Modern Stochastics: Theory and Applications 14 (1) (2027) 1–21
https://doi.org/10.15559/26-VMSTA306

Limit theorems in the extended coupon collector's problem

Andrii Ilienko[a,b]

[a]*Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, CH-3012, Bern, Switzerland*
[b]*Department of Mathematical Analysis and Probability Theory, Igor Sikorsky Kyiv Polytechnic Institute, Prospect Beresteiskyi 37, 03056, Kyiv, Ukraine*

andrii.ilienko@unibe.ch (A. Ilienko)

Received: 2 June 2026, Revised: 3 August 2026, Accepted: 3 August 2026,
Published online: 11 August 2026

Abstract We consider an extended variant of the classical coupon collector's problem with an infinite number of collections. An arriving coupon is placed in the rth collection, $r \geq 0$, if r is the smallest index such that the corresponding collection still does not have a coupon of this type. We derive distributional limit theorems for the number of empty spots in different collections at the time when the 0th collection was completed, as well as after some delay. We also obtain the joint limiting distribution for completion times of different collections. All main results are given in an ultimate infinite-dimensional form in the sense of distributional convergence in $\mathbb{R}^{\infty}$. The main tool in the proofs is convergence of specially constructed point processes.

Keywords Coupon collector's problem, balls-into-bins model, multivariate geometric distribution, Gumbel distribution, logistic distribution, convergence of point processes, Poisson processes, poissonization, thinning

2020 MSC 60C05, 60F05, 60G55

1 Introduction

The coupon collector's problem is undoubtedly one of the most popular classical problems in combinatorial probability. It is not only of purely theoretical interest, but also finds numerous applications (see, e.g., [8, 25, 1, 4, 5]). More recently, an extended

© 2027 The Author(s). Published by VTeX. Open access article under the CC BY license.

version of this problem, known in the literature as the collector's brotherhood problem, has attracted considerable attention; see, e.g., [15, 14, 2, 11].

Following [14], we give its statement as follows. A person collects coupons, each of which belongs to one of $n \in \mathbb{N}$ different types. The coupons arrive one by one at discrete times, the type of each coupon being equiprobable and independent of types of preceding ones. Each time the person receives a coupon which he does not yet have, he puts it in his album. Otherwise, he gives it to his younger brother. In his turn, the latter puts it in his own album, if the coupon is new for him. Otherwise, he gives it to the next younger brother, and so on. All the brothers try to complete their own collections, using the same policy. The main collector is labeled 0, the younger brother 1, the next younger brother 2, etc.

For $r \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, let $T_r^{(n)}$ stand for the time the rth person completes his collection. By a classical result due to Erdős and Rényi [12],

$$\begin{aligned} \mathbb{E}T_r^{(n)} &= n \ln n + rn \ln\ln n + (\gamma - \ln r!)n + \mathcal{O}(n), \\ \lim_{n\to\infty} \mathbb{P}\Big\{ \frac{T_r^{(n)}}{n} - \ln n - r \ln\ln n < x \Big\} &= \exp\Big(-\frac{\mathrm{e}^{-x}}{r!}\Big), \quad x \in \mathbb{R}, \end{aligned} \tag{1}$$

with $\gamma = -\Gamma'(1)$ standing for the Euler–Mascheroni constant. So, the limiting distribution is of Gumbel type.

Denote by $U_r^{(n)}$, $r \in \mathbb{N}$, the number of empty spots in the album of the rth brother at time $T_0^{(n)}$, that is, when the main collector completed his album. In [26], by using the optional stopping theorem for a specially constructed martingale, it was shown that $\mathbb{E}U_1^{(n)} = \mathcal{H}_1^{(n)}$, where $\mathcal{H}_1^{(n)} = \sum_{k=1}^n \frac{1}{k}$ stands for the nth harmonic number. Subsequently, it turned out that this can be further generalized to

$$\mathbb{E}U_r^{(n)} = \mathcal{H}_r^{(n)}, \qquad n, r \in \mathbb{N}, \tag{2}$$

where $\mathcal{H}_r^{(n)}$ are hyperharmonic numbers defined by the recursive formula

$$\mathcal{H}_r^{(n)} = \sum_{k=1}^n \frac{\mathcal{H}_{r-1}^{(k)}}{k}, \qquad r \geq 2.$$

This was proved in [15] and [14] by using a non-elementary generating function argument, and in [2] by a simpler and more direct probabilistic reasoning. Moreover, the latter argument made it possible to obtain another, more explicit representation:

$$\mathcal{H}_r^{(n)} = \sum_{k=1}^n \binom{n}{k} \frac{(-1)^{k+1}}{k^r}, \qquad r \geq 1.$$

By (10.1) in [15], for the above hyperharmonic numbers the following asymptotic formula holds:

$$\frac{r!\,\mathcal{H}_r^{(n)}}{\ln^r n} \to 1 \quad \text{as } n \to \infty. \tag{3}$$

It should be noted that the term "hyperharmonic numbers" often refers to a different numerical sequence (see, e.g., the corresponding article in Wikipedia).

Note that this problem allows for different, though equivalent, formulations. For instance, we give another one in terms of a balls-into-bins model. Suppose we have an infinite sequence of balls and n bins of unlimited capacity. Each time a single ball is placed into one of the bins, which is chosen equiprobably and independently of the previous history. In this setting, $T_r^{(n)}$ means the first time when each bin contains at least $r+1$ balls. Similarly, $U_r^{(n)}$ stands for the number of bins containing at most r balls at time $T_0^{(n)}$, that is, when each bin first contains at least one ball.

Until now, the question of limiting distributions for $U_r^{(n)}$ as $n \to \infty$ remained open. This problem was explicitly stated on p. 446 in [11]. In order to determine these limiting distributions it would be natural to use the generating function of $U_r^{(n)}$. The latter can be easily deduced from Proposition 4.1 in [15]:

$$\mathbb{E}u^{U_r^{(n)}} = nu \sum_{\substack{a+b+c_1+ \\ +\cdots+c_r=n-1}} \binom{n-1}{a,b,c_1,\ldots,c_r} (-1)^b (u-1)^{\sum_{k=1}^r c_k} \frac{(\sum_{k=1}^r kc_k)!}{\prod_{k=1}^r (k!)^{c_k}} (n-a)^{-1-\sum_{k=1}^r kc_k}.$$

It is, however, clear that the generating function of such an ominous form can hardly be used for this purpose. Therefore, as in our previous papers [19, 20], we take a different approach based on the use of point processes. This allows not only to determine the limiting distributions for $U_r^{(n)}$ (looking ahead, they will turn out to be exponential) but also to obtain the limit theorem in an ultimate infinite-dimensional form. In other words, we will prove the distributional convergence of the sequence $\left(U_r^{(n)}, r \in \mathbb{N}\right)$, normalized in a proper way, to a random element $(E, E, \ldots)$ of $\mathbb{R}^\infty$ as $n \to \infty$, where the same $\mathsf{Exp}(1)$-distributed random variable E appears in every coordinate. This means that $U_r^{(n)}$ exhibit a remarkable asymptotic stability. The precise statement is given in Theorem 1 below.

Note that in the proof of this limit theorem we use a rather unusual trick which, hopefully, can be applied elsewhere. Namely, appealing to a certain analogy between normalization and thinning, we prove the distributional convergence for thinned random variables rather than for normalized ones. This convergence can be deduced from that of specially constructed point processes. The way we then return to normalized random variables is described in Section 4.

We also prove another limit theorem which can be called delayed one. The idea behind is as follows. Due to their definition, the random variables $U_r^{(n)}$ increase unboundedly as $n \to \infty$. Thus, in order to obtain a distributional limit, we had to use some normalization. If one still wants to obtain a limit result without any normalization, one should, instead of $U_r^{(n)}$, consider the number of empty spots in the album of the rth brother not at time $T_0^{(n)}$ but after some delay, that is, at time $T_0^{(n)} + d_r^{(n)}$ with some non-random $d_r^{(n)} \to \infty$ as $n \to \infty$. With the right choice of $d_r^{(n)}$, the corresponding numbers of empty spots will not go to infinity as n does, but will instead converge in distribution to some limit law. In Theorem 2, we find this right form of $d_r^{(n)}$ and establish convergence to a geometric distribution. Like the previous result, we state and prove this one in an infinite-dimensional form, that is, in the sense of distributional convergence in $\mathbb{R}^\infty$.

Our third main result, Theorem 3 below, describes the asymptotic behavior of the sequence of completion times, that is, of the (centered and normalized in a proper way) random elements $\big(T_r^{(n)}, r \in \mathbb{N}_0\big)$ in $\mathbb{R}^\infty$. Actually, this is an infinite-dimensional extension of (1), the limit theorem due to Erdős and Rényi. Surprisingly, the limiting random element turns out to consist of independent entries, which means that the completion times $T_r^{(n)}$, after appropriate centering and normalizing, become asymptotically independent. This may seem all the more unexpected since $T_r^{(n)}$ increase in r, which excludes any kind of independence for non-centered completion times. So, this asymptotic independence is achieved only at the expense of centering and normalizing. As a direct application of the latter result, we derive a limit theorem for inter-completion times with limiting logistic distribution.

2 Preliminaries and main results

Denote by $Y_{i,r}^{(n)}$, $n \in \mathbb{N}$, $i \le n$, $r \in \mathbb{N}_0$, the arrival time of the $(r+1)$th coupon of type i. So, at time $Y_{i,r}^{(n)}$ the collector labeled r receives such a coupon into his collection. Hence,

$$\mathbb{P}\big\{Y_{i,r}^{(n)} = k\big\} = \binom{k-1}{r}\Big(\frac{1}{n}\Big)^{r+1}\Big(1-\frac{1}{n}\Big)^{k-r-1}, \quad k \ge r+1,$$

which is one version of the negative binomial distribution, namely, the one that counts trials up to the $(r+1)$th success. Note that

$$T_r^{(n)} = \max_{i \le n} Y_{i,r}^{(n)}, \tag{4}$$

$$U_r^{(n)} = \operatorname{card}\big\{i\colon Y_{i,r}^{(n)} > \max_{i \le n} Y_{i,0}^{(n)}\big\}. \tag{5}$$

For different i, the random variables $Y_{i,r}^{(n)}$ are identically distributed but not independent. There is a standard way to get rid of this dependence, namely, poissonization. This technique goes back at least to Karlin and was systematically exploited by Holst in the context of random allocations (see [18] or [19] for details). Consider a poissonized scheme, that is, assume that coupons arrive at random times with independent $\mathsf{Exp}(1)$-distributed intervals E_j, $j \in \mathbb{N}$. Similarly to the above, denote by $Z_{i,r}^{(n)}$ the arrival time of the $(r+1)$th coupon of type i in the poissonized scheme. Then, for different i, $Z_{i,r}^{(n)}$ are independent gamma-distributed random variables, $Z_{i,r}^{(n)} \sim \Gamma\big(r+1, \frac{1}{n}\big)$, which is a consequence of the thinning theorem for Poisson point processes (see, e.g., Theorem 5.8 in [24]). For any fixed n, the sequences $\big(Y_{i,r}^{(n)}\big)$ and $\big(Z_{i,r}^{(n)}\big)$ can be given on a common probability space and coupled by

$$Z_{i,r}^{(n)} = \sum_{j=1}^{Y_{i,r}^{(n)}} E_j, \quad i \le n,\ r \in \mathbb{N}_0. \tag{6}$$

Moreover, $\big(Y_{i,r}^{(n)}, i \le n, r \in \mathbb{N}_0\big)$ is independent of $(E_j, j \in \mathbb{N})$, because the inter-arrival times E_j do not affect the order in which the coupons of different types arrive.

Finally, note that, along with (5), a similar formula in terms of $Z_{i,r}^{(n)}$ holds:

$$U_r^{(n)} = \operatorname{card}\bigl\{i\colon Z_{i,r}^{(n)} > \max_{i\le n} Z_{i,0}^{(n)}\bigr\}. \tag{7}$$

We now state our first main result, an infinite-dimensional distributional limit theorem for $\bigl(U_r^{(n)}, r\in\mathbb{N}\bigr)$ in the space $\mathbb{R}^\infty$ equipped with the product topology.

Theorem 1. *Let* $E\sim\mathsf{Exp}(1)$. *Then*

$$\left(\frac{r!\,U_r^{(n)}}{\ln^r n}, r\in\mathbb{N}\right)\xrightarrow{d}(E,E,\ldots)\quad \text{in } \mathbb{R}^\infty \text{ as } n\to\infty. \tag{8}$$

Let us now turn to the delayed limit theorem. By analogy with $U_r^{(n)}$, for fixed $d_r^{(n)}\in\mathbb{N}$ denote by $\hat U_r^{(n)}$ the number of empty spots in the album of the rth brother at time $T_0^{(n)}+d_r^{(n)}$:

$$\hat U_r^{(n)} = \operatorname{card}\bigl\{i\colon Y_{i,r}^{(n)} > \max_{i\le n} Y_{i,0}^{(n)} + d_r^{(n)}\bigr\}. \tag{9}$$

The following result asserts that, with the right choice of $d_r^{(n)}$, the random elements $\bigl(\hat U_r^{(n)}, r\in\mathbb{N}\bigr)$ of $\mathbb{R}^\infty$ converge in distribution to some limiting random element with geometrically distributed marginals.

Theorem 2. *Let*

$$d_r^{(n)} = rn\ln\ln n + \mathcal{O}(n),\quad r\in\mathbb{N}, \tag{10}$$

where $\mathcal{O}(n)$, *of course, may differ for different* r. *Let also* $(G_r, r\in\mathbb{N})$ *be a random element of* $\mathbb{R}^\infty$ *with the probability generating function of the following form:*

$$\mathbb{E}\Bigl(\prod_{r=1}^\infty u_r^{G_r}\Bigr) = \Bigl(\mathrm{e} - \sum_{r=1}^\infty \frac{u_r}{r!}\Bigr)^{-1},\quad (u_r, r\in\mathbb{N})\in[0,1]^\infty. \tag{11}$$

Equivalently, $(G_r, r\in\mathbb{N})$ *may be defined as the random sequence* $\bigl(N_r(E/r!), r\in\mathbb{N}\bigr)$, *where* $E\sim\mathsf{Exp}(1)$ *and* N_r *stand for unit-rate Poisson counting processes, independent of each other and of* E.

Then

$$\bigl(\hat U_r^{(n)}, r\in\mathbb{N}\bigr)\xrightarrow{d}(G_r, r\in\mathbb{N})\quad \text{in } \mathbb{R}^\infty \text{ as } n\to\infty. \tag{12}$$

Remark 1. It follows from (11) that the marginal probability generating functions of the limiting random element take the form:

$$\mathbb{E}u^{G_r} = \frac{r!}{r!+1-u},\quad u\in[0,1].$$

This implies that $G_r\sim\mathsf{Geom}\bigl(\frac{r!}{r!+1}\bigr)$, that is,

$$\mathbb{P}\{G_r = k\} = \Bigl(\frac{1}{r!+1}\Bigr)^k\frac{r!}{r!+1},\quad k\in\mathbb{N}_0.$$

It is worth noting that various sums of G_r also have a geometric distribution. Indeed, let $I \subset \mathbb{N}$ and $S_I = \sum_{r\in I} G_r$. Then by (11),

$$\mathbb{E}u^{S_I} = \mathbb{E}\prod_{r\in I} u^{G_r} = \left(\mathrm{e} - \sum_{r\notin I}\frac{1}{r!} - u\sum_{r\in I}\frac{1}{r!}\right)^{-1} = \left(1 + \sum_{r\in I}\frac{1}{r!} - u\sum_{r\in I}\frac{1}{r!}\right)^{-1}, \ u \in [0,1].$$

This means that $S_I \sim \mathsf{Geom}(P_I)$ with $P_I = \left(1 + \sum_{r\in I}\frac{1}{r!}\right)^{-1}$. In particular,

$$\sum_{r=1}^{\infty} G_r = S_{\mathbb{N}} \sim \mathsf{Geom}\left(\mathrm{e}^{-1}\right).$$

Remark 2. The limiting random sequence $(G_r, r \in \mathbb{N})$ allows for an interpretation in terms of a balls-into-bins model. Consider an infinite set of bins of unlimited capacity, numbered $0, 1, 2, \ldots$ Each time a ball is placed into one of the bins, choosing the rth one with probability

$$p_r = \frac{1}{\mathrm{e}r!}, \quad r \in \mathbb{N}_0, \tag{13}$$

independently of the previous choices. In other words, the indices of the successively selected bins form an i.i.d. sequence of $\mathsf{Pois}(1)$-distributed random variables. This process continues until a ball is placed into 0th bin. Then G_r, $r \in \mathbb{N}$, describes the number of balls in the rth bin.

Indeed, taking (13) into account, we may expand the right-hand side of (11) in a geometric series:

$$\begin{aligned}\mathbb{E}\left(\prod_{r=1}^{\infty} u_r^{G_r}\right) &= \sum_{k=0}^{\infty} p_0\left(\sum_{r=1}^{\infty} p_r u_r\right)^k \\ &= \sum\left(\frac{(k_1+k_2+\cdots)!}{k_1!\cdot k_2!\cdot\ldots} p_0 \prod_{r=1}^{\infty} p_r^{k_r}\right)\left(\prod_{r=1}^{\infty} u_r^{k_r}\right),\end{aligned} \tag{14}$$

where the sum on the right-hand side is over the set of all infinite sequences $(k_r, r \in \mathbb{N})$ of non-negative integers, only a finite number of which are non-zero. It is easy to see that the probability generating function on the right-hand side of (14) corresponds to the above balls-into-bins model.

Note that finite-dimensional counterparts of infinite-dimensional geometric (and, more generally, negative binomial) distributions like that of $(G_r, r \in \mathbb{N})$ were introduced and studied in [7] and [21]; see also a survey of various multivariate geometric and negative binomial distributions in [10].

As our final result, we give an infinite-dimensional extension of (1), the limit theorem by Erdős and Rényi.

Theorem 3. *Let B_r, $r \in \mathbb{N}_0$, be independent Gumbel-distributed random variables with distribution functions*

$$\mathbb{P}\{B_r < x\} = \exp\left(-\frac{\mathrm{e}^{-x}}{r!}\right), \quad x \in \mathbb{R}. \tag{15}$$

Then

$$\left(\frac{T_r^{(n)}}{n} - \ln n - r\ln\ln n, r \in \mathbb{N}_0\right) \xrightarrow{d} (B_r, r \in \mathbb{N}_0) \quad \text{in } \mathbb{R}^{\infty} \text{ as } n \to \infty. \tag{16}$$

In particular, this result allows obtaining limit distributions for times between completions of different collections. Denote by $\Delta^{(n)}_{r_1,r_2} = T^{(n)}_{r_2} - T^{(n)}_{r_1}$, $r_2 > r_1$, such an inter-completion time.

Corollary 1. *Let L_{r_1,r_2} be a logistic random variable with distribution function*

$$\mathbb{P}\{L_{r_1,r_2} < x\} = \frac{r_2!}{r_2! + r_1!\,\mathrm{e}^{-x}}, \quad x \in \mathbb{R}.$$

Then

$$\frac{\Delta^{(n)}_{r_1,r_2}}{n} - (r_2 - r_1)\ln\ln n \xrightarrow{d} L_{r_1,r_2} \quad \text{as } n \to \infty. \tag{17}$$

It is well known (and may be easily checked by means of characteristic functions) that the difference of two independent Gumbel-distributed random variables has a logistic distribution. Particularly, $B_{r_2} - B_{r_1} \stackrel{d}{=} L_{r_1,r_2}$. Thus, (17) follows from (16).

3 Convergence of associated point processes I. Thinned processes

One of the key points in the proof of Theorem 1 is convergence of specially constructed point processes to a Poisson one. We now proceed to the corresponding construction.

For fixed $n \in \mathbb{N}$ and $r \in \mathbb{N}_0$, let

$$\psi^{(n)}(x) = \frac{x}{n} - \ln n, \quad x \in \mathbb{R}, \tag{18}$$

$$\eta^{(n)}_r = \sum_{i=1}^{n} \delta_{\psi^{(n)}\left(Z^{(n)}_{i,r}\right)}, \tag{19}$$

where δ_u stands for the Dirac measure $\mathbb{1}\{u \in \cdot\}$. The processes $\eta^{(n)}_r$ describe $(r+1)$th arrivals of different coupon types in the poissonized scheme. Note that in a similar case in [19], in order to provide the necessary convergence, we used an r-dependent centering/normalizing function

$$\psi^{(n)}_r(x) = \frac{x}{n} - \ln n - r\ln\ln n, \quad x \in \mathbb{R}. \tag{20}$$

In our case, however, we want to consider the numbers of empty spots in different albums at the same point in time, namely when the 0th album is completed. Hence, we have to deal with the same centering for different r. So, we will achieve the desired convergence in a different way — by means of thinnings.

For a proper point process η, let us denote by $T_p\eta$, $p \in [0,1]$, its p-thinning, i.e. the point process which independently keeps the points of η with probability p and removes otherwise (see, e.g., [24], Section 5.3 for details). On the space

$$\mathbb{X} = \mathbb{N}_0 \times \left(\mathbb{R} \cup \{+\infty\}\right), \tag{21}$$

endowed with some relevant metric, say,

$$d\big((r_1,x),(r_2,y)\big) = |r_2 - r_1| + |\mathrm{e}^{-y} - \mathrm{e}^{-x}|, \tag{22}$$

consider the point processes $H^{(n)}$, $n \geq 3$, given as follows: for $B_r \in \mathfrak{B}\big(\mathbb{R} \cup \{+\infty\}\big)$, $r \in \mathbb{N}_0$, define

$$H^{(n)}\Big(\bigcup_{r=0}^{\infty}\big(\{r\} \times B_r\big)\Big) = \sum_{r=0}^{\infty}\big(T_{\ln^{-r} n}\eta_r^{(n)}\big)(B_r), \tag{23}$$

where all thinnings are performed independently across levels and of the underlying coupon process. (We require $n \geq 3$ in order to have $\ln^{-r} n \in [0,1]$ for all $r \in \mathbb{N}_0$.) In other words, we are thinning out the processes $\eta_r^{(n)}$ and glue them into one "multilevel" point process $H^{(n)}$. Note that, by construction in (19),

$$\big(T_{\ln^{-r} n}\eta_r^{(n)}\big)\big(\{+\infty\}\big) = \eta_r^{(n)}\big(\{+\infty\}\big) = 0 \tag{24}$$

anyway. The reason we, nevertheless, consider the semi-compactified real axis $\mathbb{R} \cup \{+\infty\}$ instead of just $\mathbb{R}$ will become clear from what follows (see Section 5).

Before stating the main theorem of this section, we reformulate the basic definitions related to convergence of point processes as applied to our problem (for a detailed exposition in the abstract setting, see [28, 29], or [22]). Let $M_p(\mathbb{X})$ denote the space of all locally finite (with respect to d in (22)) point measures on $\mathbb{X}$ given by (21). For $\mu, \mu_1, \mu_2, \ldots \in M_p(\mathbb{X})$, μ_n are said to converge vaguely to μ (denoted by $\mu_n \xrightarrow{v} \mu$) if $\int_{\mathbb{X}} f\,\mathrm{d}\mu_n \to \int_{\mathbb{X}} f\,\mathrm{d}\mu$ for each continuous compactly supported non-negative test function f defined on $\mathbb{X}$. In our case, this means that

$$\int_{\mathbb{R}\cup\{+\infty\}} g(t)\,\mu_n\big(\{r\} \times \mathrm{d}t\big) \to \int_{\mathbb{R}\cup\{+\infty\}} g(t)\,\mu\big(\{r\} \times \mathrm{d}t\big)$$

for each $r \in \mathbb{N}_0$ and each continuous compactly supported $g\colon \mathbb{R} \cup \{+\infty\} \to [0, +\infty)$. As usual, the set $M_p(\mathbb{X})$, equipped with the topology of the above convergence, can be metrized as a complete separable metric space. This setting allows us to consider the distributional convergence of point processes $H^{(3)}, H^{(4)} \ldots$, denoted as $H^{(n)} \xrightarrow{vd} H$. The main result of this section, Theorem 4 below, asserts that the point processes $H^{(n)}$ converge in this sense toward a non-homogeneous Poisson process.

Theorem 4. *Let H be a Poisson point process on $\mathbb{X}$ with intensity measure λ given by*

$$\lambda\Big(\bigcup_{r=0}^{\infty}\big(\{r\} \times B_r\big)\Big) = \sum_{r=0}^{\infty}\frac{1}{r!}\int_{B_r} \mathrm{e}^{-x}\,\mathrm{d}x, \quad B_r \in \mathfrak{B}\big(\mathbb{R} \cup \{+\infty\}\big). \tag{25}$$

Then $H^{(n)} \xrightarrow{vd} H$ as $n \to \infty$.

Remark 3. The superposition theorem for Poisson processes (see, e.g., Theorem 3.3 in [24]) implies that different levels $H\big(\{r\} \times \cdot\big)$, $r \in \mathbb{N}_0$, of the limiting process H are independent. In other words, the point processes $T_{\ln^{-r} n}\eta_r^{(n)}$ are asymptotically independent. This fact is rather surprising, since without thinning no asymptotic independence would have been expected. Indeed, by (18) and (19), $\eta_r^{(n)}$ increase in r: $\psi^{(n)}\big(Z_{i,r_1}^{(n)}\big) < \psi^{(n)}\big(Z_{i,r_2}^{(n)}\big)$ for $r_1 < r_2$ and fixed n, i, which excludes any independence.

Remark 4. The levels $H\big(\{r\}\times\cdot\big)$ allow for a simple interpretation (see Remark 3.2 in [19]). Let ζ be a stationary unit-rate Poisson point process restricted to $(0,+\infty)$, and put

$$h(x) = -\ln r! - \ln x, \quad x > 0.$$

Then, $H\big(\{r\}\times\cdot\big) \stackrel{d}{=} \sum_{x\in\operatorname{supp}\zeta}\delta_{h(x)}$.

Proof of Theorem 4. Let $\mathcal{U}$ denote the ring of all Borel subsets in $\mathbb{X}$, bounded with respect to d in (22). So,

$$\mathcal{U} = \Big\{\bigcup_{r=0}^{s}\big(\{r\}\times B_r\big), s\in\mathbb{N}_0, \\ B_r\in\mathfrak{B}\big(\mathbb{R}\cup\{+\infty\}\big) \text{ and are bounded from below}\Big\}. \tag{26}$$

Since the measure λ given by (25) is diffuse, H is a simple Poisson point process (see, e.g., Proposition 6.9 in [24]). Thus, by Theorem 4.18 in [22], it suffices only to prove that, for each $U\in\mathcal{U}$,

$$\lim_{n\to\infty}\mathbb{P}\{H^{(n)}(U)=0\} = \mathbb{P}\{H(U)=0\}, \tag{27}$$

$$\lim_{n\to\infty}\mathbb{E}H^{(n)}(U) = \mathbb{E}H(U). \tag{28}$$

Note that, by (24) and (25), $H^{(n)}\big(\mathbb{N}_0\times\{+\infty\}\big) = H\big(\mathbb{N}_0\times\{+\infty\}\big) = 0$ a.s. So, we may assume that $U\subset\mathbb{N}_0\times\mathbb{R}$.

First we take up the proof of (27). According to (26) and (23),

$$\begin{aligned}&\mathbb{P}\{H^{(n)}(U)=0\}\\ =&\mathbb{P}\Big\{H^{(n)}\Big(\bigcup_{r=0}^{s}\big(\{r\}\times B_r\big)\Big)=0\Big\} = \mathbb{P}\Big\{\sum_{r=0}^{s}\big(T_{\ln^{-r}n}\eta_r^{(n)}\big)(B_r)=0\Big\}\\ =&\mathbb{P}\Big\{\big(T_{\ln^{-r}n}\eta_r^{(n)}\big)(B_r)=0 \text{ for } r=0,\dots,s\Big\}.\end{aligned} \tag{29}$$

For $0\le r_1<\dots<r_m\le s$, denote

$$P^{(n)}_{r_1,\dots,r_m} = \mathbb{P}\big\{\psi^{(n)}\big(Z^{(n)}_{i,r_1}\big)\in B_{r_1},\dots,\psi^{(n)}\big(Z^{(n)}_{i,r_m}\big)\in B_{r_m}\big\}, \tag{30}$$

which does not depend on i by virtue of the i.i.d. property of $Z^{(n)}_{i,r}$. By this property again, (29) and inclusion-exclusion yield

$$\begin{aligned}\mathbb{P}\{H^{(n)}(U)=0\} = \Big(1 &- \sum_{0\le r_1\le s}\ln^{-r_1}n\cdot P^{(n)}_{r_1}\\ &+\sum_{0\le r_1<r_2\le s}\ln^{-r_1-r_2}n\cdot P^{(n)}_{r_1,r_2} - \dots + (-1)^{s+1}\ln^{-1-\dots-s}n\cdot P^{(n)}_{0,1,\dots,s}\Big)^n.\end{aligned}$$

Hence,

$$\lim_{n\to\infty}\ln\mathbb{P}\{H^{(n)}(U)=0\} = -\sum_{0\le r_1\le s}\lim_{n\to\infty}\big(n\ln^{-r_1}n\cdot P^{(n)}_{r_1}\big)$$

$$+\sum_{0\le r_1<r_2\le s}\lim_{n\to\infty}\left(n\ln^{-r_1-r_2}n\cdot P^{(n)}_{r_1,r_2}\right)-\cdots$$
$$+(-1)^{s+1}\lim_{n\to\infty}\left(n\ln^{-1-\cdots-s}n\cdot P^{(n)}_{0,1,\dots,s}\right).$$

We now prove that the limits in the first sum equal $\frac{1}{r_1!}\int_{B_{r_1}}\mathrm{e}^{-x}\,\mathrm{d}x$, while those in the second sum vanish. Since all subsequent terms are dominated by the latter ones, this combined with (25) will prove (27).

As $Z^{(n)}_{i,0}, Z^{(n)}_{i,1}-Z^{(n)}_{i,0},\dots,Z^{(n)}_{i,s}-Z^{(n)}_{i,s-1}$ are independent $\mathsf{Exp}\left(\frac{1}{n}\right)$, the densities $f^{(n)}_{r_1}$ and $f^{(n)}_{r_1,r_2}$ of $\psi^{(n)}\left(Z^{(n)}_{i,r_1}\right)$ and $\left(\psi^{(n)}\left(Z^{(n)}_{i,r_1}\right),\psi^{(n)}\left(Z^{(n)}_{i,r_2}\right)\right)$, respectively, can be easily calculated:

$$f^{(n)}_{r_1}(x)=\frac{(x+\ln n)^{r_1}\mathrm{e}^{-x}}{n\,r_1!}\cdot\mathbb{1}\{-\ln n\le x\},\quad x\in\mathbb{R},\tag{31}$$

$$f^{(n)}_{r_1,r_2}(x,y)=\frac{(x+\ln n)^{r_1}(y-x)^{r_2-r_1-1}\mathrm{e}^{-y}}{n\,r_1!\,(r_2-r_1-1)!}\cdot\mathbb{1}\{-\ln n\le x\le y\},\quad x,y\in\mathbb{R}.\tag{32}$$

Recall that B_{r_1} and B_{r_2} were assumed to be bounded from below. Hence, by (30),

$$n\ln^{-r_1}n\cdot P^{(n)}_{r_1}=n\ln^{-r_1}n\cdot\int_{B_{r_1}}f^{(n)}_{r_1}(x)\,\mathrm{d}x$$
$$=\frac{1}{r_1!}\int_{B_{r_1}}\left(1+\frac{x}{\ln n}\right)^{r_1}\mathrm{e}^{-x}\cdot\mathbb{1}\{-\ln n\le x\}\,\mathrm{d}x\to\frac{1}{r_1!}\int_{B_{r_1}}\mathrm{e}^{-x}\,\mathrm{d}x,\quad n\to\infty\tag{33}$$

due to dominated convergence. Next,

$$n\ln^{-r_1-r_2}n\cdot P^{(n)}_{r_1,r_2}=n\ln^{-r_1-r_2}n\cdot\iint_{B_{r_1}\times B_{r_2}}f^{(n)}_{r_1,r_2}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$
$$=\frac{\ln^{-r_2}n}{r_1!(r_2-r_1-1)!}\iint_{B_{r_1}\times B_{r_2}}\left(1+\frac{x}{\ln n}\right)^{r_1}(y-x)^{r_2-r_1-1}\mathrm{e}^{-y}\mathbb{1}\{-\ln n\le x\le y\}\,\mathrm{d}x\,\mathrm{d}y,$$

which vanishes as $n\to\infty$ again by dominated convergence. This completes the proof of (27).

We now turn to the proof of (28). Similarly to (29), we have

$$\mathbb{E}H^{(n)}(U)=\mathbb{E}H^{(n)}\left(\bigcup_{r=0}^{s}\left(\{r\}\times B_r\right)\right)$$
$$=\sum_{r=0}^{s}\mathbb{E}\left(T_{\ln^{-r}n}\eta^{(n)}_r\right)(B_r)=\sum_{r=0}^{s}\ln^{-r}n\cdot\mathbb{E}\eta^{(n)}_r(B_r).$$

It follows from (19) and the i.i.d. property of $Z^{(n)}_{i,r}$ that $\eta^{(n)}_r(B_r)\sim\mathsf{Bin}\left(n,P^{(n)}_r\right)$ with $P^{(n)}_r$ given by (30). So, by (33),

$$\mathbb{E}H^{(n)}(U)=\sum_{r=0}^{s}n\ln^{-r}n\cdot P^{(n)}_r\to\sum_{r=0}^{s}\frac{1}{r!}\int_{B_r}\mathrm{e}^{-x}\,\mathrm{d}x\quad\text{as } n\to\infty,$$

which equals $\mathbb{E}H(U)$ due to (25). This concludes the proof of (28) and thus of Theorem 4. □

4 Thinning trick

In this section, we consider a somewhat unusual approach to proving limit theorems, which is the basic tool in the proof of Theorem 1. This approach does not seem to have been used in this context before, and may prove useful elsewhere as well.

For an $\mathbb{N}_0$-valued random variable X, define its p-thinning, $p \in [0,1]$, by

$$p \odot X = \sum_{i=1}^{X} \varepsilon_i, \tag{34}$$

where ε_i are $\mathsf{Bin}(1,p)$-distributed and independent of each other and of X. Note that, in the notation of the previous section, $(T_p\eta)(B) \stackrel{d}{=} p \odot \eta(B)$ for a proper point process η and a Borel set B. This operation, going back to [27], was then used in [32] to introduce the concepts of discrete self-decomposability and stability.

The idea behind the proposed technique is to replace the normalization by thinning. In other words, instead of proving that $p^{(n)}X^{(n)} \stackrel{d}{\to} Y$ with some constants $p^{(n)} \to 0$ and a limiting random variable Y, we will prove that $p^{(n)} \odot X^{(n)} \stackrel{d}{\to} Z$ with some new limiting random variable Z. The connection between the distributions of Y and Z can be guessed from the following examples. For $X^{(n)} = n$ a.s., we have $\frac{1}{n}X^{(n)} = 1 \stackrel{d}{\to} 1$ and $\frac{1}{n} \odot X^{(n)} \stackrel{d}{\to} Z \sim \mathsf{Pois}(1)$ by the Poisson limit theorem. More generally, consider an i.i.d. sequence $(\xi_i, i \in \mathbb{N})$ with $\mathbb{E}\xi_i = \lambda$, and denote $X^{(n)} = \sum_{i=1}^{n} \xi_i$. Due to the law of large numbers, $\frac{1}{n}X^{(n)} \stackrel{d}{\to} \lambda$, and, by Theorem 3.1 in [17], $\frac{1}{n} \odot X^{(n)} \stackrel{d}{\to} Z \sim \mathsf{Pois}(\lambda)$. All this suggests that, in general, Z must have the mixed Poisson distribution with mixing distribution of Y (see, e.g., Chapter 2 in [16]). This means that

$$\mathbb{P}\{Z = k\} = \int_{[0,+\infty)} \frac{\mathrm{e}^{-y}y^k}{k!} F_Y(\mathrm{d}y), \quad k \in \mathbb{N}_0, \tag{35}$$

where F_Y stands for the distribution function of Y.

The main result of this section, Theorem 5 below, justifies this approach. With an eye to the future, we will state and prove it in multidimensional form. Let $\mathbf{Y} = (Y_1, \dots, Y_s)$ be a random vector with a.s. non-negative components. By analogy with (35), the random vector $\mathbf{Z} = (Z_1, \dots, Z_s)$ is said to have a multivariate mixed Poisson distribution with mixing distribution of $\mathbf{Y}$ (see [13] or [23]) if

$$\begin{aligned}&\mathbb{P}\{Z_1 = k_1, \dots, Z_s = k_s\}\\ &= \int \cdots \int_{[0,+\infty)^s} \frac{\mathrm{e}^{-y_1}y_1^{k_1}}{k_1!} \cdot \ldots \cdot \frac{\mathrm{e}^{-y_s}y_s^{k_s}}{k_s!} F_{\mathbf{Y}}(\mathrm{d}y_1, \dots, \mathrm{d}y_s),\ k_1, \dots, k_s \in \mathbb{N}_0. \end{aligned} \tag{36}$$

Remark 5. In what follows, we will need another equivalent interpretation of the multivariate mixed Poisson distribution. Namely, we may define $\mathbf{Z}$ by $Z_r = N_r(Y_r)$, $r = 1, \dots, s$, where N_r stand for unit-rate Poisson counting processes, independent of each other and of $\mathbf{Y}$.

Theorem 5. *Let* $\mathbf{X}^{(n)} = \left(X_1^{(n)}, \dots, X_s^{(n)}\right)$, $n \in \mathbb{N}$, *be a sequence of random vectors with* $\mathbb{N}_0$*-valued components, and* $\left(p_1^{(n)}, \dots, p_s^{(n)}\right)$, $n \in \mathbb{N}$, *a non-random sequence*

with $p_r^{(n)} \in [0,1]$ *and* $\lim_{n\to\infty} p_r^{(n)} = 0$, $r = 1,\dots,s$. *Assume that*

$$\left(p_1^{(n)} \odot X_1^{(n)}, \dots, p_s^{(n)} \odot X_s^{(n)}\right) \xrightarrow{d} \left(Z_1, \dots, Z_s\right) \quad \text{as } n \to \infty, \tag{37}$$

and the limiting random vector on the right-hand side has a multivariate mixed Poisson distribution (36). *Suppose additionally that*

$$\sup_{n\in\mathbb{N}} p_r^{(n)} \mathbb{E} X_r^{(n)} < \infty, \quad r = 1,\dots,s. \tag{38}$$

Then

$$\left(p_1^{(n)} X_1^{(n)}, \dots, p_s^{(n)} X_s^{(n)}\right) \xrightarrow{d} \left(Y_1, \dots, Y_s\right) \quad \text{as } n \to \infty.$$

Proof. Denote by $\mathcal{G}_n$ the probability generating function of $\mathbf{X}^{(n)}$:

$$\mathcal{G}_n(u_1,\dots,u_s) = \mathbb{E}\left(u_1^{X_1^{(n)}} \cdot \ldots \cdot u_s^{X_s^{(n)}}\right), \tag{39}$$

which is well defined at least for $(u_1,\dots,u_s) \in [0,1]^s$. Then, by (34), the probability generating function of $\left(p_1^{(n)} \odot X_1^{(n)}, \dots, p_s^{(n)} \odot X_s^{(n)}\right)$ takes the form

$$\begin{aligned}
&\mathbb{E}\left(u_1^{p_1^{(n)}\odot X_1^{(n)}} \cdot \ldots \cdot u_s^{p_s^{(n)}\odot X_s^{(n)}}\right) = \mathbb{E}\left(u_1^{\sum_{i=1}^{X_1^{(n)}} \varepsilon_{1,i}^{(n)}} \cdot \ldots \cdot u_s^{\sum_{i=1}^{X_s^{(n)}} \varepsilon_{s,i}^{(n)}}\right) \\
&= \mathbb{E} \sum_{k_1,\dots,k_s=0}^{\infty} u_1^{\sum_{i=1}^{k_1} \varepsilon_{1,i}^{(n)}} \cdot \ldots \cdot u_s^{\sum_{i=1}^{k_s} \varepsilon_{s,i}^{(n)}} \cdot \mathbb{1}\left\{X_1^{(n)} = k_1, \dots, X_s^{(n)} = k_s\right\} \\
&= \sum_{k_1,\dots,k_s=0}^{\infty} \left(\mathbb{E} u_1^{\varepsilon_{1,1}^{(n)}}\right)^{k_1} \cdot \ldots \cdot \left(\mathbb{E} u_s^{\varepsilon_{s,1}^{(n)}}\right)^{k_s} \cdot \mathbb{P}\left\{X_1^{(n)} = k_1, \dots, X_s^{(n)} = k_s\right\} \\
&= \mathbb{E}\left(\left(p_1^{(n)} u_1 + q_1^{(n)}\right)^{X_1^{(n)}} \cdot \ldots \cdot \left(p_s^{(n)} u_s + q_s^{(n)}\right)^{X_s^{(n)}}\right) \\
&= \mathcal{G}_n\left(p_1^{(n)} u_1 + q_1^{(n)}, \dots, p_s^{(n)} u_s + q_s^{(n)}\right).
\end{aligned} \tag{40}$$

Here $\varepsilon_{r,i}^{(n)}$, $r = 1,\dots,s$, $i \in \mathbb{N}$, are independent $\mathsf{Bin}(1, p_r^{(n)})$, and $q_r^{(n)} = 1 - p_r^{(n)}$.

Next, denote by $\mathcal{L}$ the Laplace transform of $\mathbf{Y} = (Y_1,\dots,Y_s)$:

$$\mathcal{L}(t_1,\dots,t_s) = \mathbb{E} \mathrm{e}^{-t_1 Y_1 - \dots - t_s Y_s}, \quad t_1,\dots,t_s \geq 0. \tag{41}$$

Then the probability generating function of $\mathbf{Z} = (Z_1,\dots,Z_s)$ becomes

$$\begin{aligned}
\mathbb{E}\left(u_1^{Z_1} \cdot \ldots \cdot u_s^{Z_s}\right) &= \mathbb{E}\,\mathbb{E}_{\mathbf{Y}}\left(u_1^{Z_1} \cdot \ldots \cdot u_s^{Z_s}\right) = \mathbb{E}\left(\left(\mathbb{E}_{Y_1} u_1^{Z_1}\right) \cdot \ldots \cdot \left(\mathbb{E}_{Y_s} u_s^{Z_s}\right)\right) \\
&= \mathbb{E}\left(\mathrm{e}^{-Y_1(1-u_1)} \cdot \ldots \cdot \mathrm{e}^{-Y_s(1-u_s)}\right) = \mathcal{L}(1-u_1,\dots,1-u_s),
\end{aligned} \tag{42}$$

where $\mathbb{E}_{\mathbf{Y}}, \mathbb{E}_{Y_1}, \dots, \mathbb{E}_{Y_s}$ stand for conditional means with respect to $\mathbf{Y}, Y_1, \dots, Y_s$, respectively. So, by (37), (40), and (42),

$$\lim_{n\to\infty} \mathcal{G}_n\left(p_1^{(n)} u_1 + q_1^{(n)}, \dots, p_s^{(n)} u_s + q_s^{(n)}\right) = \mathcal{L}(1-u_1,\dots,1-u_s), \; u_1,\dots,u_s \in [0,1],$$

and thus

$$\lim_{n\to\infty} \mathcal{G}_n\big(1-p_1^{(n)}t_1,\dots,1-p_s^{(n)}t_s\big) = \mathcal{L}(t_1,\dots,t_s), \quad t_1,\dots,t_s\in[0,1]. \tag{43}$$

It follows from the multivariate mean value theorem that

$$\begin{aligned}\Big|\mathcal{G}_n\big(1-p_1^{(n)}t_1,\dots,1-p_s^{(n)}t_s\big) - \mathcal{G}_n\big(\mathrm{e}^{-p_1^{(n)}t_1},\dots,\mathrm{e}^{-p_s^{(n)}t_s}\big)\Big| \\ \le \sum_{r=1}^{s} \sup_{u_1,\dots,u_s\in[0,1]} \frac{\partial\mathcal{G}_n}{\partial u_r}(u_1,\dots,u_s)\Big(\mathrm{e}^{-p_r^{(n)}t_r} - 1 + p_r^{(n)}t_r\Big).\end{aligned} \tag{44}$$

Since $\mathrm{e}^{-p_r^{(n)}t_r} - 1 + p_r^{(n)}t_r = \mathcal{O}\big(p_r^{(n)}\big)$ as $n\to\infty$, and

$$p_r^{(n)} \sup_{u_1,\dots,u_s\in[0,1]} \frac{\partial\mathcal{G}_n}{\partial u_r}(u_1,\dots,u_s) = p_r^{(n)}\frac{\partial\mathcal{G}_n}{\partial u_r}(1,\dots,1) = p_r^{(n)}\mathbb{E}X_r^{(n)} = \mathcal{O}(1)$$

by (38), then the right-hand side of (44) is $\mathcal{O}(1)$ as $n\to\infty$. Thus, (43) implies

$$\lim_{n\to\infty} \mathcal{G}_n\big(\mathrm{e}^{-p_1^{(n)}t_1},\dots,\mathrm{e}^{-p_s^{(n)}t_s}\big) = \mathcal{L}(t_1,\dots,t_s), \quad t_1,\dots,t_s\in[0,1].$$

Due to (39) and (41), the left-hand and right-hand sides of the preceding relation are Laplace transforms of $\big(p_1^{(n)}X_1^{(n)},\dots,p_s^{(n)}X_s^{(n)}\big)$ and $\big(Y_1,\dots,Y_s\big)$, respectively. So, the claim follows from Theorem 2 in [33]. □

5 Proof of Theorem 1

In this short section, we apply Theorems 4 and 5 to the proof of Theorem 1. Let

$$\begin{aligned}\tilde{T}_0^{(n)} &= \inf\big\{x\in\mathbb{R}\colon H^{(n)}\big(\{0\}\times(x,+\infty)\big)=0\big\},\\ V_r^{(n)} &= H^{(n)}\big(\{r\}\times\big(\tilde{T}_0^{(n)},+\infty\big)\big)\end{aligned} \tag{45}$$

for $r\in\mathbb{N}$. In the notation of Section 2, $\tilde{T}_0^{(n)} = \max_{i\le n}\psi^{(n)}\big(Z_{i,0}^{(n)}\big)$, and thus means the (centered and normalized) time when the main collector completes his album in the poissonized scheme. Next, $V_r^{(n)}$ is the number of points of the thinned process $T_{\ln^{-r}n}\eta_r^{(n)}$ to the right of $\tilde{T}_0^{(n)}$. So, according to (7), it is easily seen to be equal in distribution to $\ln^{-r}n\odot U_r^{(n)}$, where the operation $\odot$ is defined in (34).

By Theorem 4 and the Skorokhod coupling (see, e.g., [29], p. 41), we may assume that $H^{(n)}\xrightarrow{v}H$ a.s. as $n\to\infty$. Define $\tilde{T}_0$ and V_r similarly to (45) but with H instead of $H^{(n)}$. Proposition 3.13 in [28], together with the simplicity of H and the independence of its levels, implies that the mappings corresponding to $\tilde{T}_0$ and V_r are continuous at H a.s. Hence, $\lim_{n\to\infty}\tilde{T}_0^{(n)} = \tilde{T}_0$ and $\lim_{n\to\infty}V_r^{(n)} = V_r$ a.s. It is here that we use the fact that $H^{(n)}(\{r\}\times\cdot)$ are given on the semi-compactified space $\mathbb{R}\cup\{+\infty\}$, because we need $(x,+\infty)$ to be relatively compact.

Let $(N_r(t), t\ge 0)$, $r\in\mathbb{N}_0$, be independent unit-rate Poisson counting processes. Denote by E the first jump time of N_0, and note that $E\sim\mathsf{Exp}(1)$. Due to Remarks

3 and 4, $\tilde{T}_0 \stackrel{d}{=} -\ln E$, and $(V_1, \ldots, V_s) \stackrel{d}{=} (N_r(E/r!), r = 1, \ldots, s)$ for any $s \in \mathbb{N}$. Summarizing all the above, we have

$$\left(\ln^{-r} n \odot U_r^{(n)},\ r = 1, \ldots, s\right) \xrightarrow{d} (N_r(E/r!),\ r = 1, \ldots, s), \qquad n \to \infty.$$

Then, applying Theorem 5 and taking into account Remark 5, we obtain

$$\left(\ln^{-r} n \cdot U_r^{(n)},\ r = 1, \ldots, s\right) \xrightarrow{d} (E/r!,\ r = 1, \ldots, s), \qquad n \to \infty,$$

condition (38) being satisfied due to (2) and (3). Thus, the convergence in (8) holds in the sense of finite-dimensional distributions. To complete the proof, it only remains to note that in $\mathbb{R}^\infty$ the notions of finite-dimensional convergence and convergence in distribution are equivalent (see, e.g., [29], pp. 53–54). □

6 Convergence of associated point processes II. An r-dependent centering

For the proofs of Theorems 2 and 3, we will also use a specially constructed sequence of point processes. This time, however, their convergence will be achieved not at the expense of thinning, as in Section 3, but due to an r-dependent centering.

On the metric space $(\mathbb{X}, d)$ given by (21) and (22), consider the point processes $\hat{H}^{(n)}$, $n \in \mathbb{N}$, defined as follows: for $B_r \in \mathfrak{B}\big(\mathbb{R} \cup \{+\infty\}\big)$, $r \in \mathbb{N}_0$, let

$$\hat{H}^{(n)}\left(\bigcup_{r=0}^{\infty}\big(\{r\} \times B_r\big)\right) = \sum_{r=0}^{\infty} \hat{\eta}_r^{(n)}(B_r), \tag{46}$$

where

$$\hat{\eta}_r^{(n)} = \sum_{i=1}^{n} \delta_{\hat{\psi}_r^{(n)}\left(Z_{i,r}^{(n)}\right)},$$

and the centering/normalizing functions $\hat{\psi}_r^{(n)}$, slightly generalizing $\psi_r^{(n)}$ in (20), are defined as follows:

$$\hat{\psi}_r^{(n)}(x) = \frac{x}{n} - \ln n - r \ln\ln n - \varkappa_r^{(n)}, \quad x \in \mathbb{R}. \tag{47}$$

Here, for each $r \in \mathbb{N}_0$, $\varkappa_r^{(n)}$ is a fixed numerical sequence such that $\lim_{n\to\infty} \varkappa_r^{(n)} = 0$. We will also use counterparts of $\hat{H}^{(n)}$ in the original, non-poissonized scheme:

$$\hat{\Xi}^{(n)}\left(\bigcup_{r=0}^{\infty}\big(\{r\} \times B_r\big)\right) = \sum_{r=0}^{\infty} \hat{\xi}_r^{(n)}(B_r),$$

where

$$\hat{\xi}_r^{(n)} = \sum_{i=1}^{n} \delta_{\hat{\psi}_r^{(n)}\left(Y_{i,r}^{(n)}\right)},$$

and $Y_{i,r}^{(n)}$ are defined at the beginning of Section 2.

The following result is an analogue of Theorem 4 but differs in two significant aspects. Firstly, we now use the r-dependent centering (47) instead of thinning, and secondly, the original scheme instead of the poissonized one. The latter will require a special depoissonization procedure based on the coupling formula (6) and similar to that used in [19]. Unlike Theorem 4, we need to consider this result in the original setting rather than in the poissonized one due to the lack of any counterpart to (7) for $\hat{U}_r^{(n)}$.

Theorem 6. *Let H be defined as in Theorem 4. Then $\hat{\Xi}^{(n)} \xrightarrow{vd} H$ as $n \to \infty$.*

Remark 6. This theorem may be regarded as an infinite-dimensional extension of Theorem 3.1 in [19].

Proof of Theorem 6. To begin with, we will return for a while to the poissonized scheme and prove that $\hat{H}^{(n)} \xrightarrow{vd} H$ as $n \to \infty$. Similarly to (27) and (28), it is enough only to show that

$$\lim_{n\to\infty} \mathbb{P}\{\hat{H}^{(n)}(U) = 0\} = \mathbb{P}\{H(U) = 0\}, \tag{48}$$

$$\lim_{n\to\infty} \mathbb{E}\hat{H}^{(n)}(U) = \mathbb{E}H(U), \tag{49}$$

for each $U \in \mathcal{U}$, where the ring $\mathcal{U}$ is defined in (26). Moreover, as before we may assume that $U \subset \mathbb{N}_0 \times \mathbb{R}$. Then, by (46)

$$\begin{aligned}\mathbb{P}\{\hat{H}^{(n)}(U) = 0\} &= \mathbb{P}\Big\{\hat{H}^{(n)}\Big(\bigcup_{r=0}^{s}\big(\{r\} \times B_r\big)\Big) = 0\Big\} \\ &= \mathbb{P}\Big\{\sum_{r=0}^{s} \hat{\eta}_r^{(n)}(B_r) = 0\Big\} = \mathbb{P}\big\{\hat{\eta}_r^{(n)}(B_r) = 0 \text{ for } r = 0, \dots, s\big\}. \end{aligned} \tag{50}$$

Hence, using the i.i.d. property of $Z_{i,r}^{(n)}$ and inclusion-exclusion, we have

$$\mathbb{P}\{\hat{H}^{(n)}(U) = 0\} = \Big(1 - \sum_{0\le r_1\le s} \hat{P}_{r_1}^{(n)} + \sum_{0\le r_1<r_2\le s} \hat{P}_{r_1,r_2}^{(n)} - \dots + (-1)^{s+1}\hat{P}_{0,1,\dots,s}^{(n)}\Big)^n,$$

where the probabilities $\hat{P}_{r_1,\dots,r_m}^{(n)}$ are defined similarly to $P_{r_1,\dots,r_m}^{(n)}$ in (30), but with $\hat{\psi}_r^{(n)}$ instead of $\psi^{(n)}$. So,

$$\begin{aligned}&\lim_{n\to\infty} \ln \mathbb{P}\{\hat{H}^{(n)}(U) = 0\} \\ &\quad = -\sum_{0\le r_1\le s} \lim_{n\to\infty} n\hat{P}_{r_1}^{(n)} + \sum_{0\le r_1<r_2\le s} \lim_{n\to\infty} n\hat{P}_{r_1,r_2}^{(n)} - \dots + (-1)^{s+1} \lim_{n\to\infty} n\hat{P}_{0,1,\dots,s}^{(n)}.\end{aligned}$$

As before, in order to prove (48), it suffices to show that the limits in the first sum equal $\frac{1}{r_1!}\int_{B_{r_1}} \mathrm{e}^{-x}\,\mathrm{d}x$, and those in the second sum vanish.

Using (31) and (32), we may easily calculate the densities $\hat{f}_{r_1}^{(n)}$ and $\hat{f}_{r_1,r_2}^{(n)}$ of $\hat{\psi}_{r_1}^{(n)}\big(Z_{i,r_1}^{(n)}\big)$ and $\big(\hat{\psi}_{r_1}^{(n)}\big(Z_{i,r_1}^{(n)}\big), \hat{\psi}_{r_2}^{(n)}\big(Z_{i,r_2}^{(n)}\big)\big)$, respectively:

$$\hat{f}_{r_1}^{(n)}(x) = \frac{\big(\frac{x}{\ln n} + \frac{r_1 \ln\ln n}{\ln n} + \frac{\varkappa_{r_1}^{(n)}}{\ln n} + 1\big)^{r_1} \mathrm{e}^{-x-\varkappa_{r_1}^{(n)}}}{n\, r_1!} \cdot \mathbb{1}\big\{-\ln n - r_1 \ln\ln n - \varkappa_{r_1}^{(n)} \le x\big\},$$

$$\hat{f}^{(n)}_{r_1,r_2}(x,y) = \frac{\left(\frac{x}{\ln n} + \frac{r_1 \ln\ln n}{\ln n} + \frac{\varkappa^{(n)}_{r_1}}{\ln n} + 1\right)^{r_1} \mathrm{e}^{-y-\varkappa^{(n)}_{r_2}}}{n\, r_1!\, (r_2 - r_1 - 1)!}$$
$$\times \left(y - x + (r_2 - r_1)\ln\ln n + \left(\varkappa^{(n)}_{r_2} - \varkappa^{(n)}_{r_1}\right)\right)^{r_2 - r_1 - 1} (\ln n)^{-(r_2 - r_1)}$$
$$\times \mathbb{1}\left\{-\ln n - r_2 \ln\ln n - \varkappa^{(n)}_{r_2} \le x - (r_2 - r_1)\ln\ln n - \left(\varkappa^{(n)}_{r_2} - \varkappa^{(n)}_{r_1}\right) \le y\right\}.$$

Hence,

$$n\hat{P}^{(n)}_{r_1} = n\int_{B_{r_1}} \hat{f}^{(n)}_{r_1}(x)\,\mathrm{d}x = \frac{1}{r_1!}\int_{B_{r_1}} \left(\frac{x}{\ln n} + \frac{r_1 \ln\ln n}{\ln n} + \frac{\varkappa^{(n)}_{r_1}}{\ln n} + 1\right)^{r_1} \mathrm{e}^{-x-\varkappa^{(n)}_{r_1}}$$
$$\times \mathbb{1}\{-\ln n - r_1 \ln\ln n - \varkappa^{(n)}_{r_1} \le x\}\,\mathrm{d}x \to \frac{1}{r_1!}\int_{B_{r_1}} \mathrm{e}^{-x}\,\mathrm{d}x \quad \text{as } n \to \infty \tag{51}$$

by dominated convergence since B_{r_1} is bounded from below.

The proof that $\lim_{n\to\infty} n\hat{P}^{(n)}_{r_1,r_2} = 0$ is a bit more technical than a similar piece in the proof of Theorem 4. Letting $\alpha = \inf B_{r_1}$, $\beta = \inf B_{r_2}$, and noting that $\alpha, \beta > -\infty$, we have

$$n\hat{P}^{(n)}_{r_1,r_2} \le n\int_{\alpha}^{+\infty}\int_{\beta}^{+\infty} \hat{f}^{(n)}_{r_1,r_2}(x,y)\,\mathrm{d}x\,\mathrm{d}y$$
$$= \frac{(\ln n)^{-(r_2-r_1)}}{r_1!\,(r_2 - r_1 - 1)!}\int_{\alpha}^{+\infty}\int_{\beta}^{+\infty} \left(\frac{x}{\ln n} + \frac{r_1 \ln\ln n}{\ln n} + \frac{\varkappa^{(n)}_{r_1}}{\ln n} + 1\right)^{r_1}$$
$$\times \left(y - x + (r_2 - r_1)\ln\ln n + \left(\varkappa^{(n)}_{r_2} - \varkappa^{(n)}_{r_1}\right)\right)^{r_2 - r_1 - 1} \mathrm{e}^{-y-\varkappa^{(n)}_{r_2}}$$
$$\times \mathbb{1}\left\{-\ln n - r_2 \ln\ln n - \varkappa^{(n)}_{r_2} \le x - (r_2 - r_1)\ln\ln n - \left(\varkappa^{(n)}_{r_2} - \varkappa^{(n)}_{r_1}\right) \le y\right\}\mathrm{d}x\,\mathrm{d}y.$$

Setting $z = y - x + (r_2 - r_1)\ln\ln n + \left(\varkappa^{(n)}_{r_2} - \varkappa^{(n)}_{r_1}\right)$, we then obtain

$$n\hat{P}^{(n)}_{r_1,r_2} \le \frac{1}{r_1!\,(r_2 - r_1 - 1)!}\int_{\alpha}^{+\infty}\int_{0}^{+\infty} \left|\frac{x}{\ln n} + \frac{r_1 \ln\ln n}{\ln n} + \frac{\varkappa^{(n)}_{r_1}}{\ln n} + 1\right|^{r_1}$$
$$\times z^{r_2 - r_1 - 1} \mathrm{e}^{-(x + z + \varkappa^{(n)}_{r_1})} \cdot \mathbb{1}\left\{x + z \ge \beta + (r_2 - r_1)\ln\ln n + \left(\varkappa^{(n)}_{r_2} - \varkappa^{(n)}_{r_1}\right)\right\}\mathrm{d}x\,\mathrm{d}z.$$

Again by dominated convergence, the right-hand side vanishes as $n \to \infty$. This completes the proof of (48).

The proof of (49) is similar to that of (28). Analogously to (50),

$$\mathbb{E}\hat{H}^{(n)}(U) = \mathbb{E}\hat{H}^{(n)}\left(\bigcup_{r=0}^{s}\left(\{r\} \times B_r\right)\right) = \sum_{r=0}^{s} \mathbb{E}\hat{\eta}^{(n)}_r(B_r).$$

Since $\hat{\eta}^{(n)}_r(B_r) \sim \mathsf{Bin}\left(n, \hat{P}^{(n)}_r\right)$, (51) yields

$$\mathbb{E}\hat{H}^{(n)}(U) = \sum_{r=0}^{s} n\hat{P}^{(n)}_r \xrightarrow[n\to\infty]{} \sum_{r=0}^{s} \frac{1}{r!}\int_{B_r} \mathrm{e}^{-x}\,\mathrm{d}x = \mathbb{E}H(U).$$

This proves (49), which, along with (48), delivers $\hat{H}^{(n)} \xrightarrow[n\to\infty]{vd} H$.

Our next goal is to carry out a depoissonization procedure which enables to turn $\hat{H}^{(n)} \xrightarrow{vd} H$ into $\hat{\Xi}^{(n)} \xrightarrow{vd} H$. Let $\mathcal{C}$ stand for the ring of finite unions of disjoint closed segments on different levels of $\mathbb{X}$, which are bounded with respect to d in (22):

$$\mathcal{C} = \Big\{ \bigcup_{r=0}^{s} \bigcup_{k=1}^{l_r} \big(\{r\} \times [a_{r,k}, b_{r,k}] \big) \colon s, l_0, \dots, l_s \in \mathbb{N}_0 \Big\}.$$

Here $a_{r,k}, b_{r,k} \in \mathbb{R} \cup \{+\infty\}$, and we will always use the convention that $+\infty \pm \varepsilon = +\infty$. In the terminology of [22], $\mathcal{C}$ is a dissecting ring in $(\mathbb{X}, d)$.

Fix any $C \in \mathcal{C}$. By Lemma 3.4 in [19], which was proved on the basis of the coupling formula (6), for any $n \in \mathbb{N}$, $\varepsilon > 0$, and all r, k we have

$$\begin{aligned} &\mathbb{P}\big\{ \hat{\xi}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \neq \hat{\eta}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \big\} \\ &\leq c_r \varepsilon^{-4} n^{-1} + \mathbb{P}\big\{ \hat{\eta}_r^{(n)}\big([a_{r,k} - \varepsilon, a_{r,k} + \varepsilon]\big) \geq 1 \big\} + \mathbb{P}\big\{ \hat{\eta}_r^{(n)}\big([b_{r,k} - \varepsilon, b_{r,k} + \varepsilon]\big) \geq 1 \big\} \end{aligned}$$

with some $c_r > 0$. (In fact, this lemma was proved for the centering/normalizing function $\psi_r^{(n)}$, not $\hat{\psi}_r^{(n)}$, and for $a_{r,k}, b_{r,k} \in \mathbb{R}$, not $\mathbb{R} \cup \{+\infty\}$, but its proof remains valid also in our case.) Taking $\varepsilon = n^{-\frac{1}{5}}$, we then obtain:

$$\begin{aligned} \mathbb{P}\big\{ \hat{\xi}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \neq \hat{\eta}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \big\} &\leq c_r n^{-\frac{1}{5}} \\ &+ \mathbb{P}\big\{ \hat{\eta}_r^{(n)}\big([a_{r,k} - n^{-\frac{1}{5}}, a_{r,k} + n^{-\frac{1}{5}}]\big) \geq 1 \big\} \\ &+ \mathbb{P}\big\{ \hat{\eta}_r^{(n)}\big([b_{r,k} - n^{-\frac{1}{5}}, b_{r,k} + n^{-\frac{1}{5}}]\big) \geq 1 \big\}. \end{aligned} \tag{52}$$

Let η_r, $r \in \mathbb{N}_0$, stand for the Poisson point process on $\mathbb{R} \cup \{+\infty\}$ with intensity measure λ_r given by

$$\lambda_r(B) = \frac{1}{r!} \int_B \mathrm{e}^{-x} \, \mathrm{d}x, \quad B \in \mathfrak{B}\big(\mathbb{R} \cup \{+\infty\}\big).$$

Actually, η_r are the single-level point processes which make up the limiting multilevel process H. Since $\hat{H}^{(n)} \xrightarrow[n\to\infty]{vd} H$ implies $\hat{\eta}_r^{(n)} \xrightarrow[n\to\infty]{vd} \eta_r$, we get

$$\mathbb{P}\big\{ \hat{\eta}_r^{(n)}\big([a_{r,k} - n^{-\frac{1}{5}}, a_{r,k} + n^{-\frac{1}{5}}]\big) \geq 1 \big\} \xrightarrow[n\to\infty]{} \mathbb{P}\big\{ \eta_r\big(\{a_{r,k}\}\big) \geq 1 \big\} = 0,$$

and the same holds for $b_{r,k}$. Here the simultaneous passage to the limit in the measure and its argument is justified by Lemma 3.5 in [19]. Hence, by (52)

$$\mathbb{P}\big\{ \hat{\xi}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \neq \hat{\eta}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \big\} \to 0, \qquad n \to \infty.$$

So, by the coupling inequality we have for any $m \in \mathbb{N}_0$

$$\begin{aligned} \big| \mathbb{P}\big\{ \hat{\Xi}^{(n)}(C) = m \big\} - \mathbb{P}\big\{ \hat{H}^{(n)}(C) = m \big\} \big| &\leq \mathbb{P}\big\{ \hat{\Xi}^{(n)}(C) \neq \hat{H}^{(n)}(C) \big\} \\ &\leq \sum_{r=0}^{s} \sum_{k=1}^{l_r} \mathbb{P}\big\{ \hat{\xi}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \neq \hat{\eta}_r^{(n)}\big([a_{r,k}, b_{r,k}]\big) \big\} \to 0, \qquad n \to \infty. \end{aligned}$$

Thus, it follows from $\hat{H}^{(n)} \xrightarrow[n\to\infty]{vd} H$ that

$$\lim_{n\to\infty} \mathbb{P}\big\{\hat{\Xi}^{(n)}(C) = m\big\} = \lim_{n\to\infty} \mathbb{P}\big\{\hat{H}^{(n)}(C) = m\big\} = \mathbb{P}\big\{H(C) = m\big\}$$

for any $C \in \mathcal{C}$ and $m \in \mathbb{N}_0$. Since $\mathcal{C}$ is a dissecting ring and the limiting point process H is simple, Theorem 4.15 in [22] yields $\hat{\Xi}^{(n)} \xrightarrow[n\to\infty]{vd} H$. This completes the proof. □

7 Proof of Theorems 2 and 3

In this section, we apply Theorem 6 to the proof of Theorems 2 and 3. As in Section 5, it suffices only to prove the convergence of finite-dimensional distributions.

Similarly to (45), let

$$\begin{aligned} \hat{T}_r^{(n)} &= \inf\big\{x \in \mathbb{R}\colon \hat{\Xi}^{(n)}\big(\{r\} \times (x, +\infty)\big) = 0\big\}, \\ \hat{V}_r^{(n)} &= \hat{\Xi}^{(n)}\big(\{r\} \times \big(\hat{T}_0^{(n)}, +\infty\big)\big) \end{aligned} \tag{53}$$

for $r \in \mathbb{N}_0$. Particularly, $\hat{V}_0^{(n)} = 0$ a.s. By definition of $\hat{\Xi}^{(n)}$ and $\hat{\xi}_r^{(n)}$, and according to (4), it holds that

$$\hat{T}_r^{(n)} = \max_{i\le n} \hat{\psi}_r^{(n)}\big(Y_{i,r}^{(n)}\big) = \hat{\psi}_r^{(n)}\big(T_r^{(n)}\big). \tag{54}$$

Thus, $\hat{T}_r^{(n)}$ means the (centered and normalized) time when the rth collector completes his album in the original, non-poissonized scheme. Next, $\hat{V}_r^{(n)}$ is the number of points of the single-level process $\hat{\xi}_r^{(n)}$ to the right of $\hat{T}_0^{(n)}$. By definition of $\hat{\xi}_r^{(n)}$ and according to (47), (54), we have

$$\hat{V}_r^{(n)} = \operatorname{card}\big\{i\colon Y_{i,r}^{(n)} > T_0^{(n)} + rn\ln\ln n + n\big(\varkappa_r^{(n)} - \varkappa_0^{(n)}\big)\big\}.$$

Comparing this with (9), we can see that, with any choice of $\mathcal{O}(n)$ in (10), one may choose $\varkappa_r^{(n)}$, $r \in \mathbb{N}_0$, in such a way that $\hat{V}_r^{(n)} = \hat{U}_r^{(n)}$.

Just as it was done in Section 5, Theorem 6 and the Skorokhod coupling imply that, for any $s \in \mathbb{N}$,

$$\big(\hat{U}_r^{(n)},\ r = 1, \dots, s\big) \xrightarrow{d} (N_r(E/r!),\ r = 1, \dots, s), \qquad n \to \infty,$$

where $E \sim \mathsf{Exp}(1)$, and N_r are unit-rate Poisson counting processes, independent of each other and of E. This completes the proof of (12). The generating function of the random vector on the right-hand side can be easily calculated by means of conditioning with respect to E:

$$\begin{aligned} &\mathbb{E}\Big(\prod_{r=1}^{s} u_r^{N_r(E/r!)}\Big) = \mathbb{E}\Big(\prod_{r=1}^{s} \mathbb{E}_E\big(u_r^{N_r(E/r!)}\big)\Big) = \mathbb{E}\Big(\prod_{r=1}^{s} \mathrm{e}^{-E(1-u_r)/r!}\Big) \\ &= \int_0^{+\infty} \mathrm{e}^{-x - x\sum_{r=1}^{s}\frac{1-u_r}{r!}}\,\mathrm{d}x = \Big(1 + \sum_{r=1}^{s}\frac{1-u_r}{r!}\Big)^{-1}, \qquad (u_1, \dots, u_s) \in [0,1]^s. \end{aligned}$$

Letting here $s \to \infty$ yields (11).

Finally, again Theorem 6 and the Skorokhod coupling, combined with (54), (53), and Remarks 3, 4 imply that, for any $s \in \mathbb{N}_0$,

$$\begin{aligned}\left(\hat{\psi}_r^{(n)}\left(T_r^{(n)}\right), r = 0, \dots, s\right) &= \left(\hat{T}_r^{(n)}, r = 0, \dots, s\right) \\ &\xrightarrow{d} \left(-\ln r! - \ln E_r, r = 0, \dots, s\right), \qquad n \to \infty,\end{aligned}$$

where E_r are independent $\mathsf{Exp}(1)$-distributed random variables. Here E_r actually means the first jump time of N_r. Thus, $-\ln r! - \ln E_r$ on the right-hand side, by Remark 4, is the rightmost point of $H\left(\{r\} \times \cdot\right)$, the rth level of the limiting process H. To complete the proof of (16), it suffices to note that $-\ln r! - \ln E_r \stackrel{d}{=} B_r$ from (15), and to put in (47) $\varkappa_r^{(n)} = 0$. □

8 Concluding remarks and open problems

In Theorems 1 and 2, the reference point is $T_0^{(n)}$, the time when the main collector completed his album. That is, we studied the number of empty spots in the album of the rth brother at time $T_0^{(n)}$ in Theorem 1 and at time $T_0^{(n)} + d_r^{(n)}$ in Theorem 2. Similarly, we could fix some $r_0 \in \mathbb{N}_0$ and consider the number of empty spots in the collection of the rth brother at times $T_{r_0}^{(n)}$ and $T_{r_0}^{(n)} + d_{r,r_0}^{(n)}$ with some $d_{r,r_0}^{(n)} \to +\infty$ as $n \to \infty$. The former makes sense only for $r > r_0$, while, in the latter case, we may consider $r < r_0$ with $d_{r,r_0}^{(n)} < 0$, replacing $d_{r,r_0}^{(n)} \to +\infty$ by $d_{r,r_0}^{(n)} \to -\infty$. Repeating the above proofs with minimal changes, we can derive analogous results with similar limiting distributions (exponential and geometric, respectively).

Another possible reference point is $T_0^{(n)}(m)$, $m \in \mathbb{N}_0$, the time when the main collector first assembled some $n - m$ (unspecified) of n coupons, or $T_{r_0}^{(n)}(m)$, the similar value for the r_0^{th} collector. In this notation, $T_{r_0}^{(n)}(0) = T_{r_0}^{(n)}$. In this case, we come to more general limiting distributions — gamma and negative binomial, respectively. This can be proved along the above lines.

Let us now dwell on some possible directions for further research. Computer simulations show that the rate of convergence in Theorems 1 and 2 is rather slow. It would be interesting to obtain, for each fixed $r \in \mathbb{N}$, upper bounds on the rates of convergence of the rth coordinates in the Wasserstein distance for Theorem 1 and in the total variation distance for Theorem 2, similarly to the recent work [9] on the Erdős–Rényi limit theorem (1), which is the one-dimensional version of Theorem 3. The most natural tool for this is Stein's method. The relevant techniques for exponential and geometric approximation are summarized, for example, in Sections 5 and 6 of [30]. Some applications of Stein–Chen Poisson approximation to the coupon collector's problem are given in Chapter 6 of [6] and in Examples 4.34 and 4.35 of [30].

The extended coupon collector's problem with unequal probabilities provides a natural direction for future research. Some steps in this direction were taken in [11]. Finally, related questions can also be posed for variants of the coupon collector's problem with group drawings, as considered, e.g., in [31] and in the recent work [3].

Acknowledgments

The author thanks the anonymous referee for suggestions that helped improve the exposition.

Funding

The author was supported by the Swiss National Science Foundation grant no. 229505.

References

[1] Abraham, S., Brockman, G., Sapp, S., Godbole, A.P.: Omnibus sequences, coupon collection, and missing word counts. Methodol. Comput. Appl. Probab. **15**(2), 363–378 (2013) MR3053962. https://doi.org/10.1007/s11009-011-9247-6

[2] Adler, I., Oren, S., Ross, S.M.: The coupon-collector's problem revisited. J. Appl. Probab. **40**(2), 513–518 (2003) MR1978107. https://doi.org/10.1017/s002190020001946x

[3] Athreya, S., Mukherjee, S., Mukherjee, S.S.: On a generalisation of the coupon collector problem. J. Theor. Probab. **38**(3) (2025) MR4958906. https://doi.org/10.1007/s10959-025-01432-x

[4] Barak-Pelleg, D., Berend, D.: The time for reconstructing the attack graph in DDoS attacks. J. Math. Anal. Appl. **531**(2) (2024) MR4665508. https://doi.org/10.1016/j.jmaa.2023.127889

[5] Barak-Pelleg, D., Berend, D.: Fast simulations of the multi-album collector. Theor. Comput. Sci. **1053** (2025) MR4929594. https://doi.org/10.1016/j.tcs.2025.115432

[6] Barbour, A.D., Holst, L., Janson, S.: Poisson Approximation, 277 pp. Oxford: Clarendon Press (1992) MR1163825

[7] Bates, G.E., Neyman, J.: Contribution to the theory of accident proneness. II: true or false contagion. Univ. Calif. Publ. Statist. **1**(10), 255–276 (1952) MR0050837

[8] Boneh, A., Hofri, M.: The coupon-collector problem revisited – a survey of engineering problems and computational methods. Commun. Stat., Stoch. Models **13**(1), 39–66 (1997) MR1430927. https://doi.org/10.1080/15326349708807412

[9] Costacèque, B., Decreusefond, L.: Convergence rate for the coupon collector's problem with Stein's method. Stoch. Model. Appl. **193** (2026) MR5000285. https://doi.org/10.1016/j.spa.2025.104835

[10] Davy, P.J., Rayner, J.C.W.: Multivariate geometric distributions. Commun. Stat., Theory Methods **25**(12), 2971–2987 (1996)

[11] Doumas, A.V., Papanicolaou, V.G.: The siblings of the coupon collector. Theory Probab. Appl. **62**(3), 444–470 (2018) MR3684649. https://doi.org/10.4213/tvp5129

[12] Erdős, P., Rényi, A.: On a classical problem of probability theory. Publ. Math. Inst. Hung. Acad. Sci., Ser. A **6**, 215–220 (1961) MR0150807

[13] Ferrari, A., Letac, G., Tourneret, J.-Y.: Multivariate mixed Poisson distributions. In: 2004 12th European Signal Processing Conference, 1067–1070 (2004)

[14] Foata, D., Zeilberger, D.: The collector's brotherhood problem using the Newman-Shepp symbolic method. Algebra Univers. **49**(4), 387–395 (2003) MR2022346. https://doi.org/10.1007/s00012-003-1815-3

[15] Foata, D., Han, G.-N., Lass, B.: Les nombres hyperharmoniques et la fratrie du collectionneur de vignettes. Sémin. Lothar. Comb. **47** (2001) MR1894021

[16] Grandell, J.: Mixed Poisson Processes, 268 pp. London: Chapman & Hall (1997) MR1463943. https://doi.org/10.1007/978-1-4899-3117-7

[17] Harremoës, P., Johnson, O., Kontoyiannis, I.: Thinning, entropy, and the law of thin numbers. IEEE Trans. Inf. Theory **56**(9), 4228–4244 (2010) MR2807322. https://doi.org/10.1109/TIT.2010.2053893

[18] Holst, L.: On birthday, collectors', occupancy and other classical urn problems. Int. Stat. Rev. **54**, 15–27 (1986) MR0959649. https://doi.org/10.2307/1403255

[19] Ilienko, A.: Convergence of point processes associated with coupon collector's and dixie cup problems. Electron. Commun. Probab. **24**, 9 (2019) MR4003125. https://doi.org/10.1214/19-ecp263

[20] Ilienko, A., Stamatiieva, V.: Limit theorems in the generalized birthday problem. Bull. T. Shevchenko Natl. Univ. Kyiv Phys. Math. **79**(2), 14–19 (2024) MR5029736. https://doi.org/10.17721/1812-5409

[21] Johnson, N.L., Kotz, S.: Urn Models and Their Application. An Approach to Modern Discrete Probability Theory, 415 pp. John Wiley & Sons, Hoboken, NJ (1977) MR0488211

[22] Kallenberg, O.: Random Measures, Theory and Applications, 694 pp. Cham: Springer (2017) MR3642325. https://doi.org/10.1007/978-3-319-41598-7

[23] Kuba, M., Panholzer, A.: On moment sequences and mixed Poisson distributions. Probab. Surv. **13**, 89–155 (2016) MR3550936. https://doi.org/10.1214/14-PS244

[24] Last, G., Penrose, M.: Lectures on the Poisson Process, 293 pp. Cambridge: Cambridge University Press (2017) MR3791470

[25] Luko, S.N.: The "coupon collector's problem" and quality control. Qual. Eng. **21**(2), 168–181 (2009)

[26] Pintacuda, N.: Coupons collectors via the martingales. Boll. Unione Mat. Ital., V. Ser., A **17**, 174–177 (1980) MR0562133

[27] Rényi, A.: A characterization of Poisson processes. Publ. Math. Inst. Hung. Acad. Sci. **1**, 519–527 (1957) MR0094861

[28] Resnick, S.I.: Extreme Values, Regular Variation, and Point Processes, 334 pp. New York etc.: Springer-Verlag (1987) MR0900810. https://doi.org/10.1007/978-0-387-75953-1

[29] Resnick, S.I.: Heavy-tail Phenomena. Probabilistic and Statistical Modeling, 404 pp. New York: Springer (2007) MR2271424

[30] Ross, N.: Fundamentals of Stein's method. Probab. Surv. **8**, 210–293 (2011) MR2861132. https://doi.org/10.1214/11-PS182

[31] Stadje, W.: The collector's problem with group drawings. Adv. Appl. Probab. **22**(4), 866–882 (1990) MR1082197. https://doi.org/10.2307/1427566

[32] Steutel, F.W., van Harn, K.: Discrete analogues of self-decomposability and stability. Ann. Probab. **7**, 893–899 (1979) MR0542141

[33] Yakymiv, A.L.: A generalization of the Curtiss theorem for moment generating functions. Math. Notes **90**(6), 920–924 (2011) MR2962966. https://doi.org/10.1134/S0001434611110290