Couplings for irregular combinatorial assemblies

A. D. Barbour*and A. Pósfai†
Universität Zürich and Tufts University

Abstract

When approximating the joint distribution of the component counts of a decomposable combinatorial structure that is 'almost' in the logarithmic class, but nonetheless has irregular structure, it is useful to be able first to establish that the distribution of a certain sum of non-negative integer valued random variables is smooth. This distribution is not like the normal, and individual summands can contribute a non-trivial amount to the whole, so its smoothness is somewhat surprising. In this paper, we consider two coupling approaches to establishing the smoothness, and contrast the results that are obtained.

Keywords: Logarithmic combinatorial structures, Dickman's distribution, Mineka coupling, compound Poisson
AMS subject classification: 60C05, 60F05, 05A16
Running head: Couplings for combinatorial assemblies

1 Introduction

Many of the classical random decomposable combinatorial structures, such as random permutations and random polynomials over a finite field, have component structure satisfying a *conditioning relation*: if $C_i^{(n)}$ denotes the number of components of size i, the distribution of the vector of component counts $(C_1^{(n)}, \ldots, C_n^{(n)})$ of a structure of size n can be expressed as

$$\mathscr{L}\bigl(C_1^{(n)}, \ldots, C_n^{(n)}\bigr) = \mathscr{L}\bigl(Z_1, \ldots, Z_n \bigm| T_{0,n} = n\bigr)\,, \tag{1.1}$$

where $(Z_i,\, i \geq 1)$ is a fixed sequence of independent non-negative integer valued random variables, and $T_{a,n} := \sum_{i=a+1}^{n} iZ_i$, $0 \leq a < n$. If, as in the examples above, the Z_i also satisfy

$$i\mathbb{P}[Z_i = 1] \;\to\; \theta \qquad \text{and} \qquad \theta_i \;:=\; i\mathbb{E}Z_i \;\to\; \theta\,, \tag{1.2}$$

the combinatorial structure is called *logarithmic*. It is shown in Arratia, Barbour & Tavaré (2003) [ABT] that combinatorial structures satisfying the conditioning relation

*Angewandte Mathematik, Universität Zürich, Winterthurertrasse 190, CH-8057 ZÜRICH; E-mail a.d.barbour@math.uzh.ch;

†Department of Mathematics, Tufts University, 503 Boston Avenue, Medford, MA 02155, USA and Analysis and Stochastics Research Group of the Hungarian Academy of Sciences, Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, Szeged 6720, Hungary; E-mail anna.posfai@tufts.edu

and slight strengthenings of the logarithmic condition share many common properties. For instance, if $L^{(n)}$ is the size of the largest component, then

$$n^{-1}L^{(n)} \to_d L, \tag{1.3}$$

where L has probability density function $f_\theta(x) := e^{\gamma\theta}\Gamma(\theta+1)x^{\theta-2}p_\theta((1-x)/x)$, $x \in (0,1]$, and p_θ is the density of the Dickman distribution P_θ with parameter θ, given in Vervaat (1972, p. 90). Furthermore, for any sequence $(a_n,\, n \ge 1)$ with $a_n = o(n)$,

$$\lim_{n\to\infty} d_{\mathrm{TV}}\bigl(\mathscr{L}(C_1^{(n)},\ldots,C_{a_n}^{(n)}), \mathscr{L}(Z_1,\ldots,Z_{a_n})\bigr) = 0\,. \tag{1.4}$$

Both of these convergence results can be complemented by estimates of the approximation error, under appropriate conditions.

If the logarithmic condition is not satisfied, as in certain of the additive arithmetic semigroups introduced in Knopfmacher (1979), the results in [ABT] are not directly applicable. However, in Manstavičius (2009) and in Barbour & Nietlispach (2010) [BN], it is shown that the logarithmic condition can be relaxed to a certain extent, without disturbing the validity of (1.4), and that (1.3) can also be recovered, if the convergence in (1.2) is replaced by a weaker form of convergence. A key step in the proofs of these results is to be able to show that, for sequences $a_n = \mathrm{o}(n)$, the normalized sum $n^{-1}T_{a_n,n}$ converges both in distribution and locally to the Dickman distribution P_θ, and that the error rates in these approximations can be controlled. To do so, it is in turn necessary to be able to show that, under suitable conditions,

$$\lim_{n\to\infty} d_{\mathrm{TV}}\bigl(\mathscr{L}(T_{a_n,n}), \mathscr{L}(T_{a_n,n}+1)\bigr) \;=\; 0\,, \qquad \text{for all } a_n = \mathrm{o}(n), \tag{1.5}$$

and that the error rate can be bounded by a power of $\{(a_n+1)/n\}$. In this note, we explore ways of using coupling to prove such estimates, in the simplest case in which the Z_i have Poisson distributions. The first of these, an improvement over the Mineka coupling, was introduced in [BN]. It is extremely flexible in obtaining error rates bounded by a power of $\{(a_n+1)/n\}$ for a wide variety of choices of the means θ_i, and it is in no way restricted to Poisson distributed Z_i's. Here, we show that, despite its attractions, it does not achieve the best possible error rate under ideal circumstances. The second approach works only in much more restricted situations, but is then capable of attaining the theoretically best results.

In the case of Poisson distributed Z_i, the distribution of T_{an} is a particular compound Poisson distribution, with parameters determined by n and by the θ_i, and it is tempting to try to approximate the distribution of $n^{-1}T_{a_n,n}$ by first approximating by the distribution that would be obtained if $\theta_i = \theta$ for all i. A natural way of obtaining compound Poisson approximation is then to use Stein's method (Barbour, Chen & Loh 1992). Difficulties arise, however, because the conditions of their Theorem 5 (needed to get useful bounds on the solution to the Stein equation) are not satisfied unless $a = 0$, and, even then, the bounds obtained are not as useful as they might be; better information for this particular case can be found in [ABT, Chapter 9]. And, even using this approach, it still seems necessary first to bound the error in (1.5), in order to obtain useful results.

2 A Mineka–like coupling

Let $\{X_i\}_{i\in\mathbb{N}}$ be mutually independent $\mathbb{Z}$-valued random variables, and let $S_n := \sum_{i=1}^n X_i$. The Mineka coupling, developed independently by Mineka (1973) and Rösler (1977) (see also Lindvall (2002, Section II.14)) yields a bound of the form

$$d_{\mathrm{TV}}\bigl(\mathscr{L}(S_n), \mathscr{L}(S_n+1)\bigr) \;\le\; \Bigl(\frac{\pi}{2}\sum_{i=1}^n u_i\Bigr)^{-1/2}, \tag{2.1}$$

where

$$u_i \;:=\; \Bigl(1 - d_{\mathrm{TV}}\bigl(\mathscr{L}(X_i), \mathscr{L}(X_i+1)\bigr)\Bigr);$$

see Mattner & Roos (2007, Corollary 1.6). The proof is based on coupling copies $\{X_i'\}_{i\in\mathbb{N}}$ and $\{X_i''\}_{i\in\mathbb{N}}$ of $\{X_i\}_{i\in\mathbb{N}}$ in such a way that

$$V_n \;:=\; \sum_{i=1}^n \bigl(X_i - X_i'\bigr), \qquad n\in\mathbb{N},$$

is a symmetric random walk with steps in $\{-1,0,1\}$. Writing $S_i' := 1 + \sum_{j=1}^i X_j' \sim S_i + 1$ and $S_i'' := \sum_{j=1}^i X_j'' \sim S_i$, so that $V_i + 1 = S_i' - S_i''$, the coupling inequality (Lindvall 1992, Section I.2) then shows that

$$d_{\mathrm{TV}}\bigl(\mathscr{L}(S_n), \mathscr{L}(S_n+1)\bigr) \;\le\; \mathbb{P}[\tau > n] \;=\; \mathbb{P}[V_n = \{-1,0\}],$$

where τ is the time at which $\{V_n\}_{n\in\mathbb{Z}_+}$ first hits level -1, and the last equality follows from the reflection principle. However, this inequality gives slow convergence rates, if $X_i = iZ_i$ and the Z_i are as described in the Introduction; typically, $d_{\mathrm{TV}}\bigl(\mathscr{L}(iZ_i), \mathscr{L}(iZ_i+1)\bigr)$ is equal to 1, and, if X_i is taken instead to be $(2i-1)Z_{2i-1} + 2iZ_{2i}$, we still expect to have $1 - d_{\mathrm{TV}}\bigl(\mathscr{L}(X_i), \mathscr{L}(X_i+1)\bigr) \asymp i^{-1}$, leading to bounds of the form

$$d_{\mathrm{TV}}\bigl(\mathscr{L}(T_{a_n,n}), \mathscr{L}(T_{a_n,n}+1)\bigr) = \mathrm{O}\bigl((\log(n/\{a_n+1\}))^{-1/2}\bigr). \tag{2.2}$$

The reason that the Mineka coupling does not work efficiently in our setting is that, once the random walk V_n takes some value k, it has to achieve a preponderance of $k+1$ negative steps, in order to get to the state -1, and this typically requires many jumps to realize. Since, at the i-th step, the probability of there being a jump is of order i^{-1}, it thus takes a very long time for such an event to occur, and the probability of this not happening before time n is then relatively large. In [BN], the difficulty is overcome by observing that the Mineka random walk can be replaced by another Markov chain $(\widetilde{V}_n, n\ge 1)$, still constructed from copies $(Z_i', i\ge 1)$ and $(Z_i'', i\ge 1)$ of the original sequence, but now associated differently with one another. The basic idea is to note that, if $\widetilde{V}_i = k$, then the random variables $X_{i+1}' := jZ_j' + (j+k+1)Z_{j+k+1}'$ and $X_{i+1}'' := jZ_j'' + (j+k+1)Z_{j+k+1}''$ can be coupled in such a way that $X_{i+1}' - X_{i+1}'' \in \{-(k+1), 0, (k+1)\}$, for any j such that the indices j and $j+k+1$ have not previously been used in the construction. Hence a single jump has probability 1/2 of making $\widetilde{V}$ reach -1. The construction starts as for the Mineka walk, but if the first jump takes $\widetilde{V}$ to $+1$, then the chain switches to jumps in $\{-2,0,2\}$; and subsequently, if $\widetilde{V}$ is in the state $k = 2^r - 1$, the chain makes jumps in

$\{-2^r, 0, 2^r\}$. Clearly, this construction can be used with $Z_i \sim \mathrm{Po}(i^{-1}\theta_i)$, even when many of the θ_i are zero. A number of settings of this kind are explored in detail in [BN]; for instance, when $\theta_i \geq \theta^*$ for all i in $\{r\mathbb{Z}_+ + t\} \cup \{s\mathbb{Z}_+ + u\}$, where r and s are coprime. Very roughly, provided that a non-vanishing fraction of the θ_i exceed some fixed value $\theta_* > 0$, the probability that $\widetilde{V}$ reaches -1 before time n is of order $n^{-\alpha}$, for some $\alpha > 0$, an error probability exponentially smaller than that in (2.2).

Here, we make the following observation. Suppose that we have the ideal situation in which $\theta_i = \theta^* > 0$ for *every* i. Then the probability that a coupling, constructed as above, should fail is at least of magnitude $n^{-\theta^*/2}$. In Section 3, it is shown that the total variation distance in (1.5) is actually of order $n^{-\min\{\theta^*,1\}}$ under these circumstances, so that the estimates of this distance obtained by the [BN] coupling are typically rather weaker. It is thus of interest to find ways of attaining sharper results. The coupling given in Section 3 is one such, but it is much less widely applicable.

The coupling approach given in [BN] evolves by choosing a pair of indices $M_{i1} < M_{i2}$ at each step i, with the choice depending on the values previously used: no index can be used more than once, and $M_{i2} - M_{i1} = \widetilde{V}_{i-1} + 1$, so that one jump in the right direction leads immediately to a successful coupling. Then, if $(M_{i1}, M_{i2}) = (j, j+k+1)$, the pair X_i' and X_i'' is constructed as above, by way of copies of the random variables jZ_j and $(j+k+1)Z_{j+k+1}$. The probability of a jump taking place is then roughly $2\theta^*/(j+k+1)$, and, if a jump occurs, it has probability $1/2$ of taking the value $-(k+1)$, leading to success. The main result of this section is the following lower bound for the failure probability of such a procedure.

Theorem 2.1. *For any coupling constructed as above, the probability* $\mathbb{P}[F]$ *that the coupling is not successful is bounded below by*

$$P[F] \;\geq\; \prod_{i=1}^{\lfloor n/2 \rfloor} \left(1 - \tfrac{1}{2}\min\{\theta^*/i, 1/e\}\right) \;\asymp\; n^{-\theta^*/2}.$$

Proof. In order to prove the lower bound, we couple two processes, one of which makes more jumps than the other. We start by letting $(U_i,\, i \geq 1)$ be independent uniform random variables on $[0,1]$. The first process is much as discussed above. It is defined by a sequence of pairs of indices $M_{i1} < M_{i2}$, $1 \leq i \leq I^*$, from $[n] := \{i \in \mathbb{N}\colon\, i \leq n\}$, with $I^* \leq \lfloor n/2 \rfloor$ the last index for which a suitable pair can be found. No index is ever used twice, and the choice of (M_{i1}, M_{i2}) is allowed to depend on $((M_{j1}, M_{j2}, U_j),\, 1 \leq j < i)$. We set $Y_i = I[U_i \leq p(M_{i1}, M_{i2})]$, where

$$p(m_1, m_2) \;:=\; 2e^{-\theta^*/m_1}\,(\theta^*/m_2)e^{-\theta^*/m_2} \;<\; 1/e,$$

for $m_1 < m_2$, representing the indicator of a jump of $\pm(m_2 - m_1)$ being made by the first process at time i. For the second, we inductively define $R_i := \{\rho(1), \ldots, \rho(i)\}$ by taking $R_0 = \emptyset$ and

$$\rho(i) \;:=\; \max\{r \in [n/2] \setminus R_{i-1}\colon\; 2r \leq M_{i2}\};$$

we shall check at the end of the proof that $\rho(i)$ always exists. (The second process, that we do not really need in detail, uses the pair $(2\rho(i) - 1, 2\rho(i))$ at stage i.) We then define

$Z_i := I[U_i \le \min\{\theta^*/\rho(i), 1/e\}]$, noting that $p(M_{i1}, M_{i2}) \le \min\{\theta^*/\rho(i), 1/e\}$, entailing $Z_i \ge Y_i$ a.s. for all i. Finally, let $(J_i,\, i \ge 1)$ be distributed as $\mathrm{Be}(1/2)$, independently of each other and everything else.

The event that the first process makes no successful jumps can be described as the event

$$F \;:=\; \left\{ \sum_{i=1}^{I^*} Y_i J_i = 0 \right\}.$$

We thus clearly have

$$F \;\supset\; \left\{ \sum_{i=1}^{\lfloor n/2 \rfloor} Z_i J_i = 0 \right\},$$

where, for $I^* < i \le n/2$, we take $\rho(i) := \min\{r \in [n/2] \setminus R_{i-1}\}$, and $R_i := R_{i-1} \cup \{\rho(i)\}$. But now the Z_i, suitably reordered, are just independent Bernoulli random variables with means $\min\{\theta^*/r, 1/e\}$, $1 \le r \le n/2$, and hence

$$P[F] \;\ge\; \prod_{i=1}^{\lfloor n/2 \rfloor} (1 - \tfrac{1}{2} \min\{\theta^*/i, 1/e\}) \;\asymp\; n^{-\theta^*/2}.$$

It remains to show that the $\rho(i)$ are well defined at each stage, which requires that

$$S_i \;:=\; \{r \in [n/2] \setminus R_{i-1}\colon\; 2r \le M_{i2}\} \;\ne\; \emptyset,$$

$1 \le i \le I^*$. For $i = 1$, $m_{12} \ge 2$, so the start is successful. Now, for $2 \le i \le n/2$, suppose that

$$r(i-1) \;:=\; \max\{s\colon\; R_{i-1} \supset \{1, 2, \ldots, s\}\}.$$

Then $1, 2, \ldots, r(i-1)$ can be expressed as $\rho(i_1), \rho(i_2), \ldots, \rho(i_{r(i-1)})$, for some indices $i_1, i_2, \ldots, i_{r(i-1)}$. For these indices, we have $M_{i_l,2} \le 2r(i-1)+1$, $1 \le l \le r(i-1)$, since $r(i-1)+1 \notin R_{i-1}$ and, from the definition of $\rho(\cdot)$, we could thus not choose $\rho(i_l) \le r(i-1)$ if $M_{i_l,2} \ge 2r(i-1)+2$. Hence, also, $M_{i_l,1} \le 2r(i-1)+1$, and, because all the M_{is} are distinct, $\{M_{i_l,s},\, 1 \le s \le 2, 1 \le l \le r(i-1)\}$ is a set of $2r(i-1)$ elements of $[2r(i-1)+1]$. Thus, when choosing the pair (M_{i1}, M_{i2}), there is only at most one element of $[2r(i-1)+1]$ still available for choice, from which it follows that $M_{i2} \ge 2r(i-1)+2$: so $r(i-1)+1 \in S_i$, and hence S_i is not empty. □

3 A Poisson–based coupling

In this section, we show that a coupling can be constructed that gives good error rates in (1.5) when $Z_j \sim \mathrm{Po}(j^{-1}\theta^*)$, for some fixed $\theta^* > 0$. If $Z_j \sim \mathrm{Po}(j^{-1}\theta_j)$ with $\theta_j \ge \theta^*$, the same order of error can immediately be deduced (though it may no longer be optimal), since, for Poisson random variables, we can write $T_{an} = T^*_{an} + T'$, with T^*_{an} constructed from independent random variables $Z^*_j \sim \mathrm{Po}(j^{-1}\theta^*)$, and with T' independent of T^*_{an}.

Because of the Poisson assumption, the distribution of $T_{an} := \sum_{j=a+1}^{n} jZ_j$ can equivalently be re-expressed as that of a sum of a random number $N \sim \mathrm{Po}(\theta^* h_{an})$ of independent copies of a random variable X having $\mathbb{P}[X = j] = 1/\{jh_{an}\}$, $a+1 \le j \le n$, where $h_{an} := \sum_{j=a+1}^{n} j^{-1}$. Fix $c > 1$, define $j_r := \lfloor c^r \rfloor$, and set

$$r_0 \ := \ r_0(a) \ := \ \lceil \log_c(a+1) \rceil, \qquad r_1 \ := \ r_1(n) \ := \ \lfloor \log_c n \rfloor.$$

Define independent random variables $(X_{ri},\, r_0 \le r < r_1,\, i \ge 1)$ and $(N_r,\, r_0 \le r < r_1)$, with $N_r \sim \mathrm{Po}(\theta^* h_{an} p_r)$ and

$$\mathbb{P}[X_{ri} = j] \ = \ 1/\{jh_{an}p_r\}, \qquad j_r \le j < j_{r+1},$$

where

$$p_r \ := \ \sum_{j=j_r}^{j_{r+1}-1} \frac{1}{jh_{an}};$$

define $\overline{P}_r := \sum_{s=r}^{r_1-1} p_s \le 1$. Then we can write T_{an} in the form

$$T_{an} \ = \ Y + \sum_{r=r_0}^{r_1-1} \sum_{i=1}^{N_r} X_{ri},$$

where Y is independent of the sum; the X_{ri} represent the realizations of the copies of X that fall in the interval $C_r := [j_r, j_{r+1})$, and Y accounts for all X-values not belonging to one of these intervals. The idea is then to construct copies T'_{an} and T''_{an} of T_{an} with T'_{an} coupled to $T''_{an} + 1$, by using the same N_r for both, and trying to couple one pair X'_{ri} and $X''_{ri} + 1$ exactly, declaring failure if this doesn't work. Clearly, such a coupling can only be attempted for an r for which $N_r \ge 1$. Then exact coupling can be achieved between X'_{r1} and $X''_{r1} + 1$ with probability $1 - 1/\{j_r h_{an} p_r\}$, since the point probabilities for X_{r1} are decreasing. Noting that the p_r are all of the same magnitude, it is thus advantageous to try to couple with r as large as possible. This strategy leads to the following theorem.

Theorem 3.1. *With $Z_j \sim \mathrm{Po}(j^{-1}\theta^*)$, $j \ge 1$, we have*

$$d_{\mathrm{TV}}\big(\mathscr{L}(T_{an}), \mathscr{L}(T_{an}+1)\big) \ = \ O(\{(a+1)/n\}^{\theta^*} + n^{-1}),$$

if $\theta^* \ne 1$; for $\theta^* = 1$,

$$d_{\mathrm{TV}}\big(\mathscr{L}(T_{an}), \mathscr{L}(T_{an}+1)\big) \ = \ O(\{(a+1)/n\} + n^{-1}\log\{n/(a+1)\}).$$

Proof. We begin by defining

$$B_r \ := \ \left(\bigcap_{s=r+1}^{r_1-1} \{N_s = 0\} \right) \cap \{N_r \ge 1\}, \qquad r_0 \le r < r_1,$$

and setting $B_0 := \bigcap_{s=r_0}^{r_1-1}\{N_s = 0\}$. On the event B_r, write $X''_{r1} = X'_{r1} - 1$ if $X'_{r1} \ne j_r$, with X''_{r1} so distributed on the event $A_r := \{X'_{r1} = j_r\}$ that its overall distribution is correct. All other pairs of random variables $X'_{r'i}$ and $X''_{r'i}$, $(r', i) \in ([r_0, \ldots, r_1 - 1] \times \mathbb{N}) \setminus \{(r, 1)\}$, are set to be equal on B_r. This generates copies T'_{an} and T''_{an} of T_{an}, with the property that $T'_{an} = T''_{an} + 1$, except on the event

$$E \;:=\; B_0 \cup \left(\bigcup_{r=r_0}^{r_1-1} (B_r \cap A_r) \right).$$

It is immediate from the construction that

$$\mathbb{P}[B_r] \;=\; \exp\{-\theta^* h_{an} \overline{P}_{r+1}\}(1 - e^{-\theta^* h_{an} p_r}), \quad r_0 \le r < r_1,$$

and that $\mathbb{P}[B_0] = \exp\{-\theta^* h_{an} \overline{P}_{r_0}\}$; and $\mathbb{P}[A_r \mid B_r] = 1/\{j_r h_{an} p_r\}$. This gives all the ingredients necessary to evaluate the probability

$$\mathbb{P}[E] \;=\; \mathbb{P}[B_0] + \sum_{r=r_0}^{r_1-1} \mathbb{P}[B_r]\mathbb{P}[A_r \mid B_r].$$

In particular, as $r \to \infty$, $j_r \sim c^r$, $h_{an} p_r \sim \log c$ and $h_{an} \overline{P}_{r+1} \sim (r_1(n) - r) \log c$, from which it follows that $\mathbb{P}[B_r] \sim c^{-\theta^*(r_1(n)-r)}(1 - c^{-\theta^*})$, $\mathbb{P}[A_r \mid B_r] \sim 1/\{c^r \log c\}$ and

$$\mathbb{P}[B_0] \;\asymp\; c^{-\theta^*(r_1(n)-r_0(a))} \;\asymp\; \{(a+1)/n\}^{\theta^*}.$$

Combining this information, we arrive at

$$\mathbb{P}[E] \;\asymp\; \{(a+1)/n\}^{\theta^*} + \sum_{r=r_0(a)}^{r_1(n)-1} c^{-r}\, c^{-\theta^*(r_1(n)-r)}.$$

For $\theta^* > 1$, the dominant term in the sum is that with $r = r_1(n) - 1$, and it follows from the definition of $r_1(n)$ that then

$$\mathbb{P}[E] \;\asymp\; \{(a+1)/n\}^{\theta^*} + c^{-r_1(n)} \asymp \{(a+1)/n\}^{\theta^*} + n^{-1}.$$

For $\theta^* < 1$, the dominant term is that with $r = r_0(a)$, giving

$$\mathbb{P}[E] \;\asymp\; \{(a+1)/n\}^{\theta^*} + n^{-\theta^*}(a+1)^{-(1-\theta^*)} \;\asymp\; \{(a+1)/n\}^{\theta^*}.$$

For $\theta^* = 1$, all terms in the sum are of the same order, and we get

$$\mathbb{P}[E] \;\asymp\; \{(a+1)/n\} + n^{-1} \log(n/(a+1)). \qquad \square$$

Note that the element $\{(a+1)/n\}^{\theta^*}$ appearing in the errors is very easy to interpret, and arises from the probability of the event that $T_{an} = 0$, a value unattainable by $T_{an} + 1$. Furthermore, the random variable T_{an} has some point probabilities of magnitude n^{-1} [ABT, p. 91], so that n^{-1} is always a lower bound for the order of $d_{\mathrm{TV}}\big(\mathscr{L}(T_{an}), \mathscr{L}(T_{an}+1)\big)$. Hence the order of approximation in Theorem 3.1 is best possible if $\theta^* \ne 1$. However, for $a = 0$ and $\theta^* = 1$, the point probabilities of T_{0n} are decreasing, and since their maximum is of order $O(n^{-1})$, the logarithmic factor in the case $\theta^* = 1$ is not sharp, at least for $a = 0$.

The method of coupling used in this section can be extended in a number of ways. For instance, it can be used for random variables Z_j with distributions other than Poisson, giving the same order of error as long as $d_{\mathrm{TV}}(\mathscr{L}(Z_j), \mathrm{Po}(\theta_j/j)) = O(j^{-2})$. This is because, first, for some $K < \infty$,

$$\mathbb{P}[B_r] \;\le\; K\exp\{-\theta^* h_{an}\overline{P}_{r+1}\}(1 - e^{-\theta^* h_{an} p_r}), \quad r_0 \le r < r_1,$$

and $\mathbb{P}[B_0] \le K\exp\{-\theta^* h_{an}\overline{P}_{r_0}\}$, where, in the definitions of the B_r, the events $\{N_s = 0\}$ are replaced by $\{Z_j = 0,\, j_s \le j < j_{s+1}\}$. Secondly, we immediately have

$$d_{\mathrm{TV}}(\mathscr{L}((Z_j,\, j_r \le j < j_{r+1})), (\widehat{Z}_j,\, j_r \le j < j_{r+1})) \;=\; O(j_r^{-1}),$$

where the $\widehat{Z}_j \sim \mathrm{Po}(\theta_j/j)$ are independent, and hence that

$$d_{\mathrm{TV}}(\mathscr{L}(T_{j_r-1, j_{r+1}-1}), \mathscr{L}(\widehat{T}_{j_r-1, j_{r+1}-1})) \;=\; O(j_r^{-1}),$$

where $\widehat{T}_{rs}$ is defined as T_{rs}, but using the $\widehat{Z}_j$. Thus, on the event B_r, coupling can still be achieved except on an event of probability of order $O(j_r^{-1})$.

It is also possible to extend the argument to allow for gaps between the intervals on which $\theta_j \ge \theta^*$. Here, for $0 < c_1 \le c_2$, the intervals $[j_r, j_{r+1}-1]$ can be replaced by intervals $[a_r, b_r]$, such that $b_r/a_r \ge c_1$ and $a_r \ge kac_2^r$ for some k and for each $1 \le r \le R$, say. The argument above then leads to a failure probability of at most

$$O\left(c_1^{-R\theta^*} + \sum_{r=1}^{R} \frac{1}{ac_2^r}\, c_1^{-\theta^*(R-r)}\right).$$

If $c_1^{\theta^*} > c_2$, the failure probability is thus at most of order $O(c_1^{-R\theta^*} + 1/\{ac_2^R\})$; if $c_1^{\theta^*} < c_2$, it is of order $O(c_1^{-R\theta^*})$. In Theorem 3.1 above, we have $c_1 = c_2 = c$, $k = 1$ and $c^{-R} \asymp (a+1)/n$, and the results are equivalent.

However, the method is still only useful if there are long stretches of indices j with θ_j uniformly bounded below. This is in contrast to that discussed in the previous section, which is flexible enough to allow sequences θ_j with many gaps. It would be interesting to know of other methods that could improve the error bounds obtained by these methods.

References

R. Arratia, A. D. Barbour & S. Tavaré (2003). *Logarithmic combinatorial structures: a probabilistic approach*, European Mathematical Society Press.

A. D. Barbour, L. H. Y. Chen & W.-L. Loh (1992). Compound Poisson approximation for nonnegative random variables via Stein's method, *Ann. Probab.* **20**, 1843–1866.

A. D. Barbour & B. Nietlispach (2010). Approximation by the Dickman distribution and quasi–logarithmic combinatorial structures, `arXiv:1007.5269`.

J. Knopfmacher (1979). *Analytic arithmetic of algebraic number fields*, Lecture notes in pure and applied mathematics **50**, Marcel Dekker, New York.

T. Lindvall (2002). *Lectures on the coupling method*, Dover Publications, Mineola, NY.

E. Manstavičius (2009). Strong convergence on weakly logarithmic combinatorial assemblies, arXiv:0903.1051.

L. Mattner & B. Roos (2007). A shorter proof of Kanter's Bessel function concentration bound, *Prob. Theory Rel. Fields* **139**, 191–205.

J. Mineka (1973). A criterion for tail events for sums of independent random variables, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **25**, 163–170.

U. Rösler (1977). Das 0-1-Gesetz der terminalen σ-Algebra bei Harrisirrfahrten, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **37**, 227–242.

W. Vervaat (1972). *Success epochs in Bernoulli trials with applications in number theory*, Mathematical Centre Tracts **42**, Math. Centrum, Amsterdam.