

Probabilistic transforms for combinatorial urn models

Olgica Milenkovic

Electrical and Computer Engineering Department

University of Colorado, Boulder

Kevin J. Compton

Department of Electrical Engineering and Computer Science

University of Michigan, Ann Arbor

December 11, 2002

Abstract

In this paper, we present several probabilistic transforms related to classical urn models. These transforms render the dependent random variables describing the urn occupancies into independent random variables with appropriate distributions. This simplifies the analysis of a large number of problems for which a function under investigation depends on the urn occupancies. The approach used for constructing the transforms involves generating functions of combinatorial numbers. We also show, by using Tauberian theorems derived in this paper, that under certain conditions the asymptotic expressions of target functions in the transform domain are identical to the asymptotic expressions in the inverse-transform domain. Therefore, asymptotic information about certain statistics can be gained without evaluating the inverse transform.

1 Introduction

The theory of urn models has a rich history [13], [14], [26], [16], since many problems in the area of physics, communication theory, computer science, combinatorial analysis of algorithms can be described in terms of distributing balls (objects) into specified urns (locations), according to some predefined set of conditions. Among the most commonly encountered urn models in physics are the so called Maxwell-Boltzman, Bose-Einstein and Fermi-Dirac model [16]. In computer science, urn models are used for database performance evaluations [36] and for modeling and analyzing algorithms. Two well known examples of the latter kind are hashing and sorting algorithms [24]. In communication theory, some transmission channels can be described in terms of contagion urn models [2]. There are many problems in the area of network analysis that can be described in terms of urn models - see for example [32] and [15] for an application in optical switch dimensioning and [43] for applications in network rerouting problems. Finally, ubiquitous urn models like Polya's and Ehrenfest's model have found applications in almost every area of applied mathematics and physics [16, p.118].

Since for an urn model the distributions of balls usually have to satisfy a number of constraints, it becomes hard to answer questions concerning the occupancies of the urns. Most of the problems in analyzing these questions arise because the occupancies are not independent. On the other hand, the

dependence of the variables representing the occupancies enters only through the requirement that the number of distributed balls is fixed. This suggests “randomizing” the problem by redefining it in such a way as to make the distribution process stochastic, and the random variables specifying the occupancies independent. A well known method used to achieve this goal for the case of a uniform distribution of balls is the Poisson transform [20]. It replaces the variables describing the urn occupancies by independent Poisson random variables. The exact expression for a function of these variables can be determined by applying the inverse Poisson transform. The inverse Poisson transform can be evaluated both by algebraic and analytical methods [39]. Analytic poissonization and depoissonization have been successfully applied for a precise analysis of some well known compression algorithms, such as the Ziv-Lempel algorithm [27]. On the other hand, the Poisson transform is not a universally applicable tool, since many urn models do not imply a uniform distribution of the balls, in the sense that every possible ball is placed in any urn with equal likelihood. For every individual urn model, one has to develop an appropriate transform that would achieve the same goal as the Poisson transform - replace the dependent variables describing the occupancies by independent variables. We will demonstrate how to construct these transforms on some well known classical examples, described in many textbooks on combinatorial theory [37]. Our approach is based on making the number of balls a random variable that has a probability distribution defined in terms of the generating function of the total number of possible distributions of balls into urns. The transform of a function is defined as its average with respect to this distribution. The idea of constructing probability distributions from generating functions first appeared in quantum physics [28], while in [29] (and in a similar way in [5]) it was shown that variables obeying these distributions can be approximated by sums of independent variables, conditioned on having a fixed number of balls. Although the variables introduced in [5] have probability distributions that match some of the distributions presented in this work, our techniques differ substantially in their application and derivation. First, the results in [5] are *approximations*, since they are not based on transform techniques presented in this work. This is a consequence of the fact that the number of balls used in the analysis in [5] is assumed to be fixed, and not a random variable, which is of crucial importance for defining a transform. The approximation error in [5] is measured in terms of the total variation distance, and the variation distance depends on the choice of the parameters of the distributions. In our approach, the parameters of the distributions are variables of the transform. Hence, one can always obtain an *exact solution* for the problem, and this solution does not depend on the particular value the parameter takes. We will also show, by analyzing the properties of the inverse transform, that in certain cases the asymptotic form of a function in the transform domain is identical to the asymptotic form of the exact function, provided that the parameter of the distribution has an appropriate “optimal” value. The optimal value of this parameter becomes intuitively clear and can easily be deduced from the definition of the transform, without resorting to the use of probabilistic distance measures (as suggested in [5]). Other results related to the problem of

finding good approximations for dependent random variables in terms of independent random variables appeared in [19] and [35]. The results in [19] deal exclusively with partitions, and there is no reference to a transform technique, although the number of balls is assumed to be a random number. Hence, the results are of approximate or asymptotic nature, and the parameter of the distributions is a variable that is optimized to give good approximations in terms of the Prohorov distance. The results in [35] are of special interest, because we believe that they include the first definition of a probabilistic transform of the form described in this paper. The transform in [35] is valid only for permutations, and we will show in Section 5 that this transform can also be derived using the general approach followed in this paper.

This paper is organized as follows. Section 2 contains the problem statement, as well as the outline of the approach used for constructing the transforms and their inverses. Section 3 presents several theorems that show under what set of conditions the asymptotic behavior of a variable in the original domain is identical to its asymptotic behavior in the transform domain. Section 4 and 5 contain our main results. Section 4 describes transforms for urn models with labeled urns, while Section 5 presents transforms for urn models with unlabeled urns. These sections also describe the optimal choices for the parameters of the probability distributions used to specify the transform. Section 6 contains the description of two transforms connected to the cycle structure of random permutations. Section 7 contains the concluding remarks.

2 Problem statement

An urn model can be described in terms of a function mapping a set of balls to a set of urns [37]. The function is described in terms of allowed placements of balls into urns. For two finite sets \mathcal{D} and \mathcal{E} of cardinality n and N , respectively, we are interested in counting the number of all possible functions $f: \mathcal{D} \rightarrow \mathcal{E}$ that obey a certain set of constraints. In this setting, the cardinality of the inverse image of i , $f^{-1}(i)$, represents the occupancy of urn number i .

Assume that we want to distribute n balls that may be labeled (distinguishable) or unlabeled (indistinguishable) into N urns that also may be labeled or unlabeled, according to a certain set of rules. Let the number of ways to perform such a distribution be $D(n, N)$. By specifying for each pair (n, N) an equiprobable probability measure on the space of all possible $D(n, N)$ distributions of balls into urns, the urn occupancies become random variables. If the urns are labeled, let these random variables be denoted by X_i , $i = 1, \dots, N$, where the index specifies the urn number. Clearly, the random variables X_i are dependent since $X_1 + \dots + X_N = n$. If the urns are unlabeled, there is no loss of information if one distinguishes urns only by their occupancy. Hence, we can restrict our attention only to the variables Y_i counting the number of urns with occupancy i . This changes the analysis in so far that $Y_1 + 2Y_2 + 3Y_3 + \dots + nY_n = n$ rather than $X_1 + \dots + X_N = n$ holds, and additionally $Y_1 + Y_2 + \dots + Y_n = N$. Therefore, the combinatorial model for the case of unlabeled urns has a different combinatorial structure and interpretation than the

case of labeled urns. Assume next that our goal is to compute a statistic $g(n, N)$ associated with an urn model that depends on the distributions of X_1, X_2, \dots, X_N (or Y_1, Y_2, \dots, Y_n , for unlabeled urns). We will refer to $g(n, N)$ as the *target statistic* of the model. Unless the target statistic is of a very simple form, evaluating it can be a very difficult task. In order to overcome this problem, one can proceed in the following way: instead of analyzing the problem for a fixed number of balls n , one can assume that the number of balls is a random variable with a discrete probability distribution that is associated with the urn model and has parameters that depend both on n and N . This implies that the target statistic $g(k, N)$ can be averaged over all possible values of k , with respect to an appropriately chosen distribution that depends on the parameters n and N . Under certain conditions, this average can be shown to be a good approximation of $g(n, N)$. We will show in the next sections that one can obtain exact results for $g(n, N)$ rather than approximations by using a probabilistic transform $T : g(n, N) \rightarrow \gamma(n, N)$, constructed from the number of possible distributions $D(n, N)$. The transform maps a space of statistics into another space of statistics, and one aspect of the mapping is that the occupancy variables in the transform domain may exhibit different properties than the variables in the original domain. Hence, the key for the success of the transform technique depends on two factors. Firstly, since the transform represents an average, the probability distribution with respect to which the averaging is performed has to assure that the random variables X_1, \dots, X_N (Y_1, Y_2, \dots, Y_N for unlabeled urns) become independent random variables $Z_i, i = 1, 2, \dots, N$ (V_1, V_2, \dots for unlabeled urns) in the transform domain. Secondly, the inverse transform has to be tractable for either exact evaluation or asymptotic analysis.

Let $\Omega(n, N)$ be the sample space of all possible distributions of n balls associated with an N -urn model, and let the probability measure on this space be uniform. Clearly, $D(n, N) = |\Omega(n, N)|$. If the balls are to be considered labeled, let $f(x)$ be the exponential generating function of $D(k, N)$ with respect to the parameter k ,

$$f(x) = \sum_{k=0}^{\infty} D(k, N) \frac{x^k}{k!}, \quad (1)$$

and if the balls are to be considered unlabeled, let $f(x)$ be the ordinary generating function of $D(k, N)$ with respect to k ,

$$f(x) = \sum_{k=0}^{\infty} D(k, N) x^k. \quad (2)$$

In general, define a generating function of the form

$$f(x) = \sum_{k=0}^{\infty} F_k(N) x^k, \quad (3)$$

where

$$F_k(N) = \frac{D(k, N)}{q(k)},$$

$q(k) = k!$, if the balls are labeled, and $q(k) = 1$, if the balls are unlabeled. Let λ be a real number for which the generating function $f(\lambda)$ converges. Define the probability distribution of a random variable

X in terms of the generating function $f(\lambda)$ as

$$P\{X = n\} = \frac{F_n(N)\lambda^n}{f(\lambda)}. \quad (4)$$

Equation (4) defines a discrete probability distribution on the non-negative integers. We will refer to this distribution as the *defining distribution* of the transform. The defining distribution represents, in some sense, the probability of choosing a distribution with n balls among distributions with any number of balls; one has to resort to the use of generating functions, since the total number of distributions is not finite. Let the transform $\gamma(\lambda, N)$ of the target statistic $g(n, N)$ be its average with respect to the defining distribution:

$$\gamma(\lambda, N) = \sum_{k=0}^{\infty} P\{X = k\} g(k, N) = \sum_{k=0}^{\infty} \frac{F_k(N) \lambda^k}{f(\lambda)} g(k, N). \quad (5)$$

From (5) it follows that

$$f(\lambda) \gamma(\lambda, N) = \sum_{k=0}^{\infty} F_k(N) g(k, N) \lambda^k, \quad (6)$$

so that it is straightforward to recover $g(n, N)$ (i.e. compute the inverse transform) as

$$g(n, N) = \frac{1}{F_n(N)} [\lambda^n] \{f(\lambda) \gamma(\lambda, N)\}, \quad (7)$$

with $[\lambda^n] h(\lambda)$ denoting the coefficient of λ^n in the series $h(\lambda)$.

If $\gamma(\lambda, N)$ is analytic in some neighborhood of the origin, then it has a Taylor series expansion, say of the form

$$\gamma(\lambda, N) = \sum_{m=0}^{\infty} \Gamma_m(N) \lambda^m. \quad (8)$$

This implies that the target function $g(n, N)$ in this case can be computed as

$$g(n, N) = \sum_{m=0}^n \Gamma_m(N) \frac{F_{n-m}(N)}{F_n(N)}. \quad (9)$$

The probability generating function $G(z)$ of the defining distribution is given by

$$G(z) = \sum_{k=0}^{\infty} P\{X = k\} z^k = \frac{f(\lambda z)}{f(\lambda)}, \quad (10)$$

and we will prove in Lemma 1 that for the case of labeled urns, $G(z)$ is the product of N identical probability generating functions, i.e.

$$G(z) = G_{\lambda}^N(z), \quad (11)$$

and an infinite product of probability generating functions with different parameters for the case of unlabeled urns, i.e.

$$G(z) = \prod_{k=1}^{\infty} G_{\lambda^k}(z^k). \quad (12)$$

This implies that there exist independent random variables Z_i in the transform domain, such that $X = Z_1 + \dots + Z_N$ (for the case of labeled urns), and independent random variables V_i , such that $X = V_1 + 2V_2 + 3V_3 + \dots$ (for the case of unlabeled urns).

A legitimate point is to ask why the transforms described by Equation (5) are useful. The answer to this question will be provided in Lemma 1, Proposition 2 and Corollary 3. These results will show that $\gamma(\lambda, N)$ can be interpreted as the target statistic for a model for which the occupancy variables are the independent variables Z_i (or the independent variables V_i). The probability distributions of the variables Z_i and V_i depend on the form of the defining distribution and the parameter λ .

Lemma 1. *Equations (11) and (12) hold for all probability generating functions of defining distributions constructed from the combinatorial numbers $D(n, N)$.*

Proof: We start by proving the claimed result for the case of labeled urns. As described before, let \mathcal{S} be the set of allowed occupancy values for labeled urns, and consider the following function

$$\left(\sum_{n \in \mathcal{S}} \frac{\lambda^n}{q(n)} \right)^N. \quad (13)$$

Expanding (13) into a product gives

$$\sum_{n_1 \in \mathcal{S}} \dots \sum_{n_N \in \mathcal{S}} \frac{\lambda^{n_1 + \dots + n_N}}{q(n_1) \dots q(n_N)}.$$

Hence,

$$[\lambda^n] \left(\sum_{n \in \mathcal{S}} \frac{\lambda^n}{q(n)} \right)^N = \sum_{\substack{n_1 + \dots + n_N = n, \\ n_i \in \mathcal{S}}} \frac{1}{q(n_1) \dots q(n_N)} = \frac{1}{q(n)} \sum_{\substack{n_1 + \dots + n_N = n, \\ n_i \in \mathcal{S}}} \frac{q(n)}{q(n_1) \dots q(n_N)},$$

where it can be easily seen that

$$\sum_{\substack{n_1 + \dots + n_N = n, \\ n_i \in \mathcal{S}}} \frac{q(n)}{q(n_1) \dots q(n_N)} = D(n, N).$$

This shows that

$$\left(\sum_{n \in \mathcal{S}} \frac{\lambda^n}{q(n)} \right)^N = f(\lambda),$$

where $f(\lambda)$ is the generating function given by (3). From the definition of the generating function $f(\lambda)$ and the previous equation it follows that

$$G(z) = \frac{f(\lambda z)}{f(\lambda)} = \left(\frac{\sum_{n \in \mathcal{S}} \frac{(\lambda z)^n}{q(n)}}{\sum_{n \in \mathcal{S}} \frac{\lambda^n}{q(n)}} \right)^N, \quad (14)$$

and this implies that

$$G_\lambda(z) = \frac{\sum_{n \in \mathcal{S}} \frac{(\lambda z)^n}{q(n)}}{\sum_{n \in \mathcal{S}} \frac{\lambda^n}{q(n)}}.$$

This completes the proof of the claimed result for the case of labeled urns.

For the case of unlabeled urns, we will restrict our attention only to the case when there is no restriction on the number of urns N (see Section 3). For the case when there is a predefined value of N one would have to generalize this result in terms of a generating function in two variables, and the proof of this result is straightforward. Additionally, we will assume that for the case of unlabeled urns, \mathcal{S} stands for the set of all possible values for the number of urns with a given occupancy. Hence, if we consider

$$\prod_{k=1}^{\infty} \left(\sum_{n \in \mathcal{S}} \frac{1}{q(n)} \left(\frac{\lambda^k}{q(k)} \right)^n \right),$$

we obtain, similarly as for the case of labeled urns, that

$$G_{\lambda^k}(z^k) = \frac{\sum_{n \in \mathcal{S}} \frac{1}{q(n)} \left(\frac{(\lambda z)^k}{q(k)} \right)^n}{\sum_{n \in \mathcal{S}} \frac{1}{q(n)} \left(\frac{\lambda^k}{q(k)} \right)^n}.$$

This completes the proof of the lemma. ■

At this point, we pause to observe that there exist interesting connections between the combinatorial result of Lemma 1 and the characterization of infinitely divisible distributions. Warde and Katti [41] give very simple sufficient conditions for a discrete random variable X to be infinitely-divisible. This condition amounts to requiring $P\{X = n\}$ to be a log-convex sequence. Therefore, even without the result given in Lemma 1, one could deduce divisibility properties of the defining distribution by examining log-concavity/convexity properties of the sequence $F_n(N)$ with respect to n . For some combinatorial numbers, connected to urn models to be considered in this paper, questions regarding log-convexity or log-concavity have been settled only in two very recent papers [6],[7]. The authors of [7] proved that provided t_j is a log-concave sequence of non-negative real numbers, and provided the sequences a_n and b_n are defined by

$$\sum_{n=0}^{\infty} a_n u^n = \sum_{n=0}^{\infty} b_n \frac{u^n}{n!} = \exp \left(\sum_{j=1}^{\infty} \frac{t_j u^j}{j} \right), \quad (15)$$

it holds that a_n is a log-concave and that b_n is a log-convex sequence. It is interesting to point out that this is not the only place where the log-concavity/convexity properties of the numbers $F_n(N)$ will be used. In Sections 2, 3, 4 we will show that the validity of approximating the exact asymptotic formula by one obtained in the transform domain depends directly on the log-concave/convex behavior of $F_n(N)$ and hence on the results of [6] and [7].

It is crucial to construct the defining distribution with respect to $D(n, N)$. For if we were not to use $D(n, N)$, the transform would not allow for recovering the exact result. We will prove this claim and give a useful interpretation of the probabilistic transform (5) in the following proposition. For the case of labeled urns, we will use $P_x(m_1, \dots, m_N)$ to describe the probability

$$P\{X_1 = m_1, \dots, X_N = m_N\}, \quad (16)$$

while for the case of unlabeled urns we will use $P_y(m_1, m_2, \dots)$ to describe

$$P\{Y_1 = m_1, Y_2 = m_2, \dots\}. \quad (17)$$

Similarly, we will use $P_z(m_1, \dots, m_N)$ and $P_v(m_1, m_2, \dots)$ to describe the same probabilities with respect to the transform variables Z_i , and V_i , $i = 1, 2, \dots$, respectively.

Proposition 2. *For the case of labeled urns, the defining distribution of X given by (4) and the distributions of Z_i computed from (10) satisfy*

$$P\{Z_1 = m_1, \dots, Z_N = m_N / X = n\} = P_x(m_1, \dots, m_N), \quad (18)$$

where

$$X = Z_1 + \dots + Z_N. \quad (19)$$

The same result holds for the case of unlabeled urns, with $P_x\{m_1, m_2, \dots\}$ replaced by $P_y\{m_1, m_2, \dots\}$, X replaced by $V_1 + 2V_2 + 3V_3 + \dots$ and the Z_i 's replaced by V_i 's.

Remark: The result of Proposition 2 shows that the transform technique allows for exact recovery of the joint distribution of the X_i 's or Y_i 's, since the joint distribution of the independent variables Z_i or V_i conditioned on their weighted sum being equal to n produces the original joint distribution of the X_i 's or Y_i 's.

Proof: From the definition and the independence property of the variables Z_i it follows that

$$P_z(m_1, \dots, m_n) = [z^{m_1}]G_\lambda(z) \dots [z^{m_N}]G_\lambda(z).$$

The defining probability is chosen in such a way as to ensure that

$$\begin{aligned} [z^{m_1}]G_\lambda(z) \dots [z^{m_N}]G_\lambda(z) &= P_x(m_1, \dots, m_N) [z^\mu] G_\lambda^N(z), \\ &= P_x(m_1, \dots, m_N) [z^\mu] G(z) \end{aligned}$$

$$\text{where } \mu = m_1 + \dots + m_N, \quad G(z) = G_\lambda^N(z)$$

holds (notice that by extracting one coefficient instead of N coefficients with pre-described exponents corresponds to randomly choosing one possible occupancy distribution with pre-described occupancies out of all possible distributions). The above equation shows that

$$P_z(m_1, \dots, m_N) = P_x(m_1, \dots, m_N) [z^\mu] G(z) = P_x(m_1, \dots, m_N) P\{X = \mu\}.$$

For $\mu = n$, it holds that

$$P_z(m_1, \dots, m_N / X = n) = \frac{P\{Z_1 = m_1, \dots, Z_N = m_N, X = n\}}{P\{X = n\}} = P_x(m_1, \dots, m_N),$$

which proves the first claimed result. The result for the case of unlabeled urns can be proved along the same lines. ■

Corollary 3. *The statistic $\gamma(\lambda, N)$ defined by (5) is the value of the target statistic under the condition that the urn occupancies Z_i (for labeled urns) or V_i (for unlabeled urns) are independent random variables, with distributions that can be determined from the factorization of the probability generating function given by (10).*

Proof: The result of Proposition 2 shows that conditioned on $Z_1 + \dots + Z_N = n$ (or $V_1 + 2V_2 + 3V_3 + \dots = n$ for unlabeled urns), the joint distribution of the occupancies arising from the independent variables Z_i (V_i , for unlabeled urns) is equal to the joint distribution of the X_i 's (Y_i 's for unlabeled urns). If a statistic $\gamma(\lambda, N)$ is generated with respect to independent urn occupancies, then the probability of $g(n, N)$ contributing to this statistic is exactly $P\{X = n\}$. This shows that averaging the target statistic over all values of n , with respect to the defining distribution, gives a statistic for which the occupancies are independent and there is no restriction on their sum. ■

For example, suppose that Ψ is some statistic dependent on the distribution of the urn occupancies. If $E_z[\Psi]$ is the average of Ψ computed with respect to the occupancies being independent variables Z_1, Z_2, \dots , then

$$\begin{aligned} E_z[\Psi] &= E_X[E_z[\Psi/X]] = \sum_{n=0}^{\infty} P\{X = n\} E_z[\Psi/X = n] \\ &= \sum_{n=0}^{\infty} P\{X = n\} E[\Psi(n, N)], \end{aligned} \tag{20}$$

where $E[\Psi(n, N)]$ is the expected value of Ψ with respect to the joint distribution of X_1, X_2, \dots and the last equality follows from Proposition 2. But Equation (20) is exactly the definition of the transform (5), and consequently $\gamma(\lambda, N)$ defined by (5) is the value of the desired function *under the condition that the urn occupancies are independent random variables*, with distributions that can be determined from (10). The observation in Corollary 3 is crucial for defining the probabilistic transforms, and it is the key difference between the approach taken in [5] and the approach taken in this paper.

Remark: From (7) it follows that the results obtained in the transform domain correspond either to the exponential or ordinary generating function of the variable $g(n, N)$. Therefore, using the transform approach it becomes very easy to compute generating functions of the number of distributions of the balls under investigation; the most complicated task is usually to compute the inverse transform.

3 Asymptotic inversion

In Section 2, we showed that evaluating the inverse transform can be reduced to extracting a coefficient in the product of a generating function and a statistic $\gamma(\lambda, N)$. There exist many well known results, such as the Tauberian theorems by Schur [10], Compton [11] and Woods [44], that establish asymptotic expressions for the coefficient of the product of two generating functions based on the asymptotic expressions of the terms of the contributing generating functions. Using a similar type of argument as for the

derivation of Compton's Tauberian theorem [11], [10], we will prove several results for computing asymptotic formulas for the inverse transforms. The proofs presented in this paper are based on the assumption that N grows proportionally with n . As N grows, the singularities of the generating functions $f(\lambda)$ become "larg" and asymptotic information about the coefficients can be obtained by using the saddle point method. The concept of a large singularity is not a precise one, although it is extensively used in literature [33]. It basically indicates how fast a function grows as its argument approaches the singularity point. For example, $z = 1$ is a small singularity for $(1 - z)^{-1}$ and a large singularity for $\exp((1 - z)^{-1})$. Additionally, the fact that N is allowed to grow with n allows us to apply the Tauberian theorems for a much wider range of values of λ than originally stated in any of the known Tauberian theorems.

Throughout this section, $f(n) = O(g(n))$ implies that $|f(n)/g(n)| \leq C$, for some constant C and sufficiently large values of n . The notation $u(n) \sim \omega(n)$ is used to describe the property that $\lim_{n \rightarrow \infty} u(n)/\omega(n) = 1$. Notice that if the last condition is satisfied, then for sufficiently large n , there exist constants C_1 and C_2 such that

$$C_1 \omega(n) \leq u(n) \leq C_2 \omega(n). \quad (21)$$

When applied to the problem under consideration, \sim describes an asymptotic relation satisfied by two sequences of statistics defined over different probability spaces, with different induced probability measures. The elements of a particular sequence are also defined over different probability spaces with different probability measures.

Theorem 4. *Let $f(\lambda, N)$ and $\gamma(\lambda, N)$ be two functions that have Taylor series with respect to λ , and for all N , given by*

$$\begin{aligned} f(\lambda, N) &= \sum_{k=0}^{\infty} F_k(N) \lambda^k, \\ \gamma(\lambda, N) &= \sum_{k=0}^{\infty} \Gamma_k(N) \lambda^k, \end{aligned} \quad (22)$$

where it is assumed that $F_k(N)$ is strictly positive for sufficiently large N . Let

$$g(n, N) = \frac{1}{F_n(N)} [\lambda^n] \{f(\lambda, N) \gamma(\lambda, N)\}, \quad (23)$$

and let N be a function of n , $\alpha(n)$, such that $\lim_{n \rightarrow \infty} F_{n-1}(\alpha(n))/F_n(\alpha(n))$ exists, and is equal to r . Furthermore, let $A_n \leq r$ be a sequence such that

$$A_n \sim \frac{F_{n-1}(\alpha(n))}{F_n(\alpha(n))} \sim r. \quad (24)$$

Then

$$g(n, \alpha(n)) \sim \gamma(A_n, \alpha(n)), \quad (25)$$

provided the following conditions are satisfied:

(A1) For all sufficiently large n and for all m it holds that

$$\left| \frac{\Gamma_m(\alpha(n))}{\gamma(A_n, \alpha(n))} \right| \leq p_m, \quad (26)$$

where $\sum_{m=0}^{\infty} p_m r^m < \infty$.

(A2) For some constant C , and for all sufficiently large $m \leq n$ and large n , it holds that

$$\left| \frac{F_{n-m}(\alpha(n))}{F_n(\alpha(n))} \right| \leq C A_n^m. \quad (27)$$

Remark: The result of Theorem 4 has a very simple interpretation in terms of Equation (9). It basically establishes the conditions under which one is allowed to replace $F_{n-m}(N)/F_n(N)$ by its asymptotic formula inside the series in (9). The only condition for the validity of the theorem that depends on the transform itself is condition (A2). The other restrictions in the statement of the theorem apply only to $\gamma(\lambda, N)$. Finally, Theorem 4 provides a choice for the parameter λ (namely A_n) for which the function $\gamma(\lambda, N)$ is asymptotically equal to its inverse $g(n, N)$. Therefore, in the case that we are only interested in the asymptotic behavior of $g(n, N)$ there is no need for computing the inverse transform.

Proof: We will prove the following result equivalent to (25):

$$\lim_{n \rightarrow \infty} \frac{\gamma(A_n, N) - g(n, N)}{\gamma(A_n, N)} = 0. \quad (28)$$

Based on (23) one has

$$g(n, \alpha(n)) = \sum_{m=0}^n \Gamma_m(\alpha(n)) \frac{F_{n-m}(\alpha(n))}{F_n(\alpha(n))},$$

and from the definition of $\gamma(\lambda, N)$ it follows that

$$\gamma(A_n, \alpha(n)) = \sum_{m=0}^{\infty} \Gamma_m(\alpha(n)) A_n^m.$$

Therefore, $\gamma(A_n, \alpha(n)) - g(n, \alpha(n))$ can be written as

$$\sum_{m=0}^n \Gamma_m(\alpha(n)) A_n^m \left(1 - \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right) + \sum_{m=n+1}^{\infty} \Gamma_m(\alpha(n)) A_n^m. \quad (29)$$

Consider the first term in Equation (29). Based on (26), it is possible to show that

$$\begin{aligned} & \left| \frac{\sum_{m=0}^n \Gamma_m(\alpha(n)) A_n^m \left(1 - \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right)}{\left| \sum_{m=0}^{\infty} \Gamma_m(\alpha(n)) A_n^m \right|} \right| \\ & \leq \sum_{m=0}^n p_m A_n^m \left| 1 - \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right|. \end{aligned} \quad (30)$$

Notice that from property (A2), it follows that

$$\left| 1 - \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right| \leq \left| \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right| + 1 \leq C \frac{A_n^m}{A_n^m} + 1 = C + 1. \quad (31)$$

Therefore,

$$A_n^m \left| 1 - \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right| \leq (C+1) A_n^m. \quad (32)$$

Since $A_n \leq r$, $\sum_{m=0}^{\infty} p_m r^m$ converges absolutely, and the Lebesgue dominated convergence theorem can be applied in order to show that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \sum_{m=0}^n p_m A_n^m \left| 1 - \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right| \\ &= \sum_{m=0}^{\infty} p_m \lim_{n \rightarrow \infty} A_n^m \lim_{n \rightarrow \infty} \left| 1 - \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} \right| \\ &= \sum_{m=0}^{\infty} p_m r^m \lim_{n \rightarrow \infty} \left| \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} - 1 \right| = 0, \end{aligned} \quad (33)$$

where the first inequality follows from the fact that the limit of both the terms in the product exists. The last equality in (33) is a consequence of the facts that both

$$\lim_{n \rightarrow \infty} A_n^m = r^m, \quad (34)$$

and

$$\lim_{n \rightarrow \infty} \frac{F_{n-m}(\alpha(n))}{F_n(\alpha(n))} = r^m \quad (35)$$

hold, and therefore

$$\lim_{n \rightarrow \infty} \frac{F_{n-m}(\alpha(n))}{A_n^m F_n(\alpha(n))} = 1. \quad (36)$$

The second term in (29) can be bounded in a similar way as

$$\frac{|\sum_{m=n+1}^{\infty} \Gamma_m(\alpha(n)) A_n^m|}{|\sum_{m=0}^{\infty} \Gamma_m(\alpha(n)) A_n^m|} \leq \sum_{m=n+1}^{\infty} p_m A_n^m. \quad (37)$$

The limit with respect to n of the sum on the right-hand side equals to zero. This follows since the dominated convergence theorem applies, and $\sum_{m=n+1}^{\infty} p_m r^m$ is the tail of a convergent series and therefore converges to zero. This proves the claimed result. \blacksquare

Example: Assume that for a transform under consideration the choice for A_n is appropriate, and that (A2) holds. Let

$$\gamma(\lambda, N) = e^{-\lambda} \frac{\lambda^s}{s!},$$

where s is some constant. Then

$$\Gamma_m(N) = [\lambda^m] \gamma(\lambda, N) = \frac{(-1)^{m-s}}{s! (m-s)!},$$

and consequently, for all sufficiently large values of n ,

$$\left| \frac{\Gamma_m(N)}{\gamma(A_n, N)} \right| = \frac{\exp(A_n)}{A_n^s (m-s)!} \leq C \frac{e^r}{r^s (m-s)!} = p_m,$$

where C is a constant. The last inequality follows from the definition of A_n and (21). If for the transform under consideration condition (A2) of Theorem 4 holds, any choice of N (for which r exists) ensures that

$\sum p_m r^m$ converges, and the result of Theorem 4 holds.

The next theorem represents a specialization of Theorem 4 for the case when N is a fixed constant and the radius of convergence of $f(\lambda, N)$ is less than or equal to the radius of convergence of $\gamma(\lambda, N)$.

Theorem 5. *Let N be an arbitrary constant, and let $f(\lambda, N)$, $\gamma(\lambda, N)$, $g(n, N)$ be as defined in Theorem 4. Assume that the following conditions hold:*

(C1) *The radius of convergence of $\gamma(\lambda, N)$ and $f(\lambda, N)$ are r_1 and r_2 , respectively; r_2 can be obtained from the ratio test and $r_2 \sim A_n$.*

(C2) *$\gamma(\lambda, N)$ converges absolutely at $\lambda = r_2$ to a non-zero value;*

(C3) *For some constant C , all sufficiently large $m \leq n$, and large n assumption (A2) of Theorem 4 holds.*

Then

$$g(n, N) \sim \gamma(A_n, N). \quad (38)$$

Proof: Assume that the conditions in the theorem are satisfied. Since

$$g(n, N) = \frac{1}{F_n(N)} [\lambda^n] \{f(\lambda) \gamma(\lambda, N)\}, \quad (39)$$

it follows from Compton's Tauberian theorem [11] that

$$g(n, N) \sim \frac{1}{F_n(N)} \gamma(r_2, N) [\lambda^n] \{f(\lambda)\}. \quad (40)$$

so that since $[\lambda^n] \{f(\lambda)\} = F_n(N)$, one consequently has

$$g(n, N) \sim \gamma(r_2, N). \quad (41)$$

In order to finish the proof, we observe that the third assumption in the statement of the theorem allows us to use Lebesgue's dominated convergence theorem and to replace r_2 by its asymptotic formula. For this purpose consider $\gamma(r_2, N) = \sum_{m=0}^{\infty} \Gamma_m(N) r_2^m$. Using the fact that

$$\lim_{n \rightarrow \infty} A_n^m = r_2^m,$$

one can rewrite $\gamma(r_2, N)$ as

$$\gamma(r_2, N) = \sum_{m=0}^{\infty} \Gamma_m(N) \lim_{n \rightarrow \infty} A_n^m.$$

Then

$$\gamma(r_2, N) = \lim_{n \rightarrow \infty} \left(\sum_{m=0}^K \Gamma_m(N) A_n^m \right) + \sum_{m=K+1}^{\infty} \Gamma_m(N) \lim_{n \rightarrow \infty} A_n^m,$$

Since $\gamma(\lambda, N)$ converges absolutely at r_2 , the same arguments as in Theorem 4 can be used to show that the dominated convergence theorem applies and hence

$$\sum_{m=K+1}^{\infty} \Gamma_m(N) \lim_{n \rightarrow \infty} A_n^m = \lim_{n \rightarrow \infty} \sum_{m=K+1}^{\infty} \Gamma_m(N) A_n^m.$$

Therefore, (38) holds and this proves the theorem. \blacksquare

We conclude this section by proving a generalization of Theorem 4, that applies not only to functions analytic in some disk around the origin, but also to functions analytic in some open ring around the origin. Furthermore, the claimed results can be proved by only considering an asymptotic expression for $\gamma(\lambda, \alpha(n))$.

Theorem 6. *Let A_n , r , $F_n(N)$ and $f(\lambda, \alpha(n))$ have the same definitions and properties as given in Theorem 4. Let $\gamma(\lambda, \alpha(n))$ be a regular and single-valued function in some open ring $\mathcal{R} = \{R_1 < |\lambda| < R_2, 0 \leq R_1 < R_2 < 1\}$. Assume that $\gamma(\lambda, \alpha(n))$ has the following form:*

$$\gamma(\lambda, \alpha(n)) = \hat{\gamma}(\lambda, \alpha(n)) \left(1 + O\left(\frac{y(\lambda)}{h(n)}\right) \right), \quad (42)$$

where $\hat{\gamma}(\lambda, \alpha(n)) = c(n)b(\lambda)$, and $y(\lambda)$ and $b(\lambda)$ (which do not depend on n) are regular and single valued in the ring \mathcal{R} . Let $y(\lambda)b(\lambda)$ have a single dominant singularity. Assume furthermore that the coefficient Y_m of the Laurent expansion of $y(\lambda)b(\lambda)$ with respect to λ and the coefficient B_m of the Laurent expansion of $b(\lambda)$ have the property that $Y_m/(B_m h(n)) \rightarrow 0$, uniformly in $m \leq n$, and as $n \rightarrow \infty$. Provided that $b(\lambda)$ converges at $\lambda = A_n$ to a non-zero value, and that condition (A2) of Theorem 4 holds for both positive and negative integer values of m , one has

$$g(n, \alpha(n)) \sim \hat{\gamma}(A_n, \alpha(n)). \quad (43)$$

Proof: We will prove the theorem by showing that

$$\lim_{n \rightarrow \infty} \frac{g(n, \alpha(n)) - \hat{\gamma}(A_n, \alpha(n))}{\hat{\gamma}(A_n, \alpha(n))} = 0. \quad (44)$$

From the well known Laurent theorem [24], it follows that the functions $\gamma(\lambda, \alpha(n))$, $b(\lambda)$ and $y(\lambda)b(\lambda)$ have Laurent expansions of the form

$$\begin{aligned} \gamma(\lambda, n) &= \sum_{m=-\infty}^{\infty} \Gamma_m(\alpha(n)) \lambda^m, \\ b(\lambda) &= \sum_{m=-\infty}^{\infty} B_m \lambda^m, \\ y(\lambda)b(\lambda) &= \sum_{m=-\infty}^{\infty} Y_m \lambda^m. \end{aligned}$$

The coefficients of the first two Laurent expansions in the previous equation are asymptotically connected, based on Cauchy's theorem [31]. Cauchy's theorem states that the k -th coefficient a_k in the Laurent expansion of a function $f(z)$ regular in a ring \mathcal{R} can be obtained as

$$a_k = \frac{1}{2\pi i} \int_{|z|=\rho} \frac{f(z)}{z^{k+1}} dz, \quad (45)$$

with $\rho \in \mathcal{R}$. Based on this result, one can show that for complex valued λ , one has

$$\begin{aligned}\Gamma_m(\alpha(n)) &= \frac{1}{2\pi i} \int_{|\lambda|=\rho} \frac{\gamma(\lambda, \alpha(n))}{\lambda^{m+1}} d\lambda = \frac{c(n)}{2\pi i} \int_{|\lambda|=\rho} \frac{b(\lambda) \left(1 + O\left(\frac{y(\lambda)}{h(n)}\right)\right)}{\lambda^{m+1}} d\lambda \\ &= c(n) B_m + \frac{c(n)}{2\pi i} \int_{|\lambda|=\rho} \frac{O\left(\frac{y(\lambda) b(\lambda)}{h(n)}\right)}{\lambda^{m+1}} d\lambda = c(n) B_m \left(1 + \frac{Y_m}{h(n) B_m} O(1)\right),\end{aligned}$$

provided that $m \leq n$ and $R_1 < \rho \leq R_2$. The last equality holds based on the transfer relation of [24, p. 280], and it implies that

$$\lim_{n \rightarrow \infty} \frac{\Gamma_m(\alpha(n))}{c(n) B_m} = 1, \quad \text{uniformly in } m. \quad (46)$$

Now, using the same approach as in Theorem 4, we can write the following expression for $\hat{\gamma}(A_n, \alpha(n)) - g(n, \alpha(n))$

$$c(n) \left(\sum_{m=-\infty}^n B_m A_n^m \left(1 - \frac{\Gamma_m(\alpha(n)) F_{n-m}(\alpha(n))}{c(n) B_m A_n^m F_n(\alpha(n))}\right) + \sum_{m=n+1}^{\infty} B_m A_n^m \right) \quad (47)$$

so that

$$\begin{aligned}& \frac{\hat{\gamma}(A_n, \alpha(n)) - g(n, \alpha(n))}{\hat{\gamma}(A_n, \alpha(n))} \\ &= \frac{\sum_{m=-\infty}^n B_m A_n^m \left(1 - \frac{\Gamma_m(\alpha(n)) F_{n-m}(\alpha(n))}{c(n) B_m A_n^m F_n(\alpha(n))}\right)}{b(A_n)} + \frac{\sum_{m=n+1}^{\infty} B_m A_n^m}{b(A_n)}.\end{aligned} \quad (48)$$

Similarly as for the proof of Theorem 4, one can show that in the limit, as n goes to infinity, the absolute value of (48) tends to zero. Based on the dominated convergence theorem, Equation 2.51, and the facts that

$$\begin{aligned}\lim_{n \rightarrow \infty} A_n^m &= r^m, \\ \sum_{m=-\infty}^{\infty} B_m A_n^m &= b(A_n) \neq 0 \text{ exists,} \\ \lim_{n \rightarrow \infty} \left| 1 - \frac{\Gamma_m(\alpha(n)) F_{n-m}(\alpha(n))}{c(n) B_m A_n^m F_n(\alpha(n))} \right| &= 0,\end{aligned} \quad (49)$$

it follows that the first term in (48) tends to zero. On the other hand, the absolute value of the second term converges to zero based on the fact that $\sum_{m=n+1}^{\infty} B_m A_n^m$ is the tail of a convergent power series. This proves the claimed result. \blacksquare

Theorem 6 is especially useful for problems that give rise to complicated exact, while fairly simple asymptotic expressions in the transform domain. Several applications of Theorem 6 are presented in a companion paper [30], concerned with the average case analysis of Gosper's algorithm.

Discussion: The proofs of Theorems 4, 5, and 6 starting from (29) are similar to the proof of Compton's Tauberian theorem, with three major differences. The first one is that r is given by an expression of the form

$$r = \lim_{n \rightarrow \infty} \frac{F_{n-1}(\alpha(n))}{F_n(\alpha(n))}. \quad (50)$$

n balls	N urns	occupancy ≥ 0	occupancy ≥ 1	occupancy 0 or 1
labeled	labeled	N^n	$N! S(n, N)$	N^n
unlabeled	labeled	$\binom{n+N-1}{n}$	$\binom{n-1}{n-N}$	$\binom{N}{n}$
labeled	unlabeled	$S(1, N) + \dots + S(n, N)$ if $N < n$, B_n if $N \geq n$	$S(n, N)$	1, if $n \leq N$, 0, if $n > N$
unlabeled	unlabeled	$p(1, N) + \dots + p(n, N)$, if $N < n$, $p(n)$ if $N \geq n$	$p(n, N)$	1, if $n \leq N$, 0, if $n > N$

Table 1: “The twelve-fold way”

This allows for more flexibility in the realization of the conditions of the theorems. The other difference is that one only has to assume analyticity in a ring rather than a disk, i.e., Laurent rather than Taylor series can be taken into consideration. The third difference lies in the fact that it is not required to know the exact expression in the transform domain, but only an asymptotic form for it.

We will use the results we derived in this section to construct probabilistic transforms for some classical urn models of the “twelve-fold way” [63], summarized in Table 1. The definitions of the entries of the table are given in the next two sections, along with the transforms corresponding to them. Provided that the function $\alpha(n)$ is chosen properly, the first or second set of conditions of Theorem 4 and Theorem 5 can be satisfied for all the combinatorial numbers given in Table 1. For the entries of Table 1 we also have to fully specify in what form we want the asymptotic expression A_n to be. Of course, there are many asymptotic expressions for a function, with various error terms, and we would like to use an expression that is as simple as possible. It is also possible to show that our definition of A_n is equivalent to the optimal choice of λ in terms of the variation distance, as explained in [5]. The values of A_n are also the saddle points of the respective generating functions $f(\lambda)$, which can be easily seen from the definition in [5], or proved by using our approach. The obvious connections to the saddle point were not observed, nor explicitly pointed out in [5].

The transforms related to the first two rows of Table 1 are of special interest, because there are two degrees of freedom allowed for finding the optimal asymptotic choice of the parameter λ (since λ depends on both n and N). On the other hand, for the transforms corresponding to the first entries of the second pair or rows of Table 1 it is harder to apply the results of Theorem 4–7, because there is only one parameter, n , that influences λ .

No.	$D(n, N)$	$f(x)$
1	N^n	$\sum_{k=0}^{\infty} N^k x^k / k! = \exp(Nx)$
2	$N! S(n, N)$	$N! \sum_{k=N}^{\infty} S(k, N) x^k / k! = (\exp(x) - 1)^N$
3	$N^n = N(N-1)\dots(N-n+1)$	$\sum_{k=0}^{\infty} N^k x^k / k! = (1+x)^N$

Table 2: The generating function (Row 1)

No.	$P\{X = n\}$	$\gamma(\lambda, N)$
1	$\exp(-N\lambda) (N\lambda)^k / n!$	$\sum_{k=0}^{\infty} \exp(-N\lambda) g(k, N) (N\lambda)^k / k!$
2	$N! S(n, N) \lambda^n / ((\exp(\lambda) - 1)^N n!)$	$\sum_{k=N}^{\infty} N! S(k, N) g(k, N) \lambda^k / ((\exp(\lambda) - 1)^N k!)$
3	$\binom{N}{n} \lambda^n / (1 + \lambda)^N$	$\sum_{k=0}^{\infty} \binom{N}{k} \lambda^k g(k, N) / (1 + \lambda)^N$

Table 3: The transform (Row 1)

4 Models with labeled urns

4.1 The first row of Table 1

The defining probabilities and transforms obtained for the first row of Table 1, using the approach described in Section 2, are summarized in Tables 2,3,4. We will describe each transform arising from this type of urn model separately.

Notice that throughout this section $N^{\underline{n}}$ stands for the falling factorial, $N(N-1)\dots(N-n+1)$.

- $D(n, N) = N^n$

For this case, the distributions of balls into urns is uniform while the defining distribution is Poisson, with mean value $N\lambda$, and the transform is the well known Poisson transform [20]. The Poisson urn

No.	$g(n, N)$	$g(n, N)$
1	$(n!/N^n) [\lambda^n] \{\exp(N\lambda) \gamma(\lambda, N)\}$	$\sum_{l=0}^n \Gamma_l(N) n^l / N^l$
2	$(n!/(N! S(n, N))) [\lambda^n] \{(\exp(\lambda) - 1)^N \gamma(\lambda, N)\}$	$\sum_{l=0}^{n-N} \Gamma_l(N) S(n-l, N) n^l / S(n, N)$
3	$\binom{N}{n}^{-1} [\lambda^n] \{\gamma(\lambda, N) (1 + \lambda)^N\}$	$\sum_{l=0}^n \Gamma_l(N) n^l / (N - n + l)^l$

Table 4: The inverse transform (Row 1)

statistic is also known as the Maxwell-Boltzman statistic [16], and it is of special importance in areas such as particle physics, networking (queuing theory) and analysis of algorithms. The independent random variables Z_1, \dots, Z_N , for which $X = Z_1 + \dots + Z_N$, are Poisson distributed with mean value λ . The coefficient in the expression for the inverse transform given in the first row of Table 4 can be obtained either by using the results of Section 2 or by using the analytic saddle-point method, as demonstrated in [39].

For the Poisson transform, $F_n(N) = N^n/n!$, so that $F_{n-1}(N)/F_n(N) = n/N$. Provided that $N = \alpha(n)$ is chosen so that the limit of n/N exists, one can choose $A_n = n/N$ (in this case A_n is an exact, rather than an asymptotic expression). The choice for A_n is actually the mean value of the Poisson random variables and the saddle point of the generating function $f(\lambda)$.

Notice that for a sequence $F_n(N)$ that is log-concave in n , one has

$$\frac{F_{n-1}(N)}{F_n(N)} \leq \frac{F_n(N)}{F_{n+1}(N)},$$

so that for $m \geq 0$

$$\frac{F_{n-m}(N)}{F_n(N)} = \frac{F_{n-m}(N)}{F_{n-m+1}(N)} \cdots \frac{F_{n-1}(N)}{F_n(N)} \leq \left(\frac{F_{n-1}(N)}{F_n(N)} \right)^m,$$

and for $m < 0$

$$\frac{F_{n-m}(N)}{F_n(N)} = \frac{F_{n+|m|}(N)}{F_{n+|m|-1}(N)} \cdots \frac{F_n(N)}{F_{n-1}(N)} \leq \left(\frac{F_{n+1}(N)}{F_n(N)} \right)^{|m|}.$$

Hence, if $F_n(N)$ is log-concave and $F_{n-1}(N)/F_n(N) = A_n$ or if A_n is an asymptotic expression with an error term of the form $O(n^{-1})$, condition (A2) of Theorem 4 holds.

Since $F_n(N) = N^n/n!$ is log-concave in n ,

$$\frac{F_{n-1}(N)}{F_n(N)} = \frac{N^{n-1}/(n-1)!}{N^n/n!} = \frac{n}{N} \leq \frac{n+1}{N} = \frac{F_n(N)}{F_{n+1}(N)}, \quad (51)$$

the third condition of Theorem 4 is satisfied in this case.

- $D(n, N) = N! S(n, N)$

The defining probability in this case is constructed from the N -th Stirling number of the second kind and order n , $S(n, N)$ [42], that represent the number of partitions of a set of size n into N subsets. The exponential generating function of the Stirling numbers of the second kind is shown in Table 2 [37]. To determine the distributions of the variables Z_1, \dots, Z_N for which $X = Z_1 + \dots + Z_N$, we use the probability generating function of the defining distribution

$$G(z) = \frac{(\exp(\lambda z) - 1)^N}{(\exp(\lambda) - 1)^N}. \quad (52)$$

This function is clearly the product of N identical terms of the form

$$G_\lambda(z) = \frac{\exp(\lambda z) - 1}{\exp(\lambda) - 1}, \quad (53)$$

and $G_\lambda(z)$ represent the probability generating function of a zero truncated Poisson random variable defined by the probability mass function

$$P\{Z_i = n\} = \frac{1}{\exp(\lambda) - 1} \frac{\lambda^n}{n!}, \quad n \geq 1, \quad i = 1, \dots, N. \quad (54)$$

This result is intuitively clear from the description of the urn model. The transform based on this model is presented in the second row of Table 3, and we refer to it as the *Stirling transform of the second kind*. Since for the Stirling transform $F_n(N) = N! S(n, N)/n!$, it follows that

$$\frac{F_{n-1}(N)}{F_n(N)} = \frac{n S(n-1, N)}{S(n, N)}. \quad (55)$$

We can either take A_n to be given by (55), or by an asymptotic formula for (55). Different asymptotic formulas hold for (55), based on the relationship between n and N (notice that from the definition in Table 1 one must necessarily have $N \leq n$). An asymptotic formula for $S(n, N)$ that holds for the largest possible range of values of N was derived in [40]. One of the consequences of the results of [40] is that for $n - n^{1/3} \leq N \leq n$ it holds that

$$\frac{F_{n-1}(N)}{F_n(N)} \sim \frac{2(n-N)}{n} \left(1 - \frac{1}{n}\right)^{2(n-N)} \sim \frac{2(n-N)}{n}, \quad (56)$$

so that $A_n \rightarrow 0$ as n goes to infinity. On the other hand, for constant N , $S(n, N)/S(n-1, N) \sim N$ [1]. In this case $A_n \sim (n/N)$ does not have a finite limit. Condition (A2) of Theorem 4 holds for this case as well. This follows from the fact that the Stirling numbers $S(n, N)$ are log-concave in n . We will give a short proof of this result by using the fact that $S(n, N)$ is log-concave in N (and hence unimodal) [42, p.138]. To the best of our knowledge, this proof was not previously stated.

Lemma 7. *Let*

$$T(n, N) = \frac{S(n, N)}{S(n-1, N)}. \quad (57)$$

Then $T(n, N)$ is a strictly decreasing sequence in n . That is, the Stirling numbers of the second kind are log-concave in n .

Proof: The proof is given in Appendix 1.

Without loss of generality, assume that $n - m \geq N$, since otherwise $S(n - m, N) = 0$ and the result holds trivially. Then based on Lemma 8 one has for all integer m

$$\frac{F_{n-m}(N)}{F_n(N)} = \frac{n! S(n-m, N)}{(n-m)! S(n, N)} \leq n^m \frac{S(n-m, N)}{S(n, N)} \leq \left(\frac{n S(n-1, N)}{S(n, N)} \right)^m.$$

If we choose an asymptotic formula for $F_{n-1}(N)/F_n(N)$ that has an error term of the order of $O(n^{-1})$, then

$$(1 + O(n^{-1}))^n = O(1), \quad (58)$$

and condition (A2) holds.

No.	$D(n, N)$	$f(x)$
1	$\binom{n+N-1}{n}$	$\sum_{k=0}^{\infty} \binom{k+N-1}{k} x^k = 1/(1-x)^N$
2	$\binom{n-1}{n-N}$	$\sum_{k=N}^{\infty} \binom{k-1}{k-N} x^k = x^N/(1-x)^N$
3	$\binom{N}{n}$	$\sum_{k=0}^N \binom{N}{k} x^k = (1+x)^N$

Table 5: The generating function (Row 2)

- $D(n, N) = N\underline{n} = N(N-1)\dots(N-n+1)$

The defining probability given in the third row of Table 1 gives rise to a transform we refer to as the *binomial transform*. The probability generating function of the defining distribution is of the form

$$G(z) = \sum_{k=0}^{\infty} P\{X = k\} z^k = \left(\frac{1+z\lambda}{1+\lambda} \right)^N, \quad (59)$$

Hence, for the binomial transform, the random variables Z_1, Z_2, \dots, Z_N are identically distributed according to the Bernoulli distribution:

$$p = P\{Z_i = 0\} = \frac{1}{1+\lambda},$$

$$q = 1 - p = P\{Z_i = 1\} = \frac{\lambda}{1+\lambda}.$$

We would like to point out that there exists a different transform that is known as the “binomial transform” [34]. It is defined as

$$\gamma(n) = \sum_{k=0}^n (-1)^k \binom{n}{k} g(k), \quad (60)$$

and is of special interest in the summation theory of divergent series [23].

Since

$$\frac{F_{n-1}(N)}{F_n(N)} = \frac{n! N^{n-1}}{(n-1)! N\underline{n}} = \frac{n}{N-n+1}, \quad (61)$$

one can take $A_n = n/(N-n+1)$. Notice that for this urn model one must necessarily have $N \geq n$.

Similarly as for the previous two cases, $F_n(N)$ is log-concave in n and

$$\frac{F_{n-m}(N)}{F_n(N)} = \frac{n! N^{n-m}}{(n-m)! N\underline{n}} \leq \left(\frac{n}{N-n+1} \right)^m,$$

so that condition (A2) of Theorem 4 is satisfied for this case as well.

4.2 The second row of Table 1

The second row of Table 1 gives rise to three transforms, described in Tables 5,6,7. We will discuss each of the three transforms separately.

No.	$P\{X = n\}$	$\gamma(\lambda, N)$
1	$\binom{n+N-1}{n} \lambda^n (1-\lambda)^N$	$\sum_{k=0}^{\infty} \binom{k+N-1}{k} \lambda^k (1-\lambda)^N g(k, N)$
2	$\binom{n-1}{n-N} \lambda^{n-N} (1-\lambda)^N$	$\sum_{k=0}^{\infty} \binom{k-1}{n-N} \lambda^{k-N} (1-\lambda)^N g(k, N)$
3	$\binom{N}{k} \lambda^k / (1+\lambda)^N$	$\sum_{k=0}^N \binom{N}{k} \lambda^k g(k, N) / (1+\lambda)^N$

Table 6: The transform (Row 2)

No.	$g(n, N)$	$g(n, N)$
1	$\binom{n+N-1}{n}^{-1} [\lambda^n] \{\gamma(\lambda, N) / (1-\lambda)^N\}$	$\sum_{l=0}^{\infty} \Gamma_l(N) n^l / (N+n-1)^l$
2	$\binom{n-1}{n-N}^{-1} [\lambda^{n-N}] \{\gamma(\lambda, N) / (1-\lambda)^N\}$	$\sum_{l=0}^{\infty} \Gamma_l(N) (n-N)^l / (n-1)^l$
3	$\binom{N}{n}^{-1} [\lambda^n] \{\gamma(\lambda, N) (1+\lambda)^N\}$	$\sum_{l=0}^{\infty} \Gamma_l(N) n^l / (N-n+l)^l$

Table 7: The inverse transform (Row 2)

- $D(n, N) = \binom{n+N-1}{n}$

For this case, the defining probability distribution is negative-binomial [16], and we named the corresponding transform the negative-binomial transform. The negative-binomial urn model appears under the name Bose-Einstein statistic [16] in many areas of physics, communication theory and interpolation theory. Due to many applications of Bose-Einstein statistic, the negative-binomial transform is probably the most useful new transform presented in this paper.

From the probability generating function of the defining distribution that takes the form

$$G(z) = \left(\frac{1-\lambda}{1-\lambda z} \right)^N, \quad (62)$$

it follows that the variables Z_1, \dots, Z_N are geometrically distributed, with mean λ , i.e.

$$P\{Z_i = n\} = (1-\lambda) \lambda^n, \quad \forall n \geq 0, \quad i = 1, \dots, N. \quad (63)$$

From $F_n(N) = \binom{N+n-1}{n}$, one also has

$$\frac{F_{n-1}(N)}{F_n(N)} = \frac{\binom{N+n-2}{n-1}}{\binom{N+n-1}{n}} = \frac{n}{n+N-1} \sim \frac{n}{n+N}. \quad (64)$$

Choosing $A_n = n/(n+N)$ insures the validity of the conditions of Theorem 4. This again follows from the fact that $F_n(N)$ is log-concave in n :

$$\frac{F_{n-1}(N)}{F_n(N)} = \frac{n}{n+N-1} \leq \frac{n+1}{n+N} = \frac{F_n(N)}{F_{n+1}(N)}.$$

The negative-binomial transform can be generalized along the same lines as it was done for the Poisson transform in [24], by assuming that a ball is placed in urn number k with probability π_k , where the values of the π_k 's not necessarily identical (the possible $D(n, N)$ distributions of balls into urns are not equally likely). The transform arising from this generalization is presented in Appendix 2.

- $D(n, N) = \binom{n-1}{N-1}$

For the second entry of the second row of Table 1, the random variables Z_1, \dots, Z_N are zero-truncated geometric variables, with probability distributions

$$P\{Z_i = n\} = (1 - \lambda) \lambda^{n-1}, \quad n \geq 1, \quad i = 1, \dots, N. \quad (65)$$

For this model $n \geq N$, and A_n can be determined from

$$\frac{F_{n-1}(N)}{F_n(N)} = \frac{\binom{n-2}{N-1}}{\binom{n-1}{N-1}} = \frac{n-N}{n-1}. \quad (66)$$

Additionally, since $F_n(N)$ is log-concave in n ,

$$\frac{F_{n-k}(N)}{F_n(N)} \leq \left(\frac{n-N}{n-1} \right)^k, \quad (67)$$

and condition (A2) of Theorem 4 is satisfied as well.

- $D(n, N) = \binom{N}{n}$

The urn model corresponding to this case is also known as the Fermi-Dirac model [16]. Due to the fact that the exponential generating function of the falling factorial $N^{\underline{n}}$ and the ordinary generating function of the binomial coefficient are identical, the transforms they define are the same. The difference is only clear when applying the transform, because the distributions of balls that the transform statistics enumerate are restricted in different ways.

5 Models with unlabeled urns

As described in the previous section, the combinatorial structure of models with unlabeled urns is different from the combinatorial structure of models with labeled urns. The difference comes from the fact that the urns cannot be distinguished, and hence there is no need for knowing their exact occupancies - it is only important to know how many urns have a certain occupancy. The transform techniques developed in this work allow us to assume that the random variables V_1, V_2, V_3, \dots specifying the number of urns with occupancy 1, 2, 3, ..., respectively, are independent.

5.1 The third row of Table 1

The defining probabilities and transforms derived from the entries in the third row of Table 1 are presented in Tables 8,9,10. We will not discuss the entry corresponding to the case when the occupancies are

No.	$D(n, N)$	$f(x)$
1	B_n	$\sum_{k=0}^{\infty} B_k x^k / k! = \exp(\exp(x) - 1)$
2	$S(n, N)$	$(\exp(x) - 1)^N / N!$

Table 8: The generating function (Row 3)

No.	$P\{X = n\}$	$\gamma(\lambda, N)$
1	$B_n \exp(1 - \exp(\lambda)) \lambda^n / n!$	$\sum_{k=0}^{\infty} B_k \exp(1 - \exp(\lambda)) g(k) \lambda^k / k!$
2	$(N! S(n, N) \lambda^n) / (n! (\exp(\lambda) - 1)^N)$	$\sum_{k=N}^{\infty} (N! S(k, N) g(k, N) \lambda^k) / (k! (\exp(\lambda) - 1)^N)$

Table 9: The transform (Row 3)

restricted to be zero or one, since for this case the results are of no practical interest.

- $D(n, N) = B_n$

For the first entry of the third row of Table 1, we will only analyze the case $N \geq n$. B_n is the n -th Bell number [37, pp. 33-34], the total number of partitions of a set of size n . We will refer to the defining distribution and transform constructed from the Bell numbers as the Bell distribution and transform, respectively. The probability generating function of the defining Bell distribution is

$$G(z) = \frac{\exp(\exp(\lambda z) - 1)}{\exp(\exp(\lambda) - 1)}, \quad (68)$$

so that using the formula

$$\exp(\lambda z) - 1 = \sum_{k=1}^{\infty} \frac{(\lambda z)^k}{k!},$$

it follows that Equation (68) can be rewritten as

$$G(z) = \prod_{k=1}^{\infty} \frac{\exp\left(\frac{(\lambda z)^k}{k!}\right)}{\exp\left(\frac{\lambda^k}{k!}\right)} = \prod_{k=1}^{\infty} G_{\lambda^k}(z^k),$$

No.	$g(n, N)$	$g(n, N)$
1	$(n! / B_n) [\lambda^n] \{\exp(\exp(\lambda) - 1) \gamma(\lambda, N)\}$	$\sum_{l=0}^{\infty} \Gamma_l(N) B_{n-l} n^l / B_n$
2	$(n! / (N! S(n, N))) [\lambda^n] \{(\exp(\lambda) - 1)^N \gamma(\lambda, N)\}$	$\sum_{l=0}^{\infty} \Gamma_l(N) S(n-l, N) n^l / S(n, N)$

Table 10: The inverse transform (Row 3)

where

$$G_{\lambda^k}(z^k) = \frac{\exp\left(\frac{z^k}{k!}\right)}{\exp\left(\frac{\lambda^k}{k!}\right)}.$$

This implies that the sequence of independent random variables in the transform domain V_i , $i = 1, 2, \dots$, is Poisson distributed with parameter $\lambda^i/i!$:

$$P\{V_i = n\} = \exp\left(-\frac{\lambda^i}{i!}\right) \frac{(\lambda^i/i!)^n}{n!}.$$

In order to ensure that the joint distribution of V_1, V_2, \dots defined in Proposition 2, as well as $X = V_1 + 2V_2 + 3V_3 + \dots$ are meaningful with respect to the above setting, we observe that the following inequality holds

$$P\{V_i \neq 0\} = 1 - \exp\left(-\frac{\lambda^i}{i!}\right) \leq \frac{\lambda^i}{i!}, \quad (69)$$

based on the fact that $\exp(-x) \geq 1 - x$. Hence,

$$\sum_{i=1}^{\infty} P\{V_i \neq 0\} \leq \sum_{i=1}^{\infty} \frac{\lambda^i}{i!} = \exp(\lambda) - 1.$$

This shows that $\sum_{i=1}^{\infty} P\{V_i \neq 0\}$ is finite, so that from the first Borel-Cantelli lemma [8] it follows that the probability of infinitely many V_i being non-zero is zero. Therefore, the joint distribution of V_1, V_2, \dots is well defined and X is finite with probability one.

For the Bell transform, there is no flexibility in choosing N so that to satisfy condition (A1) of Theorem 4. For the Bell transform $r = 1$, and if we want to apply the asymptotic results of Theorem 4, we need to make sure that the function $\sum p_m \lambda^m$ in the transform domain converges at $\lambda = r = 1$. An appropriate choice for A_n , and hence for the optimal value of λ , can be found by using asymptotic expressions for the Bell numbers. It is interesting to note that the same optimal value of λ was derived in [5] by minimizing the variation distance between the approximation and the exact result. More specifically, in [5] it was shown that a good choice for λ is one that satisfies

$$E[X^{(n)}] = E[V_1] + 2E[V_2] + \dots + nE[V_n] \sim n, \quad (70)$$

where $X^{(n)}$ is the value of the variable X after truncating its series at the n -th term, and V_1, V_2, \dots are parameterized by λ as described in the previous sections. Since λ has to be chosen so that $f(\lambda)$ converges, one has

$$E[X^{(n)}] \sim E[X] = \sum_{k=0}^{\infty} k \frac{F(k, N)}{f(\lambda)} \lambda^k = \lambda \frac{f'(\lambda)}{f(\lambda)},$$

so that (70) reads as

$$\lambda \frac{f'(\lambda)}{f(\lambda)} \sim n.$$

A variable λ satisfying the above equation is by definition the saddle point of $f(\lambda)$ [33]. This fact was not pointed out by the authors of [5].

The ratio of B_{n-m} and B_n can be expressed as

$$\frac{B_{n-m}}{B_n} = (1 + o(1)) \exp\left(-m\rho + O\left(\frac{m^2}{n}\right)\right) \sim \exp(-m\rho), \quad (71)$$

where ρ is the solution of the equation $\rho(n) e^{\rho(n)} = n$ [12]. The previous equation implies that A_n can be chosen to be asymptotically of the form

$$n \frac{B_{n-1}}{B_n} \sim n \exp(-\rho) \sim \rho. \quad (72)$$

where the last expression follows from the definition of ρ . We can check that this expression is the saddle point in the following way. From the expression for the generating function of the Bell numbers and the definition of the saddle point, it follows that the saddle point ρ satisfies the equation

$$\frac{f'(\rho)}{f(\rho)} = \frac{n}{\rho},$$

which reduces to

$$\frac{\exp(\exp(\rho) - 1)'}{\exp(\exp(\rho) - 1)} = \frac{n}{\rho}, \quad (73)$$

i.e. $e^\rho = n/\rho$.

The third condition of Theorem 4 is satisfied due to the fact that the generating function of the Bell numbers is of the form given in (15) and hence $B_n/n!$ is a log-concave sequence in n , i.e

$$\frac{n! B_{n-m}}{(n-m)! B_n} \leq \left(\frac{n B_{n-1}}{B_n}\right)^m \leq C \rho^m. \quad (74)$$

Due to the special form of the generating function of the Bell numbers, which is a HS-admissible [33] function, one can also compute the inverse transform analytically, using the saddle point method.

- $D(n, N) = S(n, N)$

The second rows of Tables 8, 9, 10, are essentially the same as the second rows of Tables 2, 3, 4, respectively.

5.2 The fourth row of Table 1

As for the previous case, we will restrict our attention only to the first two entries of the fourth row of Table 1. The defining probabilities and resulting transforms are given in Tables 11,12,13.

- $D(n, N) = p(n)$

For the first entry in the fourth row of Table 1, we analyze only the case $N \geq n$, for which $p(n)$ denotes the number of partitions of the integer n [3]. For this urn model, we assume that the random variable X satisfies $X = V_1 + 2V_2 + 3V_3 + \dots$, according to the results in Section 1, so that V_i corresponds to the number of parts of size i in a random partition. The generating function of

No.	$D(n, N)$	$f(x)$
1	$p(n)$	$\sum_{k=0}^{\infty} p(k) x^k = 1/((1-x)(1-x^2)(1-x^3)\dots)$
2	$p(n, N)$	$\sum_{k=0}^{\infty} p(k, N) x^k = x^N/((1-x)(1-x^2)\dots(1-x^N))$

Table 11: The generating function (Row 4)

No.	$P\{X = n\}$	$\gamma(\lambda, N)$
1	$p(k) \lambda^k \prod_{m=1}^{\infty} (1 - \lambda^m)$	$\sum_{k=0}^{\infty} p(k) \lambda^k (\prod_{m=1}^{\infty} (1 - \lambda^m)) g(k)$
2	$p(n, N) \lambda^{n-N} \prod_{k=1}^N (1 - \lambda^k)$	$\sum_{k=0}^{\infty} p(n, N) \lambda^{n-N} \left(\prod_{k=1}^N (1 - \lambda^k) \right) g(k, N)$

Table 12: The transform (Row 4)

No.	$g(n, N)$	$g(n, N)$
1	$(1/p(n)) [\lambda^n \{\gamma(\lambda)/((1-\lambda)(1-\lambda^2)\dots)\}]$	$\sum_{l=0}^{\infty} \Gamma_l p(n-l)/p(n)$
2	$(1/p(n, N)) [\lambda^{n-N} \{\gamma(\lambda, N) / ((1-\lambda)\dots(1-\lambda^N))\}]$	$\sum_{l=0}^{\infty} \Gamma_l(N) p(n-N-l, N)/p(n, N)$

Table 13: The inverse transform (Row 4)

$p(n)$ is given by the well known formula of Euler, and is shown in Table 11. Since the probability generating function of the random variable X is of the form

$$G(z) = \frac{(1-\lambda)(1-\lambda^2)(1-\lambda^3)\dots}{(1-\lambda z)(1-(\lambda z)^2)(1-(\lambda z)^3)\dots}, \quad (75)$$

it factors as

$$G_\lambda(z) G_{\lambda^2}(z^2) \dots G_{\lambda^k}(z^k) \dots = \prod_{k=1}^{\infty} \frac{1-\lambda^k}{1-(\lambda z)^k}. \quad (76)$$

This implies that V_1, V_2, \dots are independent, geometrically distributed random variables, with parameters $\lambda, \lambda^2, \lambda^3, \dots$. Similarly as for the Bell transform, one can show that for this case

$$P\{V_i \neq 0\} = 1 - (1 - \lambda^i) = \lambda^i \quad (77)$$

holds and therefore

$$\sum_{i=1}^{\infty} P\{V_i \neq 0\} = \sum_{i=1}^{\infty} \lambda^i = \frac{\lambda}{1-\lambda}. \quad (78)$$

The last expression is finite whenever $\lambda \neq 1$, so that the first Borel-Cantelli lemma implies that the joint distribution of V_1, V_2, \dots is well defined and that X is finite with probability one. The algebraic inverse of the partition transform is shown in Table 13. One may also try to use analytic methods, such as Hardy-Ramanujan's circle method [4], in order to find the asymptotic behavior of the inverse transform.

From the famous result by Hardy and Ramanujan [3] it follows that

$$\frac{p(n-1)}{p(n)} \sim \frac{\frac{1}{4\sqrt{3}(n-1)} \exp\left(\pi \sqrt{\frac{2(n-1)}{3}}\right)}{\frac{1}{4\sqrt{3}n} \exp\left(\pi \sqrt{\frac{2n}{3}}\right)}. \quad (79)$$

The last equation can be simplified as

$$\begin{aligned} \frac{p(n-1)}{p(n)} &\sim \exp\left(\pi \sqrt{\frac{2}{3}} (\sqrt{n-1} - \sqrt{n})\right) = \exp\left(\pi \sqrt{\frac{2}{3}} \sqrt{n} \left(\sqrt{1 - \frac{1}{n}} - 1\right)\right) \\ &\sim \exp\left(-\pi \sqrt{\frac{2}{3}} \sqrt{n} \frac{1}{2n}\right) = \exp\left(-\pi \sqrt{\frac{1}{6n}}\right). \end{aligned} \quad (80)$$

We would like to remark that the optimal choice for λ in Equation (80) was easy to determine based on the Tauberian theorems presented in this work. On the other hand, the authors of [19] and [5] state that they had to resort to guessing a solution in the form of an exponential and then verify that it would minimize the variation distance.

The expression in (80) is the value of λ that minimizes the asymptotic variation distance between the target statistic and its inverse, as shown in [5], but also the saddle point used in the derivations by Meinardus [3]. Meinardus found the following expression for the saddle point of $f(\lambda)$

$$\exp\left(-n^{-1/(1+\alpha)} (A\Gamma(\alpha+1) \zeta(\alpha+1))^{1/(1+\alpha)}\right), \quad (81)$$

where $\alpha = 1$ is the radius of convergence of the Riemann Zeta function,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \quad (82)$$

$\Gamma(z)$ is the well-known Gamma function, and A is the residue of the simple pole of $\zeta(s)$ at $s = \alpha = 1$ (for more details, see [3]). For the partition number $p(n)$, the residue at the pole $s = 1$ is $A = 1$. Using additionally the well known fact that $\zeta(2) = \pi^2/6$, one can see that the saddle point is exactly of the form (80). If A_n is taken to be of the form given by (80), then from the formula [33, p.18],

$$\frac{p(n-m)}{p(n)} = \left(1 + O\left(n^{-1/6}\right)\right) \exp\left(-\frac{m\pi}{\sqrt{6n}}\right),$$

it follows that

$$\frac{p(n-m)}{p(n)} \leq C \exp\left(-\frac{m\pi}{\sqrt{6n}}\right) = C A_n^m.$$

Hence, condition (A2) of Theorem 4 is satisfied for A_n as chosen above. Notice that the sequence $p(n)$ is neither log-convex nor log-concave.

- $D(n, N) = p(n, N)$

The second row of Table 11 presents the defining distribution generated from the function $p(n, N)$, enumerating the number of partitions of an integer n into exactly N parts (which is the same as the number of partitions of an integer into parts of size $\leq N$). The ordinary generating function of $p(n, N)$ is shown in Table 11 [3]. V_1, \dots, V_N are zero-truncated geometric variables, with probability distributions that have parameters $\lambda, \dots, \lambda^N$, respectively. As for the case of Stirling numbers of the second kind, there exist different asymptotic expressions for $p(n, N)$, depending on the relationship between n and N . For example, from a result by Erdős [3] which establishes that

$$p(n, N) \sim \frac{n^{N-1}}{N!(N-1)!} \quad (83)$$

holds for $N = o(n^{1/3})$, one can get

$$\frac{p(n-1, N)}{p(n, N)} \sim \frac{(n-1)^{N-1}}{n^{N-1}} \sim \left(1 - \frac{1}{n}\right)^{n \frac{N-1}{n}} \sim \exp\left(-\frac{N-1}{n}\right).$$

The same type of argument as presented for the case of the partition function can be used in this case to show that for A_n chosen of the form (84) condition (A2) of Theorem 4 holds.

6 Some other transforms related to combinatorial numbers

In this section we will derive two transforms corresponding to the distribution of the number of cycles in a permutation. For this case, $D(n, N) = C(n, N)$, where $C(n, N)$ is the absolute value of the N -th Stirling number of the first kind and order n (the absolute value of the Stirling number of the first kind with parameters n and N counts the number of permutations in the symmetric group S_n that have exactly

N cycles). The exponential generating function of the absolute value of the Stirling numbers of the first kind is

$$\sum_{k \geq N} C(k, N) \frac{x^k}{k!} = \frac{\log^N \left(\frac{1}{1-\lambda} \right)}{N!}.$$

The defining probability can be found as

$$P\{X = n\} = C(n, N) \frac{N!}{n!} \frac{\lambda^n}{\log^N \left(\frac{1}{1-\lambda} \right)},$$

while the transform, which we will refer to as the Stirling transform of the first kind, is of the form

$$\gamma(\lambda, N) = \sum_{k \geq N} C(k, N) \frac{N!}{k!} \frac{\lambda^k}{\log^N \left(\frac{1}{1-\lambda} \right)} g(k, N).$$

The inverse transform can be obtained in the following form

$$g(n, N) = \frac{n!}{N! C(n, N)} [\lambda^n] \left\{ \log^N \left(\frac{1}{1-\lambda} \right) \gamma(n, N) \right\},$$

and for $\gamma(\lambda, N)$ analytic in a disk around the origin

$$\gamma(\lambda, N) = \sum_{l=0}^{\infty} \Gamma_l(N) \frac{C(n-l, N)}{C(n, N)} \frac{n!}{N!},$$

where $\Gamma_l(N)$ is as defined as in Equation (8). The probability generating function of the defining distribution is

$$G(z) = \frac{\log^N \left(\frac{1}{1-z\lambda} \right)}{\log^N \left(\frac{1}{1-\lambda} \right)},$$

and this function is clearly the product of N identical functions of the form

$$G_\lambda(z) = \frac{\log \left(\frac{1}{1-z\lambda} \right)}{\log \left(\frac{1}{1-\lambda} \right)}. \quad (84)$$

$G_\lambda(z)$ in (84) is the probability generating function of a logarithmic random variable with probability mass function

$$P\{Z_i = n\} = \frac{1}{\log \left(\frac{1}{1-\lambda} \right)} \frac{\lambda^n}{n}, \quad \forall n \geq 0, i = 1, \dots, N.$$

Based on the discussion in Section 1, Z_i represents the length of the i -th cycle in a random permutation in S_n that has exactly N cycles. As for the case of the Stirling numbers of the second kind, the optimal choice for λ depends on the functional relationship between n and N .

For example, if $n - n^{1/3} \leq N \leq n$, then based on [40]

$$A_n \sim \frac{n C(n-1, N)}{C(n, N)} \sim \frac{2(n-N)}{N}.$$

Similarly, for $N \leq \sqrt{\log n}$,

$$A_n \sim \frac{n C(n-1, N)}{C(n, N)} \sim \exp\left(-\frac{N-1}{n(\log n + \gamma)}\right).$$

The absolute values of the Stirling numbers of the second kind $C(n, N)$ form a log-concave sequence in n , which can be proved along the same lines as Lemma 8. This fact allows one to conclude that condition (A2) of Theorem 4 is also satisfied for this transform.

Closely connected to the above transform is the geometric transform, which we will describe next. Let $D(n, N) = n!$. The generating function and defining probability for this case take the form

$$f(x) = \sum_{n=0}^{\infty} n! \frac{x^n}{n!} = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x},$$

and

$$P\{X = n\} = (1-\lambda) \lambda^n,$$

respectively. This gives rise to the geometric transform,

$$\gamma(\lambda) = \sum_{k=0}^{\infty} (1-\lambda) \lambda^k g(k),$$

with an inverse

$$g(n) = [\lambda^n] \left\{ \frac{\gamma(\lambda)}{1-\lambda} \right\}.$$

For $\gamma(\lambda)$ analytic in a disk around the origin the inverse transform is of an especially simple form

$$g(n) = \sum_{k=0}^{\infty} \Gamma_k.$$

The probability generating function of a geometric random variable is of the form

$$G(z) = \sum_{k=0}^{\infty} (1-\lambda) \lambda^k z^k = \frac{1-\lambda}{1-z\lambda}.$$

Writing

$$\begin{aligned} 1-\lambda &= \exp\left(-\log\left(\frac{1}{1-\lambda}\right)\right), \\ 1-z\lambda &= \exp\left(-\log\left(\frac{1}{1-z\lambda}\right)\right), \end{aligned}$$

shows that

$$G(z) = \frac{\exp\left(\log\left(\frac{1}{1-z\lambda}\right)\right)}{\exp\left(\log\left(\frac{1}{1-\lambda}\right)\right)} = \frac{\exp\left(\frac{\lambda z}{1} + \frac{(\lambda z)^2}{2} + \frac{(\lambda z)^3}{3} + \dots\right)}{\exp\left(\frac{\lambda}{1} + \frac{\lambda^2}{2} + \frac{\lambda^3}{3} + \dots\right)} = \prod_{k=1}^{\infty} \exp\left(\frac{\lambda^k}{k} (z^k - 1)\right).$$

Therefore, the variables V_k counting the number of cycles of length k are Poisson distributed, with mean value λ^k/k (compare this to the case of the Bell transform where the mean is given by $\lambda^k/k!$). For the

geometric transform, the ratio $F_{n-1}/F_n = 1$ for all n . Hence, under the set of conditions of Theorem 4, the optimal choice for λ would be equal to one.

Although the geometric transform is usually attributed to Grabner [21], it was first explicitly stated in a paper by Shepp and Lloyd [35].

We conclude this section by pointing out that some of the transforms presented in this paper are closely connected to various summability methods, such as the Borel and Hausdorff method [23].

7 Appendix 1

Proof of Lemma 7:

We start from the recurrence relation satisfied by the Stirling numbers of the second kind [1]

$$S(n, N) = N S(n-1, N) + S(n-1, N-1). \quad (85)$$

Dividing both sides of (85) by $S(n-1, N)$ gives

$$\frac{S(n, N)}{S(n-1, N)} = N + \frac{S(n-1, N-1)}{S(n-1, N)}.$$

Therefore,

$$T(n, N) - N = \frac{S(n-1, N-1)}{S(n-1, N)}. \quad (86)$$

Equation (86) implies

$$\frac{T(n, N) - N}{T(n-1, N) - N} = \frac{S(n-1, N-1) S(n-2, N)}{S(n-1, N) S(n-2, N-1)} = \frac{T(n-1, N-1)}{T(n-1, N)}, \quad (87)$$

as well as

$$\frac{T(n, N) - N}{T(n, N-1) - (N-1)} = \frac{S^2(n-1, N-1)}{S(n-1, N) S(n-1, N-2)} > 1, \quad (88)$$

where the last inequality holds since the Stirling numbers are log-concave in N . From Equation (88) one can see that

$$T(n, N) - N > T(n, N-1) - (N-1),$$

which is equivalent to

$$\frac{T(n, N-1)}{T(n, N)} < \frac{1}{1 + \frac{1}{T(n, N-1)}} < 1,$$

since $T(n, N-1)$ is positive. The last equation shows that

$$\frac{T(n-1, N-1)}{T(n-1, N)} < 1$$

holds, and from (87) it follows that

$$\frac{T(n, N) - N}{T(n-1, N) - N} < 1, \quad (89)$$

which proves the claimed result. ■

8 Appendix 2

Let Z_i be a geometric random variable with parameter λ_i . The probability generating function $G(z)$ of the variable $X = Z_1 + \dots + Z_N$ is in this case of the form

$$\prod_{i=1}^N \frac{1 - \lambda_i}{1 - \lambda_i z},$$

so that

$$\begin{aligned} P\{X = k\} &= [z^k] \prod_{i=1}^N \frac{1 - \lambda_i}{1 - \lambda_i z} \\ &= \prod_{i=1}^N (1 - \lambda_i) [z^k] \prod_{i=1}^N \frac{1}{1 - \lambda_i z} \\ &= h_k(\lambda_1, \dots, \lambda_N) \prod_{i=1}^N (1 - \lambda_i), \end{aligned} \quad (90)$$

where $h_k(\lambda_1, \dots, \lambda_N)$ is the complete homogeneous function of order k over the set of N variables $\lambda_1, \dots, \lambda_N$ defined as [38, p.294]

$$h_k(\lambda_1, \dots, \lambda_N) = \sum_{\substack{a_1 + \dots + a_N = k, \\ a_i \geq 0}} \lambda_1^{a_1} \dots \lambda_N^{a_N}. \quad (91)$$

The specialization of h_k for $\lambda_i = \lambda$, $i = 1, \dots, N$ is $\lambda^n \binom{n+N-1}{n}$ [38], in agreement with the results derived in the previous sections. The transform for the above case can be defined as

$$\gamma(\lambda_1, \dots, \lambda_N, N) = \sum_{k=0}^{\infty} \left(\prod_{i=1}^N (1 - \lambda_i) \right) h_k(\lambda_1, \dots, \lambda_N) g(k, N),$$

so that

$$\frac{\gamma(\lambda_1, \dots, \lambda_N, N)}{\prod_{i=1}^N (1 - \lambda_i)} = \sum_{k=0}^{\infty} h_k(\lambda_1, \dots, \lambda_N) g(k, N).$$

Let $\gamma(\lambda_1, \dots, \lambda_N, N) = \sum_{\mu} q_{\mu} \lambda^{\mu}$, where μ is an ordered N -tuple of non-negative integers (μ_1, \dots, μ_N) , $|\mu| = \sum_{i=1}^N \mu_i$, and $\lambda^{\mu} = \lambda_1^{\mu_1} \dots \lambda_N^{\mu_N}$. Furthermore, let q_{μ} be real numbers. Then

$$[\lambda^{\mu}] \left\{ \frac{\gamma(\lambda_1, \dots, \lambda_N, N)}{\prod_{i=1}^N (1 - \lambda_i)} \right\} = g(|\mu|, N). \quad (92)$$

The last result implies that

$$g(|\mu|, N) = \sum_{\eta \leq \mu} q_{\eta}, \quad \text{where } \eta = (\eta_1, \dots, \eta_N) \text{ and } \eta_i \leq \mu_i \text{ } i = 1, \dots, N. \quad (93)$$

9 Appendix 3

In this appendix, we will present several simple examples for the application of the new transforms in combinatorial theory. Since the Poisson transform is by now a standard tool in probabilistic analysis, no

examples for its application are given; for an extensive reference of applications see [24].

Example 1: We will show next how to find the number of ordered, non-negative N -tuples (Z_1, \dots, Z_N) , such that $Z_i = 0$ or $1 \pmod{3}$, satisfying the equation $Z_1 + \dots + Z_N = n$. The number of solutions of this equation without the constraint is $\binom{n+N-1}{n}$, so that we may assume that Z_1, \dots, Z_N are independent, identically distributed geometric random variables with parameter λ . The probability of having a solution that satisfies the above constraint can be found in the transform domain as

$$\begin{aligned} & P\{Z_k = 3m \text{ or } 3m + 1, m = 1, 2, \dots, k = 1, \dots, N\} \\ &= \left((1 - \lambda) \sum_{m=0}^{\infty} \lambda^{3m} + (1 - \lambda) \sum_{m=0}^{\infty} \lambda^{3m+1} \right)^N \\ &= \left(\frac{1 - \lambda}{1 - \lambda^3} + \frac{\lambda(1 - \lambda)}{1 - \lambda^3} \right)^N = \left(\frac{1 - \lambda^2}{1 - \lambda^3} \right)^N. \end{aligned} \tag{94}$$

If the number of solutions of the problem is $E_{n,N}$, then from the definition of the inverse transform one has

$$\begin{aligned} E_{n,N} &= [\lambda^n] \left(\left(\frac{1 - \lambda^2}{1 - \lambda^3} \right)^N \frac{1}{(1 - \lambda)^N} \right) = [\lambda^n] \left(\frac{1 + \lambda}{1 - \lambda^3} \right)^N \\ &= [\lambda^n] \sum_{l=0}^{\infty} \binom{N+l-1}{l} \lambda^{3l} (1 + \lambda)^N \\ &= \sum_{l=0}^{\lfloor \frac{n}{3} \rfloor} \binom{N+l-1}{l} \binom{N}{n-3l}. \end{aligned}$$

We observe that the asymptotic formula of Theorem 4 cannot be applied to this case, because the function (94) is exponential in N .

Example 2: In this example we will show how to compute the exponential generating function of the number of permutations with cycles of even length only. For permutations with a fixed number of cycles N , the probability of each cycle being even equals

$$\left(\sum_{m=1}^{\infty} P\{V_k = 2m\} \right)^N = \frac{1}{\log^N \left(\frac{1}{1-\lambda} \right)} \left(\sum_{m=1}^{\infty} \frac{\lambda^{2m}}{2m} \right)^N = \left(\frac{1}{2} \frac{\log \left(\frac{1}{1-\lambda^2} \right)}{\log \left(\frac{1}{1-\lambda} \right)} \right)^N.$$

Therefore, the exponential generating function of the number of permutations with N cycles of even length is

$$\frac{1}{N!} \left(\frac{\log \left(\frac{1}{1-\lambda^2} \right)}{2} \right)^N.$$

Summing over all values of $N \geq 0$ gives the desired result

$$\exp \left(\frac{\log \left(\frac{1}{1-\lambda^2} \right)}{2} \right) = \frac{1}{\sqrt{1-\lambda^2}}.$$

The last example is concerned with the partition transform. Before presenting the specific examples, we would like to point out several things. For the partition transform the urn occupancies are non-identically distributed random variables. Therefore, the probability of having not more than k parts in an integer partition with parts not greater than N is given by $h_k(\lambda, \lambda^2, \dots, \lambda^N) \prod_{i=1}^{\infty} (1 - \lambda^i)$, where $h_k(\lambda_1, \lambda_2, \dots, \lambda_N)$ is the complete homogeneous symmetric function of order k in N variables defined in the previous section. A property of the partition transform is that the parameters $\lambda_1, \lambda_2, \dots$ of the complete symmetric functions are consecutive powers of the same parameter λ , and there are well known formulas for the specialization of $h_k(\lambda_1, \lambda_2, \dots, \lambda_N)$ for this case [38, p.303].

If one is interested in evaluating distribution statistics for the case when no repeated parts are allowed in a partition, the proper choice for a symmetric function is the elementary symmetric function [38, p.290], given by

$$e_k(\lambda_1, \lambda_2, \dots, \lambda_N) = \sum_{\substack{a_1 + \dots + a_N = k, \\ 0 \leq a_i \leq 1}} \lambda_1^{a_1} \dots \lambda_N^{a_N}. \quad (95)$$

Example 3: This example will demonstrate how to find the generating function of the number of partitions that fit in a $k \times m$ rectangular Ferrer's diagram. For this case, one has to consider partitions with at most k parts, with each part less than or equal to m . This translates to saying that

$$\begin{aligned} V_{m+1} &= V_{m+2} = \dots = 0, \\ V_1 + V_2 + \dots + V_m &\leq k. \end{aligned} \quad (96)$$

The joint probability of these two events can be written as

$$\begin{aligned} &P\{V_{m+1} = V_{m+2} = \dots = 0, V_1 + V_2 + \dots + V_m \leq k\} \\ &= P\{V_{m+1} = V_{m+2} = \dots = 0\} \sum_{s=0}^k P\{V_1 + V_2 + \dots + V_m \leq k\}, \end{aligned} \quad (97)$$

since the set $\{V_{m+1}, V_{m+2}, \dots\}$ is independent from the set $\{V_1, \dots, V_m\}$. Now,

$$P\{V_{m+1} = V_{m+2} = \dots = 0\} = \prod_{l=m+1}^{\infty} (1 - \lambda^l). \quad (98)$$

The probability $P\{V_1 + V_2 + \dots + V_m = s\}$ can be easily expressed in terms of the complete homogeneous symmetric function of degree s and over m variables as

$$P\{V_1 + V_2 + \dots + V_m = s\} = h_s(\lambda, \dots, \lambda^m) \prod_{i=1}^m (1 - \lambda^i). \quad (99)$$

It is well known that the specialization of $h_s(\lambda_1, \dots, \lambda_m)$ for $(\lambda_1, \dots, \lambda_m) = (\lambda, \dots, \lambda^m)$ [38, p.303] equals

$$\lambda^s \left[\begin{matrix} s + m - 1 \\ s \end{matrix} \right]_{\lambda},$$

where

$$\left[\begin{matrix} b \\ a \end{matrix} \right]_{\lambda} = \frac{(1 - \lambda)(1 - \lambda^2) \dots (1 - \lambda^b)}{(1 - \lambda) \dots (1 - \lambda^a)(1 - \lambda) \dots (1 - \lambda^{b-a})}$$

is the Gaussian binomial coefficient (Gaussian polynomial). This implies that the generating function for the partitions that fit into a $k \times m$ array is

$$\sum_{s=0}^k \begin{bmatrix} s+m-1 \\ s \end{bmatrix}_\lambda \lambda^s. \quad (100)$$

The equation can be simplified by using the identity [37]

$$\sum_{j=0}^n \lambda^j \begin{bmatrix} m+j \\ m \end{bmatrix}_\lambda = \begin{bmatrix} n+m+1 \\ m+1 \end{bmatrix}_\lambda. \quad (101)$$

Hence, the generating function we considered is equal to the right-hand side of (101).

10 Conclusion

In this paper we addressed the problem of constructing probabilistic transforms that simplify the analysis of random urn occupancy statistics. The transforms and their inverses are defined in terms of a probabilistic average, with respect to a probability distribution constructed from the generating function of the number of distributions of balls into urns. The transforms map dependent random occupancies into independent random occupancies, allowing for many simplifications in analyzing statistics that depend on the distributions of these variables. Several elementary examples for the use of these transforms in combinatorial theory are presented as well.

References

- [1] M. Abramowitz, I. Stegun, *Handbook of Mathematical Functions*, Dover Publications, INC., New York, 1964.
- [2] F. Alajaji, T. Fuja, "A Communication Channel Modeled on Contagion," *IEEE Trans. on Info. Theory*, vol. 40, pp. 2035-2041, 1994.
- [3] G. Andrews, *The Theory of Partitions*, Cambridge University Press, 1984.
- [4] T. Apostol, *Modular Functions and Dirichlet Series in Number Theory*, Springer Verlag, NY, 1990.
- [5] R. Arratia, S. Tavaré, "Independent Process Approximations for Random Combinatorial Structures," *Advances in Mathematics*, vol. 104, pp. 90-154, 1994.
- [6] N. Asai, I. Kubo, H-H. Kuo, "Bell numbers, Log-Concavity and Log-Convexity," *Acta Applicandae Mathematicae*, vol. 63, pp. 79-87, 2000.
- [7] E. Bender, R. Canfield, "Log-Concavity and Related Properties of the Cycle Index Polynomials," *J. of Comb. Theory, Series A*, vol. 74, pp. 55-70, 1995.

- [8] P. Billingsley, *Probability and Measure*, John Wiley and Sons, Series in Probability and Mathematical Statistics, New York, 1995.
- [9] K. Bogart, *Introductory Combinatorics*, Harcourt Brace Jovanovich, San Diego, 1990.
- [10] S. Burris, *Density in Number Systems*, Mathematical Surveys and Monographs, vol. 86, American Mathematical Society, 2001.
- [11] K. J. Compton, "Some Methods for Computing Component Distribution Probabilities in Relational Structures," *Discrete Mathematics*, Vol. 66, pp. 59-77, 1987.
- [12] N. De Bruijn, *Asymptotic Methods in Analysis*, Dover Publications, Inc. New York, 1981.
- [13] M. Drmota, D. Gardy and B. Gittenberger, "General urn models with several types of balls and Gaussian limiting fields," in revision.
- [14] M. Drmota, D. Gardy and B. Gittenberger, "A unified presentation of some urn models," in revision.
- [15] V. Eramo, M. Listanti, C. Nuzman, and P. Whiting, "Optical Switch Dimensioning and the Classical Occupancy Problem," to appear in *International Journal on Communication Systems*.
- [16] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I, Wiley and Sons, New York, 1967.
- [17] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. II, Wiley and Sons, New York, 1967.
- [18] P. Flajolet, A. Odlyzko, "Singularity Analysis of Generating Functions," *SIAM J. on Disc. Math.*, vol. 3, no. 2, pp.216-240, 1990.
- [19] B. Fristedt, "The Structure of Random Partitions of Large Integers," *Trans. of the American Math. Soc.*, vol. 337, no. 2, 1993.
- [20] G.H. Gonnet, J.I. Munro, "The Analysis of Linear Probing Sort by the use of a New Mathematical Transform," *Journal of Algorithms*, vol. 5, pp. 451-470, 1984.
- [21] P. Grabner, "Searching for Losers," *Random Structures and Algorithms*, vol. 4, no. 1, 1993.
- [22] R. Graham, D. Knuth, O. Patashnik, *Concrete Mathematics*, Addison-Wesley Publishing Company, 1994.
- [23] G.H. Hardy, *Divergent Series*, Clarendon Press, Oxford, 1949.
- [24] M. Hofri, *Analysis of Algorithms: Computational Methods and Mathematical tools*, Oxford University Press, 1995.

- [25] L. Holst, "On Birthday, Collector's, Occupancy and other Classical Urn Problems," *International Statistical Review*, vol. 54, no. 1, pp. 15-27, 1986.
- [26] L. Holst, "The Poisson-Dirichlet Distribution and its Relatives Revisited," in revision.
- [27] P. Jacquet, W. Szpankowski, "Asymptotic Behavior of the Lempel-Ziv Parsing Scheme and Digital Search Trees," *Theoretical Computer Science*, vol. 144, pp. 161-197, 1995.
- [28] A. Y. Khinchin, "On the Analytical Apparatus of Physical Statistics," *Trudy Mat. Inst. Steklov*, vol. 33, 1950.
- [29] V. Kolchin, *Random Mappings*, Optimization Software Inc. Publishing Company, 1986.
- [30] O. Milenkovic, K.J. Compton, "Probabilistic Analysis of Gosper's Algorithm," *in submission*.
- [31] Z. Nehari, *Introduction to Complex Analysis*, Allyn and Bacon Inc., Boston, 1961.
- [32] C. Nuzman, P. Whiting, "Calculus of Variations and Sample Path Large Deviations and the Classical Occupancy Problem," *Allerton Conf. on Commun., Comp., and Control*, September 2001.
- [33] A. Odlyzko, P. Flajolet, "Asymptotic Enumeration Methods," *Handbook of Combinatorics*, vol. 1, no. 2, Elsevier, Amsterdam, pp. 1063-1229, 1995.
- [34] P.V. Pobleto, J.I. Munro, T. Papadakis, "The Binomial Transform and its Applications to the Analysis of Skip Lists," *Proceedings of the 3rd European Symposium on Algorithms, ESA95*, Lecture Notes in Computer Science 979, Corfu, Greece, 1995.
- [35] L.A. Shepp, S.P. Lloyd, "Ordered Cycle Lengths in a Random Permutation," *Trans. Amer. Math. Soc.*, vol. 121, pp. 340-357, 1966.
- [36] R. Simha, A. Majumdar, "An urn model with applications to Database performance evaluation," in submission.
- [37] R. Stanley, *Enumerative Combinatorics*, vol. 1, Cambridge University Press, 1997.
- [38] R. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge University Press, 1999.
- [39] W. Szpankowski, *Average Case Analysis of Algorithms on Sequences*, J. Wiley, New York, 2001.
- [40] N. Temme, "Asymptotic Estimates of Stirling Numbers," *Studies in Applied Mathematics*, vol. 89, no. 3, pp. 233-243, 1993.
- [41] W.D. Warde, S.K. Katti, "Infinite Divisibility of Discrete Distributions II," *Ann. Math. Stat.*, vol. 42, no. 3, pp. 1088-1090.

- [42] H. Wilf, *Generatingfunctionology*, Academic Press, Boston, 1994.
- [43] E. Wong, A. Chan, P. Yum, "Analysis of Rerouting in Circuit-Switched Networks," *IEEE/ACM Trans. on Networking*, vol. 8, no. 3, June 2000.
- [44] A. R. Woods, "Counting Finite Models," *Journal of Symbolic Logic*, Vol. 62, pp. 925-949, 1997.