) \circ \pi & \pi_1 = I, \pi_2 \neq J \\
(\pi_1 \ I) \circ (\pi_2 \ J) \circ \pi & \pi_1 \neq I, \pi_2 \neq J.
\end{cases}$$

(17)

Alternatively, we can write this more simply as

$$\pi^*: = (\pi_2 \ J) \circ \pi,$$

$$\pi' = (\pi_1 \ 1) \circ \pi^*,$$

or

$$\pi' = \begin{cases} 
(\pi_1 \pi_2 \ J) \circ \pi, & \pi_2 = I \\
(\pi_2 \pi_1 \ J) \circ \pi, & \pi_1 = J \\
(\pi_1 \ I) \circ (\pi_2 \ J) \circ \pi, & \text{otherwise.}
\end{cases}$$

(18)

The coupling just described defines the quantity $V_{\alpha} := E|W - V_{\alpha}|$, $\alpha = (1, 2)$, which is the total expected number of changes in conflict by forcing a conflict to occur between
rooks 1 and 2. Recall that $X_\alpha$ is the indicator random variable that rooks 1 and 2 are attacking. Let $X'_\alpha$ denote the the same quantity under the coupling. Then

\begin{equation}
V_\alpha = \sum_{\beta \neq \alpha} X'_\beta, \quad W = \sum_{\beta} X_\beta,
\end{equation}

and we have

\begin{equation}
|W - V_\alpha| = |X_\alpha + \sum_{\beta \neq \alpha} (X_\beta - X'_\beta)|.
\end{equation}

Now we need to consider the different values of this quantity, which depend on whether swapped rooks were visible or invisible.

**Case 1:** No other rooks are involved since there is no swap, therefore $|W - V_\alpha| = X_\alpha = 1$.

**Case 2a:** WLOG, assume we are swapping rook 2 into location $J$. In this case we assume there was a visible rook in $J$, so when we swap there are no changes in the configuration, therefore $|W - V_\alpha| = X_\alpha = 1$.

**Case 2b:** WLOG, assume we are swapping rook 2 into location $J$. In this case we assume there was an invisible rook in $J$, so when we swap there are a random number of new attacking rook pairs. We will estimate this quantity later.

**Case 3a:** We assume further that there were two visible rooks, say rooks $a$ and $b$, in locations $I$ and $J$. In this case, no new attacking pairs are formed by swapping, hence $|W - V_\alpha| = X_\alpha = 1$.

**Case 3b:** We assume that there was one visible rook and one invisible rook in locations $I$ and $J$. This case is similar to Case 2b.

**Case 3c:** We assume that there were two invisible rooks in locations $I$ and $J$. In this case, we expect at most twice the number of new attacking pairs as in cases 2b and 3b.
Case 1:

\[
\begin{array}{c|c|c|c|c|c|c}
1 & 2 & \cdots & I & J & \cdots & n \\
\end{array}
\mapsto
\begin{array}{c|c|c|c|c|c|c}
1 & 2 & \cdots & I & J & \cdots & n \\
\end{array}
\]

Case 2a:

\[
\begin{array}{c|c|c|c|c|c|c}
1 & a & \cdots & 2 \\
\end{array}
\mapsto
\begin{array}{c|c|c|c|c|c|c}
1 & 2 & \cdots & a \\
\end{array}
\]

Case 2b:

\[
\begin{array}{c|c|c|c|c|c|c}
1 & \cdots & 2 \\
\end{array}
\mapsto
\begin{array}{c|c|c|c|c|c|c}
1 & 2 & \cdots & + \\
\end{array}
\]

Case 3a:

\[
\begin{array}{c|c|c|c|c|c|c}
a & b & 1 & 2 \\
\end{array}
\mapsto
\begin{array}{c|c|c|c|c|c|c}
a & b & 1 & 2 \\
\end{array}
\]

Case 3b:

\[
\begin{array}{c|c|c|c|c|c|c}
a & \cdots & 1 & 2 \\
\end{array}
\mapsto
\begin{array}{c|c|c|c|c|c|c}
a & + & 1 & 2 \\
\end{array}
\]

Case 3c:

\[
\begin{array}{c|c|c|c|c|c|c}
\cdots & \cdots & 1 & 2 \\
\end{array}
\mapsto
\begin{array}{c|c|c|c|c|c|c}
\cdots & \cdots & + & + \\
\end{array}
\]

Figure 2. Before and after positions of rooks 1 and 2 on the board relative to \( I \) and \( J \) and visible rooks \( a \) and \( b \). A * represents a location where a change in the number of conflicts can occur.

First, the contribution to the expectation \( v_\alpha \) for Cases 1, 2a, and 3a are simply the probabilities of each occurring. These are straightforward to calculate:

\[
P(\text{Case 1}) = \frac{2}{N(N-1)},
\]

\[
P(\text{Case 2a}) = \frac{2(k-2)}{N(N-1)},
\]

\[
P(\text{Case 3a}) = \frac{2(k-2)(k-3)}{N(N-1)}.
\]

For \( k = O(\sqrt{n}) \), these terms are \( O(1/n^2) \), which is a higher order than Case 3b, so while they do not affect the resulting asymptotic formula, they must be included in the explicit upper bound.

For the other cases, the possible changes in attacking pairs comes from those rooks that were in attacking configurations before the coupling (and were moved out of attacking), and those in attacking configurations occurring after the coupling (i.e., were moved into attacking). We thus decompose \( v_\alpha \) into positive and negative parts, namely,

\[
v_\alpha = \mathbb{E}(W - V_\alpha)^+ + \mathbb{E}(W - V_\alpha)^-.
\]
Under the coupling, we may have eliminated a previously attacking rook pair. Fix rook $a$, and define $\Gamma_a = \{ \delta = (s,t) : s = a \text{ or } t = a \}$. Then only random variables $Y_\delta$, where $\delta \in \Gamma_a$ are affected. For $\alpha, \gamma \in \Gamma$, let $R_\alpha^\alpha$ (respectively, $C_\alpha^\alpha$) denote the random variable $Y_\alpha^\alpha$ further conditioned on the coupled rooks occurring in the same row (respectively, column). We have for each $\alpha \in \Gamma$,

$$
\mathbb{E} (Y_\delta - R_\delta^\alpha)^+ \leq (k-1)q_\alpha \left[ P(\text{Case 2b}) + P(\text{Case 3b}) + 2P(\text{Case 3c}) \right].
$$

The factor of 2 in the last term comes from the fact that the coupling swaps 2 rooks, whereas in the other cases we only swap 1. We can think about these cases separately, but instead, since $P(\text{Case 3c}) = 1 - O(k/N + k^2/N^2)$, we simply use the estimate

$$
[P(\text{Case 2b}) + P(\text{Case 3b}) + 2P(\text{Case 3c})] \leq 2,
$$

which is a hard upper bound with a higher order error term. Hence,

$$
\sum_{\delta \in \Gamma_a} \mathbb{E} (Y_\delta - R_\delta^\alpha)^+ \leq 2(k-1)q_\alpha.
$$

By symmetry, the corresponding quantity for columns is the same (i.e., replacing $R$ with $C$), hence we multiply by a factor of 2, and obtain

$$
\mathbb{E} (W - V_\alpha)^+ \leq \frac{8(k-1)(n)}{N}.
$$

The term $\mathbb{E} (W - V_\alpha)^-$ is more complicated since it depends on $I$ and $J$; it is the expected number of new attacking rook pairs conditioned on the pair of rooks $\alpha = (a,b)$ in attacking configuration. There are two cases to consider when the attacking rooks are coupled to attack on the same row:

(A) Any other rook $s$ that attacks in the same row will now contribute 2 additional attacks, one for each pair $\{s,a\}$ and $\{s,b\}$. See Figure 3.

(B) An attacking rook pair in the same column as either of the two coupled rooks will still contribute just 1 attack. See Figure 4.

By symmetry, the calculations are the same when the attacking rooks are coupled to attack on the same column. We have for all $\alpha, \gamma \in \Gamma$,

$$
\mathbb{E} (Y_\gamma - Y_\gamma^\alpha)^- = \mathbb{E} (Y_\gamma - R_\gamma^\alpha)^- + \mathbb{E} (Y_\gamma - C_\gamma^\alpha)^- = 2\mathbb{E} (Y_\gamma - R_\gamma^\alpha)^-,
$$

where the last equality follows by symmetry. We consider each case starting with Case (A): the probability of the pair $\alpha = (a,b)$ conditioned on attacking in row $r$ is $\binom{r}{2}/\binom{n}{3}$.

Once this row is determined, there are $r-2$ remaining spaces for another rook to appear. Hence, we have

$$
2(k-2) \sum_{r=2}^{n-1} \frac{\binom{r}{2}}{\binom{n}{3}} \frac{r-2}{\binom{n}{2}} = 2(k-2) \frac{3\binom{n}{4}}{\binom{n}{3} \binom{n}{2}}.
$$

7Instead of letting pairs of rooks lie in any of the $N$ possible pairs of positions, they can only appear pairwise in a row, proportional to the number of pairs of positions in that row. This implies $C \sum_{r=2}^{n-1} \binom{r}{2} = 1$, hence $C = \binom{n}{3}^{-1}$. 


In Case (B) we have to sum over all possible positions of \( i \) and \( j \) within each row. The calculation is as follows:

\[
(k - 2) \frac{1}{(\binom{n}{3})} \sum_{r=2}^{n-1} \sum_{c_1=1}^{r-1} \sum_{c_2=c_1+1}^{r} \left[ \frac{n - 1 - c_1}{\binom{r}{2} - 2} + \frac{n - 1 - c_2}{\binom{r}{2} - 2} \right]
\]

\[
= \frac{k - 2}{\binom{n}{3}} \left( 2(n - 1) \binom{n}{3} - \binom{n}{3} \frac{n + 1}{4} - 2 \left( \binom{n}{4} \right) \right).
\]

\[
= \frac{k - 2}{\binom{n}{2}} \left( 2(n - 1) - \frac{n + 1}{4} - \frac{n + 1}{2} \right).
\]

\[
= \frac{(k - 2)}{\binom{n}{2}} \left( \frac{5n - 11}{4} \right).
\]

Summarizing (by combining Equations (27) and (28) and multiplying them by 2 for symmetry in columns), we have

\[
(k - 2) \frac{12 \binom{n}{4}}{\binom{n}{3}(\binom{n}{2} - 2)} + \frac{k - 2}{\binom{n}{2}} \left( \frac{5n - 11}{2} \right).
\]

The final step is to combine Equations (26), (29), along with the higher order estimates in Equations (21), (22), (23), to obtain

\[
v_\alpha \leq \frac{8(k - 1) \binom{n}{3}}{N} + (k - 2) \frac{12 \binom{n}{4}}{\binom{n}{3}(\binom{n}{2} - 2)} + \frac{(k - 2)}{\binom{n}{2}} \left( \frac{5n - 11}{2} \right)
\]

\[
+ \frac{2}{N(N - 1)} + \frac{2(k - 2)}{N(N - 1)} + \frac{2(k - 2)(k - 3)}{N(N - 1)}.
\]

Finally, one observes that our quantities do not depend on the particular rooks chosen, hence for \( \alpha' = (1, 2) \), we have

\[
\sum_\alpha q_\alpha v_\alpha = \frac{(k - 2)}{2} q_{\alpha_0} v_{\alpha_0} \leq \mu v_{\alpha_0}.
\]

5. Remarks

5.1. An alternative coupling. Instead of the setup where no two rooks may lie in the same square, we can instead place rooks independently on the board so that it is possible that two rooks occupy the same square. A similarly defined coupling that moves two rooks, say 1 and 2, into positions \( I \) and \( J \) could then simply move them instead of swapping them. In this case, however, we would need to condition on the event \( D := \{\text{all rooks lie in distinct squares}\} \), and while this presents no technical difficulty, it nonetheless diminishes the utility of attempting to simplify the proof of Theorem 5.
Nevertheless, using this coupling we can obtain similar results by a transformation of the following permutation coupling quantities:

\[
\begin{align*}
EW & \mapsto \frac{M-1}{M} EW \\
\binom{n}{2} & \mapsto \binom{n}{2} \frac{k}{k!} \\
d_3 & \mapsto \frac{M-1}{M} d_3.
\end{align*}
\]

In fact, in \(d_3\), instead of multiplying the entire term by \((M - 1)/M\), the middle two terms in the summation can be multiplied by \((M - 2)/M\) instead.

Numerical calculations for small \(n\) and \(k\) suggest that this bound is sometimes more optimal, for example, with \(n = 10\) and \(k = 2\), we obtain with the permutation bound

\[S(10, 8) \in [310.937, 1242.82],\]
whereas using the independence coupling, we obtain

\[ S(10, 8) \in [332.572, 1265.08]. \]

The true value is \( S(10, 8) = 750 \), and in this case the independence coupling has a more accurate lower bound and a smaller overall width, whereas the permutation coupling has a more accurate upper bound.

It would be tempting to conjecture that one of these bounds may be more optimal as a lower bound, and the other as an upper bound. While this may indeed be the case for a few small values of the parameters, numerical calculations for much larger values of \( n \) and \( k \) suggest that the permutation bound is overall better, both as an upper and lower bound, though naturally for such large values the difference between the two intervals becomes negligibly small. For a concrete example, consider the case when \( n = 10^{10} \) and \( k = 10^5 \).

We have from the permutation coupling,

\[ S(n, n-k)/10^{1513322} \in [9.32741404741295, 9.35813076701713] \]

whereas for the independence coupling, we have

\[ S(n, n-k)/10^{1513322} \in \{[9.32741403567849, 9.35813076944439] \}

5.2. **An alternative bound.** One can also define an index set \( J = \{(i, i+1), i = 1, 2, \ldots, n-1\} \), and random variable \( X_\alpha, \alpha \in J \), as the indicator random variable that there is a pair of attacking rooks in either row \( i \) or column \( i+1, i = 1, 2, \ldots, n-1 \); i.e., \( X_\alpha \) is the indicator that the diagonal coordinate \((i, i+1)\) detects at least one pair of attacking rooks. This will in general undercount the number of pair-wise attacking rooks, since for example any collection of three or more rooks in the same row will only be counted as one set of attacking rooks rather than \( \binom{3}{2} \) separate pair-wise attacks. The argument in this case is almost the same, but the bounds are somewhat different.

**Theorem 6.** Assume \( n \geq 3, n \geq k \geq 1 \). Let \( M = \binom{n}{2} \) and \( N = \binom{M}{2} \). Define

\[
\lambda_+ := \frac{\binom{k}{3}}{N},
\]

\[
\lambda_- := \max\left(0, \lambda_+ - 3 \frac{\binom{k}{3}}{\binom{M}{3}}\right),
\]

\[
d_1 := \frac{2\lambda_+^2}{k-1} + \frac{\binom{k}{3} \frac{(n-1)^4}{5}}{N} + \frac{k(k-4)+5}{N},
\]

\[
d_2 := \frac{2\lambda_+}{k-1} + \frac{1}{\lambda_-} \frac{\binom{k}{3} \frac{(n-1)^4}{5}}{N} + \frac{k(k-4)+5}{N},
\]

and \( C_{n,k} := \min\left(\min(d_1, d_2), 1\right) \).

\[\text{If } \lambda_- = 0 \text{ then we set } d_2 = +\infty.\]
Then we have

\[
\binom{n}{k} 2^{\lambda} \left( 1 - 2e^{2\lambda} C_{n,k} \right) \leq S(n, n - k) \leq \binom{n}{k} 2^{\lambda} \left( 1 + 2e^{2\lambda} C_{n,k} \right).
\]

We omit the proof since it follows using a similar counting method, and numerical calculations suggest that it is altogether inferior to the one in Theorem 5, although we have not carried out a formal algebraic proof. Since the error terms have the same order of magnitude, and because it is an asymmetrical bound, and because it contains a messy expression in the denominator of \(d_2\), we do not investigate further.

5.3. An historical note on small values of \(m\). For values of \(m\) fixed as \(n \to \infty\), Jordan [8] is credited by Moser and Wyman [12] for the asymptotic formula

\[
|s(n, m)| \sim \frac{(n - 1)!}{(m - 1)!} (\log n + \gamma)^{m-1}.
\]

Moser and Wyman [12] extended this first-order asymptotic formula to values of \(m = o(\log n)\). Wilf [18] extended this formula into an asymptotic expansion valid for \(m = O(1)\), and Hwang [7] extended Wilf’s asymptotic expansion to \(m = O(\log(n))\).

5.4. A more optimal Stein’s method. We have used the trivial fact that

\[
|P(X = 0) - P(Y = 0)| \leq d_{TV}(X, Y)
\]
in order to derive our bound using the Chen-Stein method, but there are other variations that specifically address bounding of point probabilities. For example, in Section 2.4 of [1], we have the following.

**Theorem 7** ([1]). Suppose we are in the setup of Theorem 4. For each \(\beta \in \Gamma\), pick any \(k_\beta, K_\beta \geq 1\). Define

\[
\epsilon = \max_{\beta \in \Gamma} P \left( |W - V_\beta| \geq k_\beta, W \leq K_\beta \right) / \mathbb{E} |W - V_\beta|,
\]

\[
\eta = \max_{\beta \in \Gamma} P \left( W > K_\beta \right) / \mathbb{E} |W - V_\beta|,
\]

\[
\mu = \max_{\beta \in \Gamma} \sum_{0 < |\ell| \leq k_\beta} |\ell| \max_{0 \leq r \leq K_\alpha} P \left( V_\beta = r + \ell | W = r \right) / \mathbb{E} |W - V_\beta|.
\]

Then we have

\[
|P(W = j) - P(P = j)| \leq 2\psi_j \left( \epsilon + \eta + \mu (2e\lambda)^{-1/2} \right),
\]
as long as \(\mu \psi \leq 1/2\), where

\[
\psi_j = \min(\lambda^{-1}, j^{-1}, 1) \sum_{\beta \in \Gamma} p_\beta \mathbb{E} |W - V_\beta|.
\]
and

\[ \psi = \max_j \psi_j = \psi_1. \]

There are many choices for \(k_\beta\) and \(K_\beta\), and we are not prepared to speculate on the best choice, nor have we worked out the details, but we believe the calculations should be more or less straightforward and provide better bounds.

5.5. **Rook and File numbers.** We have taken our board \(B\) to be the staircase board, i.e., the board with rows of size \(n - 1, n - 2, \ldots, 2, 1\). One can generalize this to a board with rows of integer lengths \(\lambda_1, \lambda_2, \ldots\), such that \(\lambda_1 \geq \lambda_2 \geq \ldots \lambda_\ell \geq 1\). Such a sequence of positive integers is called an integer partition, and the board \(B\) is called a Ferrers board. The column sizes are given by the conjugate partition to \(\lambda\), which we denote by \(\lambda' = (\lambda'_1, \lambda'_2, \ldots)\), where \(\lambda'_i = \#\{j : \lambda_j \geq i\}\). The rook number \(r_B(k)\) is defined as the number of ways of placing \(k\) non-attacking rooks on board \(B\), and the file number \(f_B(k)\) is defined as the number of ways of placing \(k\) non-attacking rooks on board \(B\) such that attacks only occur along columns.

In this setting, the permutation coupling is the same, although the precise calculations are slightly more general. Nevertheless, the proof is no more difficult, perhaps only more cryptic in that it includes unsimplified sums, and so we simply state the analogous theorem below.

**Theorem 8.** Given a Ferrers board \(B\) with row lengths given by the integer partition \(\lambda = (\lambda_1, \lambda_2, \ldots)\), define \(M = \sum_i \lambda_i\), \(N = \binom{M}{2}\), \(L = \sum_i \binom{\lambda_i}{2}\), \(L' = \sum_i \binom{\lambda'_i}{2}\),

\[
\mu := \frac{(k) L}{N}, \quad \mu' := \frac{(k) L'}{N},
\]

\[
b_1 := \frac{4(k - 1)(L)}{N}, \quad b'_1 := \frac{4(k - 1)(L')}{N},
\]

\[
b_2 := (k - 2) \left[ \sum_i \frac{\binom{\lambda_i}{2}}{L} \frac{\lambda_i - 2}{(M - 2)} \right], \quad b'_2 := (k - 2) \left[ \sum_i \frac{\binom{\lambda'_i}{2}}{L'} \frac{\lambda'_i - 2}{(M - 2)} \right],
\]

\[
b_3 := (k - 2) \frac{1}{L} \sum_{i_1=1}^{\lambda_1-1} \sum_{c_1=1}^{\lambda_1-1} \sum_{c_2=1}^{\lambda_2} \left[ \frac{\lambda'_1 - 1}{(M - 2)} + \frac{\lambda'_2 - 1}{(M - 2)} \right],
\]

\[
b'_3 := (k - 2) \frac{1}{L'} \sum_{i_1=1}^{\lambda'_1-1} \sum_{c_1=1}^{\lambda'_1-1} \sum_{c_2=1}^{\lambda'_2} \left[ \frac{\lambda_1 - 1}{(M - 2)} + \frac{\lambda_2 - 1}{(M - 2)} \right],
\]

\[
b_4 := \frac{k(k - 4) + 5}{N},
\]

\[
d_3 := b_1 + b_2 + b_3 + b_4, \quad d'_3 := b'_1 + b'_2 + b'_3 + b'_4,
\]

\[
d_4 := \mu d_3, \quad d'_4 := \mu d'_3,
\]

and

\[
D_{n,k} := \min (d_3, d_4, 1), \quad D'_{n,k} := \min (d_3, d_4, 1),
\]
Then we have
\[
\binom{M}{k} e^{-\mu - \mu'} \left( 1 - e^{\mu + \mu'} (D_{n,k} + D'_{n,k}) \right) \leq r_B(k) \leq \binom{M}{k} e^{-\mu} \left( 1 + e^{\mu + \mu'} (D_{n,k} + D'_{n,k}) \right).
\]

\[
\binom{M}{k} e^{-\mu'} \left( 1 - e^{\mu'} (D'_{n,k}) \right) \leq f_B(k) \leq \binom{M}{k} e^{-\mu'} \left( 1 + e^{\mu'} (D'_{n,k}) \right).
\]

It should be noted that for many choices of row lengths \( \lambda \), the bounds will be trivial. It would be interesting to investigate conditions on the row lengths \( \lambda \) such that the Poisson approximation holds.

5.6. Generalizations and Extensions. There are extensions of Stirling numbers to complex-valued arguments \[3\]. We are not aware of any combinatorial arguments that would correspond to placement of rooks on a board, so we do not pursue this idea further.

One can also generalize Stirling numbers to include non-uniform weights on the squares and also one can generalize the board \( B \) so that the number of squares in each row is monotonically increasing, see for example \[13\] and the references therein. In principle one can still apply Poisson approximation to these cases, although instead of simple counting arguments one obtains weighted sums.

6. Acknowledgements

SD would like to thank Amir Dembo, Jim Haglund, Igor Pak, and Bruce Rothschild for helpful discussions.

References

[1] A.D. Barbour, L. Holst, and S. Janson. Poisson approximation. Oxford science publications. Clarendon Press, 1992.
[2] R. Chelluri, L. B. Richmond, and N. M. Temme. Asymptotic estimates for generalized stirling numbers, 1999.
[3] Philippe Flajolet and Helmut Prodinger. On stirling numbers for complex arguments and hankel contours. SIAM J. Discrete Math., 12(2):155–159, 1999.
[4] A.M Garsia and J.B Remmel. Q-counting rook configurations and a formula of frobenius. Journal of Combinatorial Theory, Series A, 41(2):246 – 275, 1986.
[5] G. Grimmett and D. Stirzaker. Probability and Random Processes. Probability and Random Processes. OUP Oxford, 2001.
[6] L. C. Hsu. Note on an asymptotic expansion of the nth difference of zero. The Annals of Mathematical Statistics, 19(2):pp. 273–277, 1948.
[7] Hsien-Kuei Hwang, Ecole Polytechnique, and Main Results. Asymptotic expansions for the stirling numbers of the first kind, 1994.
[8] C. Jordan. Calculus of Finite Differences. Chelsea Publishing Company, 1947.
[9] Irving Kaplansky and John Riordan. The problem of the rooks and its applications. Duke Mathematical Journal, 13(2):259–268, 06 1946.
[10] Guy Louchard. Asymptotics of the Stirling numbers of the second kind revisited. Applicable Analysis and Discrete Mathematics.
[11] Guy Louchard. Asymptotics of the Stirling numbers of the first kind revisited: a saddle point approach. Discrete Math. Theor. Comput. Sci., 12(2):167–184, 2010.
[12] L. Moser and M. Wyman. Asymptotic development of the Stirling numbers of the first kind. *J. London Math. Soc.*, 33:133–146, 1958.

[13] L. Moser and M. Wyman. Stirling numbers of the second kind. *Duke Mathematical Journal*, 25(1):29–43, 03 1958.

[14] Jeffrey B. Remmel and Michelle L. Wachs. Rook theory, generalized stirling numbers and (p, q)-analogues. *Electr. J. Comb.*, 11(1), 2004.

[15] Vladimir N. Sachkov. *Probabilistic methods in combinatorial analysis*. Translated by V. A. Vatutin. Reprint of the 1997 hardback ed. Cambridge: Cambridge University Press, reprint of the 1997 hardback ed. edition, 2011.

[16] Bruno Salvy and John Shackell. Symbolic asymptotics: Multiseries of inverse functions. *Journal of Symbolic Computation*, 27(6):543 – 563, 1999.

[17] R.P. Stanley. *Enumerative Combinatorics*. Number v. 1 in Cambridge studies in advanced mathematics. Cambridge University Press, 2002.

[18] Herbert S. Wilf. The asymptotic behavior of the stirling numbers of the first kind. *J. Comb. Theory, Ser. A*, 64(2):344–349, 1993.