



HAL
open science

On a problem of team hiring

Jin-Yi Liu

► **To cite this version:**

| Jin-Yi Liu. On a problem of team hiring. 2021. hal-02485153v2

HAL Id: hal-02485153

<https://hal.science/hal-02485153v2>

Preprint submitted on 26 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

On a problem of team hiring

Jin-Yi Liu

*School of Computer and Communication Engineering
Liaoning Shihua University
Fushun, Liaoning 113001, P.R. China
j-y-liu@sina.com*

Abstract

Given two positive integers n and k with $k \leq n$, let \mathcal{X} denote the family of all the k -subsets of $[n] := \{1, 2, \dots, n\}$. In this paper, assuming that the $\binom{n}{k}$ k -subsets hold some order of ranks, we consider the expected number of left-to-right maxima in some sequential ordering of \mathcal{X} , under a random permutation of $[n]$. In the case $k = 1$, it is well-known that the answer is the n th harmonic number $H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n} = \ln n + O(1)$. For general k , the contribution of this paper is about $E_{\text{LEX}}^*(n, k)$ and $E_{\text{COL}}^*(n, k)$, the expected numbers of left-to-right maxima respectively in the lexicographical and colexicographical orderings of \mathcal{X} , achieved when the k -subsets in \mathcal{X} hold their respective worst orders of ranks (to make the two values as large as possible). We show that $E_{\text{LEX}}^*(n, k) = E_{\text{COL}}^*(n, k)$, and give an exact formula (not in closed form) for them. For estimating them, we further show that when $k' := \min\{k, n - k\}$ is fixed and n is big enough, they are asymptotic to $\frac{1}{k'^!}(\ln n)^{k'}$.

The problem we consider here can be viewed as an extension to the assistant-hiring problem, presented in the textbook *Introduction to Algorithms* by Cormen, Leiserson, Rivest, and Stein, for introducing probabilistic analysis and randomized algorithms.

1. Introduction

In their legendary textbook *Introduction to Algorithms* [5, Section 5.1, both the second and third editions], Cormen et al. take an assistant-hiring problem as the primary example to introduce probabilistic analysis and randomized algorithms. Here, we extend that hiring problem to a team-hiring scenario, which may lead to closer insight into these two subjects. To adapt with team hiring, let us alter their original story to some extent. Suppose that our organization wish to find the best-qualified k -person team from n applicants and hire them permanently. The qualification of a team can only be determined by a sophisticated test (for collaborative capability, etc.), and the result of the test is unpredictable. For the work at hand not to be postponed, we decide to adopt an exhaustive testing-and-hiring process to accomplish our task. At the beginning of the process, the applicants are assumed to have been randomly numbered from 1 to n , and a least-qualified dummy team has been hired. Then, each k -person team is tested each morning, one by one in accordance with some ordering of all the $\binom{n}{k}$ teams. If the team just tested is better than the current hired team,

we discard the old team and hire the new team until a more qualified team is determined. In the end, the lastly hired team will be hired permanently. Now, under the assumption that the qualifications of all the teams are distinct, we have two concerns in our team-hiring problem: the time and space requirements for generating the ordering of all the teams, and the expected number of hired teams in the whole testing-and-hiring process (because discarding an old team and organizing a new team costs money and energy). Clearly, the latter is affected by two factors: the order in which we test the teams, and the order of qualifications (ranks) that the $\binom{n}{k}$ teams hold.

In the case $k = 1$, as analyzed in the textbook [5] and noted in many other references (e.g., [7]), the expected number of hired teams is the n th harmonic number $H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n} = \ln n + O(1)$, no matter in what order we test the n applicants, and no matter what is the order of ranks that the n applicants hold. In the case $k > 1$, if we assume that the $\binom{n}{k}$ teams hold a random order of ranks, the expected number of hired teams is ideally $H_{\binom{n}{k}} = O(k' \ln n)$ with $k' := \min\{k, n - k\}$, no matter in what order we test the teams (see Proposition 1). However, for a sequential order to test the teams, this result is not of probabilistic analysis to the team-hiring problem, because the assumption that “the teams hold a random order of ranks”, unlike the assumption that “the applicants are randomly numbered”, is not a proper assumption on the inputs (of which the “teams” are not a part).

If we want to achieve the goal that the expected number of hired teams is always $H_{\binom{n}{k}}$ for $k > 1$ without any assumption on the order of ranks that the teams hold, to test the teams in a random order may be of first choice, such that we get a randomized process. To fulfil this, we need the Knuth shuffle (or called Fisher-Yates Shuffle) [12, Algorithm P] to generate a random ordering of all the teams, either by shuffling all the k -subsets of $\{1, 2, \dots, n\}$ directly, or by shuffling the set $\{1, 2, \dots, \binom{n}{k}\}$ firstly and then reconstructing the k -subsets with an UNRANK procedure which follows some sequential ordering [14, Algorithms 2.8, 2.10, or 2.12]. Although the Knuth shuffle depends on a pseudorandom number generator for implementation on computers, it is widely accepted as a uniform random permutation generator [5]. The drawback of applying it here is that the shuffling process needs $\Omega(\binom{n}{k})$ working space.

If the above random ordering cannot be accepted for the sake of space requirements, some pseudorandom or quasi-random orderings may be of consideration. To generate such orderings, we need in a stateless way to generate a pseudorandom or quasi-random permutation of $\{1, 2, \dots, \binom{n}{k}\}$, where we can use the *linear congruential sequence* with period length $\binom{n}{k}$ [12, Theorem A], or use the truncated *bit-reversal permutation* [15, Section V]. With such methods, the space requirements are cut down, but we still need an UNRANK procedure to reconstruct each k -subset, in which a quantity of binomial coefficients need to be computed and big numbers must be involved if $\binom{n}{k}$ is big (see [14] for detail). What is worse is that we (at least the author) are not sure that these pseudorandom or quasi-random orderings can achieve the same or a similar result as the real random ordering.

In this paper, we focus on two well-known sequential orderings: the lexicographical (or lexical) and colexicographical (or colex) orderings. Compared with the above random, pseudorandom or quasi-random orderings, the sequences of k -subsets obeying these two orderings

can be generated very efficiently in both time and space. As shown in [13], [14], [17], and [20], the current k -subset to be generated is only dependent on its predecessor, so all the k -subsets can be generated with a SUCCESSOR procedure, in $O(k)$ time per subset (indeed amortized constant time if $k \leq n/2$ and the time for output is excluded), and in $O(k)$ working space, with no big numbers involved if n is not big. Although they are quite nice in this regard, we have the intuition that they should be bad in term of the expected number of hired teams they induce in the worst case. But how bad are they? This paper tries to give a quantitative answer to this question.

1.1. Problem formalization and our results

Throughout this paper, an *ordering* of a family of subsets, is referred to the sequence of these subsets generated in some *ordering* way. Here, we follow [14] (rather than [13]), for the normal forms of the lexical and colex orderings of all the k -subsets of the set $\{1, 2, \dots, n\}$, which can be seen from the examples hereafter. Given a sequence of elements from a well-order set, a (strong) *left-to-right maximum* of the sequence is the element which is greater than any element preceding it.

Given two positive integer n and k with $k \leq n$, the team-hiring problem in this paper is only determined by n and k , so we call it (n, k) -THP for short. We formalize it with the following notations:

- $[n] := \{1, 2, \dots, n\}$;
- \mathcal{X} : the family of all the $\binom{n}{k}$ k -subsets of $[n]$;
- *ord*: the variable tag for an ordering way to generate an ordering of \mathcal{X} ;
- \mathcal{X}_{ord} : the ordering of \mathcal{X} generated in the way of *ord*;
- $\sigma : [n] \rightarrow [n]$, a permutation of $[n]$;
- S_n : the set of all the $n!$ permutations of $[n]$;
- $\{t_1, t_2, \dots, t_k\}^\sigma := \{\sigma(t_1), \sigma(t_2), \dots, \sigma(t_k)\}$, supposing that $\{t_1, t_2, \dots, t_k\} \in \mathcal{X}$;
- $\mathcal{X}_{ord}^\sigma := \langle X_1^\sigma, X_2^\sigma, \dots, X_{\binom{n}{k}}^\sigma \rangle$, supposing that $\mathcal{X}_{ord} = \langle X_1, X_2, \dots, X_{\binom{n}{k}} \rangle$;
- $f : \mathcal{X} \rightarrow [\binom{n}{k}]$, a bijective ranking function on \mathcal{X} ;
- $\mathcal{M}_{ord}^{\sigma, f} := \{X_j^\sigma \in \mathcal{X}_{ord}^\sigma : 1 \leq j \leq \binom{n}{k}, f(X_i^\sigma) < f(X_j^\sigma) \text{ for all } i < j\}$, the collection of left-to-right maximum k -subsets in \mathcal{X}_{ord}^σ , for a fixed σ and a fixed f ;
- $E_{ord}^f(n, k) := (1/n!) \sum_{\sigma \in S_n} |\mathcal{M}_{ord}^{\sigma, f}|$, the expected number of left-to-right maximum k -subsets in \mathcal{X}_{ord}^σ , for a random $\sigma \in S_n$ and a fixed f .

Additionally, we let “LEX” and “COL” be the tags of the lexical and colex orderings respectively, and (from now on) we use the convention that a permutation of $[n]$ is written as a n -word and a k -subset of $[n]$ is written as a k -word with elements in the ascending order. (When the ascending order cannot be guaranteed, we still use the form $\{\dots\}$ for a subset.) For a ranking function f on \mathcal{X} , we generally write it as an ordering of \mathcal{X} to express it. That is, $f = \langle X_1, X_2, \dots, X_{\binom{n}{k}} \rangle$ means that $f(X_1) = 1, f(X_2) = 2, \dots, f(X_{\binom{n}{k}}) = \binom{n}{k}$. Besides, the inequality $f(X_1) < f(X_2)$ is generally written as $X_1 <_f X_2$.

To illustrate the above notations, let's consider $(4, 2)$ -THP. If we choose $f = \langle 13, 23, 34, 24, 12, 14 \rangle$, then for a fixed $\sigma = 3214$, we have

$$\begin{aligned} \mathcal{X}_{\text{LEX}} &= \langle 12, 13, 14, 23, 24, 34 \rangle, & \mathcal{X}_{\text{COL}} &= \langle 12, 13, 23, 14, 24, 34 \rangle, \\ \mathcal{X}_{\text{LEX}}^\sigma &= \langle 23, 13, 34, 12, 24, 14 \rangle, & \mathcal{X}_{\text{COL}}^\sigma &= \langle 23, 13, 12, 34, 24, 14 \rangle, \\ \mathcal{M}_{\text{LEX}}^{\sigma, f} &= \{23, 34, 12, 14\}, & \mathcal{M}_{\text{COL}}^{\sigma, f} &= \{23, 12, 14\}, \\ |\mathcal{M}_{\text{LEX}}^{\sigma, f}| &= 4, & |\mathcal{M}_{\text{COL}}^{\sigma, f}| &= 3. \end{aligned}$$

Summarizing over all $\sigma \in S_4$, we can further obtain that

$$E_{\text{LEX}}^f(4, 2) = \frac{61}{24}, \text{ and } E_{\text{COL}}^f(4, 2) = \frac{59}{24}.$$

Let \mathcal{F} be the collection of all the ranking functions on \mathcal{X} . Although \mathcal{F} can be partitioned into at most $\binom{n}{k}!/n!$ equivalence classes (see Section 2.5), where the functions f in the same class induce the same $E_{\text{ord}}^f(n, k)$ for any fixed ord , we simply think of \mathcal{F} as $|\mathcal{F}| = \binom{n}{k}!$. Here, we concentrate on the average and the worst behavior of $f \in \mathcal{F}$, and only is the latter essential to the probabilistic analysis of the team-hiring problem. Precisely, we have two definitions that

$$E_{\text{ord}}^\circ(n, k) := (1/\binom{n}{k}!) \sum_{f \in \mathcal{F}} E_{\text{ord}}^f(n, k), \text{ and } E_{\text{ord}}^*(n, k) := \max_{f \in \mathcal{F}} E_{\text{ord}}^f(n, k).$$

Our first result of this paper is about how much $E_{\text{ord}}^\circ(n, k)$ is, which is both obvious and intriguing. It belongs to the rare cases in the world that the ‘‘average’’ is easier to solve than the ‘‘worst’’.

Proposition 1. *For any ordering ord of \mathcal{X} , $E_{\text{ord}}^\circ(n, k) = H_{\binom{n}{k}}$.*

Proof. For any fixed $\sigma \in S_n$, $(1/\binom{n}{k}!) \sum_{f \in \mathcal{F}} |\mathcal{M}_{\text{ord}}^{\sigma, f}| = H_{\binom{n}{k}}$. □

Our main result of this paper is an exact formula (not in closed form) for $E_{\text{LEX}}^*(n, k)$ and $E_{\text{COL}}^*(n, k)$.

Theorem 1. *For $1 \leq k \leq n$, we have*

$$E_{\text{LEX}}^*(n, k) = E_{\text{COL}}^*(n, k) = \sum_{1 \leq t_1 < \dots < t_k \leq n} P(t_1 \dots t_k),$$

where

$$P(t_1 \dots t_k) = \begin{cases} 1/t_1 & \text{if } k = 1 \\ (1/t_k) \sum_{1 \leq j \leq k} P(t_1 \dots t_{j-1}(t_{j+1}-1) \dots (t_k-1)) & \text{otherwise.} \end{cases} \quad (1)$$

In recurrence (1), the expression $t_1 \dots t_{j-1}(t_{j+1}-1) \dots (t_k-1)$ is understood to be a $(k-1)$ -subset of $[n-1]$ obtained from the k -subset $t_1 \dots t_k$ of $[n]$, by removing t_j and decrementing the elements larger than t_j (if any), and $P(\cdot)$ can be temporally viewed as a function which can take any nonempty subset of $[n]$ as its argument. In Section 2, we will see that $P(t_1 \dots t_k)$

really corresponds to the probability that the subset $t_1 \cdots t_k$ achieves left-to-right maxima in \mathcal{X}_{LEX} or \mathcal{X}_{COL} . To illustrate recurrence (1), letting 247 be a 3-subset, we have $P(247) = \frac{1}{7}[P(24) + P(26) + P(36)] = \frac{1}{7}[\frac{1}{4}(\frac{1}{2} + \frac{1}{3}) + \frac{1}{6}(\frac{1}{2} + \frac{1}{5}) + \frac{1}{6}(\frac{1}{3} + \frac{1}{5})]$.

We can immediately check that Theorem 1 well fits the cases $k = 1$ and $n = k$. Applying Theorem 1 to (4, 2)-THP, we have

$$\begin{aligned} E_{\text{LEX}}^*(4, 2) = E_{\text{COL}}^*(4, 2) &= P(12) + P(13) + P(23) + P(14) + P(24) + P(34) \\ &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} + \frac{5}{24} + \frac{1}{6} \\ &= \frac{61}{24}, \end{aligned}$$

which can be verified by a simple computer program. We will give the proof of Theorem 1 for the general case in Section 2. From Theorem 1 we know that $E_{\text{LEX}}^*(n, k) = E_{\text{COL}}^*(n, k)$, thus we use a unified $E_{\text{LC}}^*(n, k)$ to denote both $E_{\text{LEX}}^*(n, k)$ and $E_{\text{COL}}^*(n, k)$. Now, it is not difficult to derive a formula in closed form for $E_{\text{LC}}^*(n, 2)$, a formula in summation form for $E_{\text{LC}}^*(n, 3)$, and the subsequent asymptotics for them as well.

Corollary 1. *We have*

$$E_{\text{LC}}^*(n, 2) = \frac{1}{2}(H_n^2 - H_n^{(2)}) + H_n - 1 = \frac{1}{2}\ln^2 n + O(\ln n), \text{ and}$$

$$E_{\text{LC}}^*(n, 3) = \sum_{3 \leq i \leq n} \left(\frac{1}{2i}(H_{i-2}^2 - H_{i-2}^{(2)}) + \left(\frac{1}{i} + \frac{1}{i-1}\right)H_{i-2} + \frac{1}{i(i-1)} \right) = \frac{1}{6}\ln^3 n + O(\ln^2 n),$$

where $H_n^{(2)} = \sum_{1 \leq i \leq n} 1/i^2 < \frac{\pi^2}{6}$, which is the second-order n th harmonic number.

The sequence of $n!E_{\text{LC}}^*(n, 2)$ for $n \geq 2$ is $\langle 2, 11, 61, 379, 2668, \dots \rangle$, and the sequence of $n!E_{\text{LC}}^*(n, 3)$ for $n \geq 3$ is $\langle 6, 50, 379, 3023, 26193, \dots \rangle$. These two sequences have been submitted to the On-line Encyclopedia of the Integer Sequences(OEIS) [21], and published at the entries A308729 and A308860 respectively.

For general k and n , analogous to the symmetry of binomial coefficients, $E_{\text{LC}}^*(n, k)$ is symmetric in k and $n - k$, as might be expected. And it holds the bounds in terms of unsigned Stirling numbers of the first kind.

Corollary 2. *Specially define $E_{\text{LC}}^*(n, 0) := 1$. Then for $0 \leq k \leq n$, we have*

$$E_{\text{LC}}^*(n, k) = E_{\text{LC}}^*(n, n - k), \text{ and } \frac{1}{n!} \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \leq E_{\text{LC}}^*(n, k) \leq \frac{k!}{n!} \begin{bmatrix} n+1 \\ k+1 \end{bmatrix}.$$

The proof of Corollary 2 will be given in Section 3. By applying the well-known asymptotic formula $\frac{1}{n!} \begin{bmatrix} n+1 \\ k+1 \end{bmatrix} \sim \frac{1}{k!} (\ln n)^k$ for fixed k (see [19, Table 4.6]), we can immediately give a conditional asymptotic bound in terms of big-theta for $E_{\text{LC}}^*(n, k)$: $E_{\text{LC}}^*(n, k) = \Theta((\ln n)^{k'})$ for fixed $k' := \min\{k, n - k\}$. However, Corollary 1 hints that the coefficient of $(\ln n)^k$ in the asymptotic expansion of $E_{\text{LC}}^*(n, k)$ for fixed k seems to be $1/k!$, so that we can derive a more accurate estimate. This is true, and consequently we have the following corollary, whose proof will be given in Section 4.

Corollary 3. *For fixed $k' := \min\{k, n - k\}$ and as $n \rightarrow \infty$, we have*

$$E_{\text{LC}}^*(n, k) = (1/k')!(\ln n)^{k'} + O((\ln n)^{k'-1}).$$

1.2. Related work

In previous references, the left-to-right maxima of a sequence were often given other names, such as *records* and *outstanding elements*. Wilf [23] attributed to A. Rényi the result that the number of permutations of n letters that have exactly k left-to-right maxima is $\left[\begin{smallmatrix} n \\ k \end{smallmatrix} \right]$, an unsigned Stirling number of the first kind. It seems to be of the folklore that the expected number of left-to-right maxima in a random permutation of n letters is H_n , and the variance is $H_n - H_n^{(2)}$. Wilf [23] presented plenty of other properties of left-to-right maxima in permutations of n letters. In recent years, the (strong or weak) left-to-right maxima in other types of sequences have also been analyzed by many authors. For example, see [10], [11], [16], and [18].

For a special problem (identifying stars from a set U of n spikes with queries on k -subsets of U), Mueller et al. [15] compared several orderings of k -subsets with experiments, and showed that the quasi-random ordering based on bit reversals is superior to the lexical ordering, in terms of the time for discovering all possible sets of stars.

There are another class of “hiring problems” (e.g., [9] and references therein), where an applicant is either irrevocably rejected or permanently hired at once after an evaluation of her/his capabilities, and the objective is to hire one or several possibly top employee(s). These problems originated from the famous “secretary problem” [6], and from the practical point of view, what they model is much closer to realistic hiring situations.

There are another kind of maxima – vector maxima, which deserve to mention. A vector in a set of d -dimensional vectors is said to be maximal if it is not dominated by (in all dimensions smaller than) any other vector. A set of n d -vectors is said to be random if for each dimension, the components in the same dimension independently correspond to a random permutation of $[n]$, i.e, it is uniformly chosen from $n!^d$ possible vector sets. Under this assumption on randomness (or its equivalences), it is amazing that our $E_{\text{LC}}^*(n, k)$ is numerically related to $A(n, d)$, the expected number of maxima in a random set of n d -vectors. Bentley et al. [3] showed that $A(n, d)$ holds the recurrence

$$A(n, d) = A(n - 1, d) + (1/n)A(n, d - 1)$$

with the initial conditions $A(1, d) = 1$ and $A(n, 1) = 1$, and gave it the bound $O((\ln n)^{d-1})$ for fixed d . Buchta [4] derived a summation formula for $A(n, d)$ and gave it an asymptotic expansion for fixed d , whose first term is $(1/(d-1)!)(\ln n)^{d-1}$. For small d , we have $A(n, 2) = H_n$, $A(n, 3) = \frac{1}{2}H_n^2 + \frac{1}{2}H_n^{(2)}$, and $A(n, 4) = \frac{1}{6}H_n^3 + \frac{1}{2}H_nH_n^{(2)} + \frac{1}{3}H_n^{(3)}$, so we can observe that $A(n, d)$ is very close to $E_{\text{LC}}^*(n, d - 1)$. But there should be a large gap between them when d is large, because d may be greater than n , and $A(n, d)$ does not hold the property of symmetry in d and $n - d$. The variance of the number of maxima in a random vector set has also received attention, but its exact formula for general d is very complicated. Bai et al. [2] gave an asymptotic formula for the variance, whose form is still not simple by any means.

2. Proof of Theorem 1

2.1. Preliminaries and Proof Outline

First, let's consider general ord and general f . Let $\mathcal{X}_{ord} = \langle X_1, X_2, \dots, X_{\binom{n}{k}} \rangle$, and let $f = \langle Y_1, Y_2, \dots, Y_{\binom{n}{k}} \rangle$. Then, we say that $T_{ord}^f(X_i, Y_l)$ is the *contribution* of Y_l to X_i , with $T_{ord}^f(X_i, Y_l)$ being defined as follows:

$$T_{ord}^f(X_i, Y_l) := |\{\sigma \in S_n : X_i^\sigma = Y_l, X_i^\sigma \in \mathcal{M}_{ord}^{\sigma, f}\}|.$$

Observation 1. (i) $T_{ord}^f(X_i, Y_l) = 0$ for $l < i$. (ii) Suppose that f is variable. Then $T_{ord}^f(X_i, Y_l), T_{ord}^f(X_i, Y_{l+1}), \dots, T_{ord}^f(X_i, Y_{\binom{n}{k}})$ all remain unchanged if the consecutive subsequence $\langle Y_l, Y_{l+1}, \dots, Y_{\binom{n}{k}} \rangle$ remains unchanged.

In terms of contributions, we have another formula for $E_{ord}^f(n, k)$:

$$E_{ord}^f(n, k) = (1/n!) \sum_{1 \leq i \leq l \leq \binom{n}{k}} T_{ord}^f(X_i, Y_l).$$

Next, let's consider two special ranking functions which are respectively consistent with \mathcal{X}_{LEX} and \mathcal{X}_{COL} in their normal forms. We respectively name them *lex* and *col*. That is, in our simplified language, we can say that $\text{lex} = \mathcal{X}_{\text{LEX}}$ and $\text{col} = \mathcal{X}_{\text{COL}}$, or say that $<_{\text{lex}}$ and $<_{\text{col}}$ are two strict well-order relations on \mathcal{X} . Without any confusion, we always stipulate that $<_{\text{lex}}$ and $<_{\text{col}}$ can be used to compare any two distinct subsets of $[n]$ with the same cardinality.

Now let $s_1 \cdots s_k$ and $t_1 \cdots t_k$ be two distinct k -subsets of $[n]$. The definitions of $<_{\text{lex}}$ and $<_{\text{col}}$ can be given recursively as follows:

- $s_1 \cdots s_k <_{\text{lex}} t_1 \cdots t_k$ if $s_1 < t_1$, or $s_1 = t_1$ and $s_2 \cdots s_k <_{\text{lex}} t_2 \cdots t_k$.
- $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$ if $s_k < t_k$, or $s_k = t_k$ and $s_1 \cdots s_{k-1} <_{\text{col}} t_1 \cdots t_{k-1}$.

Observation 2. Let $c \in s_1 \cdots s_k \cap t_1 \cdots t_k$. Then both $<_{\text{lex}}$ and $<_{\text{col}}$ hold that

$$s_1 \cdots s_k <_{\text{lex}(\text{col})} t_1 \cdots t_k \iff s_1 \cdots s_k \setminus \{c\} <_{\text{lex}(\text{col})} t_1 \cdots t_k \setminus \{c\}.$$

Let $d \in [n]$. Then the mapping $\delta_d : [n] \setminus \{d\} \rightarrow [n-1]$ is called *d -pivoted decrement* (of order n) if it satisfies

$$\delta_d(i) = \begin{cases} i & \text{if } i < d \\ i - 1 & \text{otherwise.} \end{cases}$$

Clearly, δ_d is a bijection. Letting $d \notin t_1 \cdots t_k$, like the labeling of a permutation on a subset, we use $(t_1 \cdots t_k)^{\delta_d}$ to denote the subset $\delta_d(t_1)\delta_d(t_2) \cdots \delta_d(t_k)$. (It is guaranteed that $\delta_d(t_1) < \delta_d(t_2) < \cdots < \delta_d(t_k)$ if $d \notin t_1 \cdots t_k$.) Then it is not difficult to verify the following observation.

Observation 3. Let $d \in [n]$ and $d \notin s_1 \cdots s_k \cup t_1 \cdots t_k$. Then both $<_{\text{lex}}$ and $<_{\text{col}}$ hold that

$$s_1 \cdots s_k <_{\text{lex}(\text{col})} t_1 \cdots t_k \iff (s_1 \cdots s_k)^{\delta_d} <_{\text{lex}(\text{col})} (t_1 \cdots t_k)^{\delta_d}.$$

Now let $c \in t_1 \cdots t_k$. Then $(t_1 \cdots t_k \setminus \{c\})^{\delta_c}$ is a valid $(k-1)$ -subset of $[n-1]$. For convenience, we use $t_1 \cdots t_k \ominus c$ to denote it. That is, when $c \in t_1 \cdots t_k$, we have the definition that

$$t_1 \cdots t_k \ominus c := (t_1 \cdots t_k \setminus \{c\})^{\delta_c} := t_1 \cdots t_{i-1} (t_{i+1} - 1) \cdots (t_k - 1),$$

where i satisfies $t_i = c$. Combining Observations 3 and 2 together, we immediately have the following Observation 4.

Observation 4. *Let $c \in s_1 \cdots s_k \cap t_1 \cdots t_k$. Then both $<_{\text{lex}}$ and $<_{\text{col}}$ hold that*

$$s_1 \cdots s_k <_{\text{lex}(\text{col})} t_1 \cdots t_k \iff s_1 \cdots s_k \ominus c <_{\text{lex}(\text{col})} t_1 \cdots t_k \ominus c.$$

After presenting the above preliminaries, the following part of this section focuses on showing that lex and col are just the worst ranking functions for \mathcal{X}_{LEX} and \mathcal{X}_{COL} respectively, and they induce the quantity given in Theorem 1. We present the proof by sequentially arguing the following assertions:

- (I) $E_{\text{COL}}^{\text{col}}(n, k) = \sum_{1 \leq t_1 < \cdots < t_k \leq n} P(t_1 \cdots t_k)$, with $P(t_1 \cdots t_k)$ holding recurrence (1).
- (II) For any $f \in \mathcal{F}$, $E_{\text{COL}}^f(n, k) \leq E_{\text{COL}}^{\text{col}}(n, k)$.
- (III) $E_{\text{LEX}}^{\text{lex}}(n, k) = E_{\text{COL}}^{\text{col}}(n, k)$.
- (IV) For any $f \in \mathcal{F}$, $E_{\text{LEX}}^f(n, k) \leq E_{\text{LEX}}^{\text{lex}}(n, k)$.

2.2. Proof of (I)

We now formally define $P(t_1 \cdots t_k)$ to be the probability that the subset $t_1 \cdots t_k$ in \mathcal{X}_{COL} achieves left-to-right maxima under a random σ and the fixed ranking function col . Precisely, we have the definition that

$$P(t_1 \cdots t_k) := (1/n!) |\{\sigma \in S_n : (t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, \text{col}}\}|.$$

Since the summation $\sum_{1 \leq t_1 < \cdots < t_k \leq n} P(t_1 \cdots t_k)$ traverses all the k -subsets of $[n]$, what we need to argue in the rest of this subsection, is only that in term of its probability meaning, recurrence (1) really holds. We now prove this by induction on n and k .

From recurrence (1), we have $P(t_1) = 1/t_1$ when $k = 1$, and $P(t_1 \cdots t_k) = P(1 \cdots k) = 1$ when $n = k$, so recurrence (1) holds for the basic $(n, 1)$ -THP and (k, k) -THP. In the inductive step, our assumption is that recurrence (1) holds for $(n-1, k-1)$ -THP and $(n-1, k)$ -THP with $1 < k < n$, and our goal is to argue that it holds for (n, k) -THP. Now we carry out the induction by separately considering two cases.

Case 1: $t_k = n$. In this case, the induction is from $(n-1, k-1)$ -THP. To make the exposition non-confusing, for $(n-1, k-1)$ -THP, we use the notations $\hat{\mathcal{X}}$, $\hat{\sigma}$, and $\hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \text{col}}$, corresponding to \mathcal{X} , σ , and $\mathcal{M}_{\text{COL}}^{\sigma, \text{col}}$ respectively.

First, let's consider the k -subset $t_1 \cdots t_k$ with $t_k = n$ in the left-hand side of recurrence (1). It appears in both \mathcal{X}_{COL} and col , which are the same sequence indeed. By the property of

the colex ordering, we know that all the last $\binom{n-1}{k-1}$ k -subsets in \mathcal{X}_{COL} include n , one of which is $t_1 \cdots t_k$; by Observation 1, we know that only the k -subsets in the consecutive subsequence of col starting from $t_1 \cdots t_k$ can have contribution to $t_1 \cdots t_k$ in \mathcal{X}_{COL} . So if $(t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, \text{col}}$, then there must be $1 \leq j \leq k$ such that $\sigma(t_j) = n$. Therefore in case 1, it holds that

$$P(t_1 \cdots t_k) = (1/n!) \sum_{1 \leq j \leq k} |\{\sigma \in S_n : \sigma(t_j) = n, (t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, \text{col}}\}|. \quad (2)$$

Next, for $1 \leq j \leq k$, let's consider the $(k-1)$ -subset corresponding to the current j in the right-hand side of recurrence (1). We observe that $t_1 \cdots t_{j-1}(t_{j+1}-1) \cdots (t_k-1)$ is just

$$t_1 \cdots t_k \ominus t_j.$$

Then according to the assumption that recurrence (1) holds for $(n-1, k-1)$ -THP, we have

$$P(t_1 \cdots t_k \ominus t_j) = (1/(n-1)!)|\{\hat{\sigma} \in S_{n-1} : (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \in \hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \text{col}}\}|. \quad (3)$$

Now substituting their formulas (respectively in (2) and (3)) for $P(t_1 \cdots t_k)$ and $P(t_1 \cdots t_k \ominus t_j)$ in recurrence (1), we find that the holding of recurrence (1) for (n, k) -THP is implied by the following equation for $1 \leq j \leq k$:

$$|\{\sigma \in S_n : \sigma(t_j) = n, (t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, \text{col}}\}| = |\{\hat{\sigma} \in S_{n-1} : (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \in \hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \text{col}}\}|. \quad (4)$$

To prove equation (4), we note that for $1 \leq j \leq k$, there is a bijection $\Delta_j : \{\sigma \in S_n : \sigma(t_j) = n\} \rightarrow S_{n-1}$ defined by

$$\sigma \mapsto \sigma(1) \cdots \sigma(t_j-1) \sigma(t_j+1) \cdots \sigma(n).$$

In other words, $\Delta_j(\sigma)$ is obtained by removing $\sigma(t_j) = n$, and shifting the sub-word $\sigma(t_j+1) \cdots \sigma(n)$ left (if any). Let $(\sigma, \hat{\sigma}) \in \Delta_j$, and $r_1 \cdots r_k$ be a k -subset of $[n]$ with some $r_i = t_j$. Then we can see bijection Δ_j holds that

$$\begin{aligned} (r_1 \cdots r_k)^\sigma &= (r_1 \cdots r_{i-1} t_j r_{i+1} \cdots r_k)^\sigma \\ &= \{\sigma(r_1), \dots, \sigma(r_{i-1}), \sigma(r_{i+1}), \dots, \sigma(r_k)\} \cup \{\sigma(t_j)\} \\ &= \{\hat{\sigma}(r_1), \dots, \hat{\sigma}(r_{i-1}), \hat{\sigma}(r_{i+1}-1), \dots, \hat{\sigma}(r_k-1)\} \cup \{n\} \\ &= (r_1 \cdots r_{i-1} (r_{i+1}-1) \cdots (r_k-1))^{\hat{\sigma}} \cup \{n\} \\ &= (r_1 \cdots r_k \ominus t_j)^{\hat{\sigma}} \cup \{n\} \end{aligned} \quad (5)$$

With bijection Δ_j , to prove equation (4), what we need to argue is only the following assertion for $1 \leq j \leq k$ and $(\sigma, \hat{\sigma}) \in \Delta_j$:

$$(t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, \text{col}} \iff (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \in \hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \text{col}}. \quad (6)$$

To prove assertion (6), we note that for $1 \leq j \leq k$, there is a bijection $\Lambda_j : \{r_1 \cdots r_k \in \mathcal{X} : t_j \in r_1 \cdots r_k\} \rightarrow \hat{\mathcal{X}}$ defined by

$$r_1 \cdots r_k \mapsto r_1 \cdots r_k \ominus t_j.$$

The bijectiveness of Λ_j is based on Observation 3: In terms of $<_{\text{lex}}$ or $<_{\text{col}}$, Λ_j is order-preserving. So we immediately see that the mapping rule $r_1 \cdots r_k \mapsto r_1 \cdots r_k \ominus t_j$ also establish the bijection

$$\Lambda'_j : \{s_1 \cdots s_k \in \mathcal{X} : t_j \in s_1 \cdots s_k, s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k\} \rightarrow \{r_1 \cdots r_{k-1} \in \hat{\mathcal{X}} : r_1 \cdots r_{k-1} <_{\text{col}} t_1 \cdots t_k \ominus t_j\}.$$

To illustrate bijection Λ'_j , let's consider (6, 3)-THP, and suppose that $t_1 t_2 t_3 = 236$ and the current $j = 2$ (i.e., $t_j = 3$). Then, bijection Λ'_j is as follows:

$$\begin{array}{cccccccccccc} 123 & (124) & 134 & 234 & (125) & 135 & 235 & (145) & (245) & 345 & (126) & 136 & (236) \\ \updownarrow & & \updownarrow & \updownarrow & & \updownarrow & \updownarrow & & & \updownarrow & & \updownarrow & \\ 12 & & 13 & 23 & & 14 & 24 & & & 34 & & 15 & (25) \end{array}$$

With bijection Δ_j and Λ'_j , we are ready for proving assertion (6). Combining equation (5) and Observation 2, we have the following assertion for $1 \leq j \leq k$, $(\sigma, \hat{\sigma}) \in \Delta_j$, and $(s_1 \cdots s_k, s_1 \cdots s_k \ominus t_j) \in \Lambda'_j$:

$$\begin{aligned} (s_1 \cdots s_k)^\sigma <_{\text{col}} (t_1 \cdots t_k)^\sigma &\iff (s_1 \cdots s_k \ominus t_j)^{\hat{\sigma}} \cup \{n\} <_{\text{col}} (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \cup \{n\} && \text{(by Eq. (5))} \\ &\iff (s_1 \cdots s_k \ominus t_j)^{\hat{\sigma}} <_{\text{col}} (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} && \text{(by Obs. 2).} \end{aligned} \tag{7}$$

Since bijection Λ'_j establishes an injection from $\{r_1 \cdots r_{k-1} \in \hat{\mathcal{X}} : r_1 \cdots r_{k-1} <_{\text{col}} t_1 \cdots t_k \ominus t_j\}$ to $\{s_1 \cdots s_k \in \mathcal{X} : s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k\}$, we find that assertion (7) directly implies the \Rightarrow direction of assertion (6). How about its \Leftarrow direction? Consider such $s_1 \cdots s_k$ with $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$ and $t_j \notin s_1 \cdots s_k$. We find that $n \notin (s_1 \cdots s_k)^\sigma$, and thus it inherently holds that $(s_1 \cdots s_k)^\sigma <_{\text{col}} (t_1 \cdots t_k)^\sigma$. Therefore, the \Leftarrow direction of assertion (6) also holds.

With assertion (6) being proved, we accomplish the argument for case 1.

Case 2: $t_k < n$. In this case, we know that recurrence (1) has already held for (t_k, k) -THP, by the conclusion of case 1. What we need to argue is only that $P(t_1 \cdots t_k)$ remains unchanged if we extend the problem from (t_k, k) -THP to (n, k) -THP. We argue this also by induction. That is, under the assumption that $P(t_1 \cdots t_k)$ remains unchanged from (t_k, k) -THP to $(n-1, k)$ -THP, our goal is to argue that it remains unchanged from $(n-1, k)$ -THP to (n, k) -THP.

By the property of the colex ordering, we know that every $s_1 \cdots s_k$ with $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$ remains unchanged if we extend the problem from $(n-1, k)$ -THP to (n, k) -THP, and it doesn't include n . So what we need to argue for case 2 is implied by the following equation for $1 \leq d \leq n$:

$$|\{\sigma \in S_n : \sigma(n) = d, (t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, \text{col}}\}| = |\{\tilde{\sigma} \in S_{n-1} : (t_1 \cdots t_k)^{\tilde{\sigma}} \in \tilde{\mathcal{M}}_{\text{COL}}^{\tilde{\sigma}, \text{col}}\}|, \tag{8}$$

where $\tilde{\mathcal{M}}_{\text{COL}}^{\tilde{\sigma}, \text{col}}$ is understood to be the notation of $(n-1, k)$ -THP, corresponding to $\mathcal{M}_{\text{COL}}^{\sigma, \text{col}}$.

To prove equation (8), we note that for $1 \leq d \leq n$, there is a bijection $\Upsilon_d : \{\sigma \in S_n : \sigma(n) = d\} \rightarrow S_{n-1}$ defined by

$$\sigma \mapsto \delta_d(\sigma(1))\delta_d(\sigma(2)) \cdots \delta_d(\sigma(n-1)).$$

In other words, $\mathcal{Y}_d(\sigma)$ is obtained by removing $\sigma(n) = d$ and decrementing all the elements larger than d . Now, letting $(\sigma, \tilde{\sigma}) \in \mathcal{Y}_d$, and $r_1 \cdots r_k$ be a k -subset of $[n]$ with $r_k < n$ (i.e., $d \notin (r_1 \cdots r_k)^\sigma$), we can see bijection \mathcal{Y}_d holds that

$$\begin{aligned} ((r_1 \cdots r_k)^\sigma)^{\delta_d} &= \{\delta_d(\sigma(r_1)), \delta_d(\sigma(r_2)), \dots, \delta_d(\sigma(r_k))\} \\ &= \{\tilde{\sigma}(r_1), \tilde{\sigma}(r_2), \dots, \tilde{\sigma}(r_k)\} \\ &= (r_1 \cdots r_k)^{\tilde{\sigma}}. \end{aligned} \tag{9}$$

Then, combining equation (9) and Observation 3, we have the following assertion for $1 \leq d \leq n$, $(\sigma, \tilde{\sigma}) \in \mathcal{Y}_d$, and any $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$:

$$\begin{aligned} (s_1 \cdots s_k)^\sigma <_{\text{col}} (t_1 \cdots t_k)^\sigma &\iff ((s_1 \cdots s_k)^\sigma)^{\delta_d} <_{\text{col}} ((t_1 \cdots t_k)^\sigma)^{\delta_d} && \text{(by Obs. 3)} \\ &\iff (s_1 \cdots s_k)^{\tilde{\sigma}} <_{\text{col}} (t_1 \cdots t_k)^{\tilde{\sigma}} && \text{(by Eq. (9)).} \end{aligned}$$

That directly implies equation (8).

With the argument for case 2 being completed, we accomplish the proof of (I).

2.3. Proof of (II)

In this subsection, we continue to use the definition of $P(t_1 \cdots t_k)$ as in Section 2.2, i.e., it is the probability that the subset $t_1 \cdots t_k$ in \mathcal{X}_{COL} achieves left-to-right maxima under a random σ and the fixed col, and it holds recurrence (1) by the conclusion of Section 2.2. When the ranking function f is general, we use $R_{\text{COL}}^f(t_1 \cdots t_k)$ to denote the probability that the subset $t_1 \cdots t_k$ in \mathcal{X}_{COL} achieves left-to-right maxima under a random σ . Precisely, we have the definition that

$$R_{\text{COL}}^f(t_1 \cdots t_k) := (1/n!) |\{\sigma \in S_n : (t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, f}\}|.$$

(That is, under this new definition, $P(t_1 \cdots t_k)$ is just $R_{\text{COL}}^{\text{col}}(t_1 \cdots t_k)$.) Then we can see that the inequality $E_{\text{COL}}^f(n, k) \leq E_{\text{COL}}^{\text{col}}(n, k)$ we want to prove is definitely implied by the following stronger inequality for any $t_1 \cdots t_k \in \mathcal{X}$:

$$R_{\text{COL}}^f(t_1 \cdots t_k) \leq P(t_1 \cdots t_k). \tag{10}$$

So the remainder of this subsection rests on proving inequality (10), by using an induction similar to the one in the proof of (I).

We first note that inequality (10) holds for the basic $(n, 1)$ -THP and (k, k) -THP with equalities. In the inductive step, our goal is to argue that it holds for (n, k) -THP under the assumption that it holds for $(n-1, k-1)$ -THP and $(n-1, k)$ -THP with $1 < k < n$. We achieve this goal also by separately considering two cases.

Case 1: $t_k = n$. In this case, the induction is also from $(n-1, k-1)$ -THP. To make the notations distinguished, we still add $\hat{\cdot}$ to the relevant symbols of $(n-1, k-1)$ -THP.

Letting $f = \langle Y_1, Y_2, \dots, Y_{\binom{n}{k}} \rangle$, we now consider two subcases for case 1.

Case 1.1: There is $1 \leq c \leq n$ such that $c \in \bigcap_{\binom{n-1}{k} < l \leq \binom{n}{k}} Y_l$. In other words, in this subcase, c appears only in the last $\binom{n-1}{k-1}$ k -subsets of f . Let's construct $\hat{f} = \langle \hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_{\binom{n-1}{k-1}} \rangle$ from f with the rule

$$\hat{Y}_l := Y_{\binom{n-1}{k}+l} \ominus c$$

for $1 \leq l \leq \binom{n-1}{k-1}$. For example, if $f = \langle \dots, 234, 125, 245, 124, 123, 235 \rangle$ (i.e., $c = 2$) in $(5, 3)$ -THP, then $\hat{f} = \langle 23, 14, 34, 13, 12, 24 \rangle$. This mapping rule establishes a surjection

$$\Phi_c : \{f \in \mathcal{F} : c \in \bigcap_{\binom{n-1}{k} < l \leq \binom{n}{k}} Y_l\} \rightarrow \hat{\mathcal{F}},$$

since similar to bijection Λ_j , the bijection from $\{Y_{\binom{n-1}{k}+1}, Y_{\binom{n-1}{k}+2}, \dots, Y_{\binom{n}{k}}\}$ to $\{\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_{\binom{n-1}{k-1}}\}$ is order-preserving in terms of $<_f$ and $<_{\hat{f}}$. So we have the following assertion for $c \in r_1 \cdots r_k \cap s_1 \cdots s_k$ and $(f, \hat{f}) \in \Phi_c$:

$$r_1 \cdots r_k <_f s_1 \cdots s_k \iff r_1 \cdots r_k \ominus c <_{\hat{f}} s_1 \cdots s_k \ominus c. \quad (11)$$

In $(n-1, k-1)$ -THP, for general $\hat{f} \in \hat{\mathcal{F}}$ and $1 \leq j \leq k$, it is naturally understood that

$$\hat{R}_{\text{COL}}^{\hat{f}}(t_1 \cdots t_k \ominus t_j) = (1/(n-1)!) |\{\hat{\sigma} \in S_{n-1} : (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \in \hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \hat{f}}\}|, \quad (12)$$

and according to the assumption on $(n-1, k-1)$ -THP, it holds that

$$\hat{R}_{\text{COL}}^{\hat{f}}(t_1 \cdots t_k \ominus t_j) \leq P(t_1 \cdots t_k \ominus t_j). \quad (13)$$

Now we claim that inequality (10) is implied by the following equation for $(f, \hat{f}) \in \Phi_c$:

$$R_{\text{COL}}^f(t_1 \cdots t_k) = (1/n) \sum_{1 \leq j \leq k} \hat{R}_{\text{COL}}^{\hat{f}}(t_1 \cdots t_k \ominus t_j). \quad (14)$$

That is because if equation (14) is true, then we immediately have

$$\begin{aligned} R_{\text{COL}}^f(t_1 \cdots t_k) &\leq (1/n) \sum_{1 \leq j \leq k} P(t_1 \cdots t_k \ominus t_j) && \text{(by inequality (13))} \\ &= P(t_1 \cdots t_k) && \text{(by recurrence (1)).} \end{aligned}$$

All that remains for case 1.1 is to prove equation (14). In fact, it can be proved along the way to argue case 1 of Section 2.2. Since $t_k = n$, $t_1 \cdots t_k$ is one of the last $\binom{n-1}{k-1}$ k -subsets in \mathcal{X}_{COL} ; so if $(t_1 \cdots t_k)^{\sigma} \in \mathcal{M}_{\text{COL}}^{\sigma, f}$, then there must be $1 \leq j \leq k$ such that $\sigma(t_j) = c$. Therefore in case 1.1, it holds that

$$R_{\text{COL}}^f(t_1 \cdots t_k) = (1/n!) \sum_{1 \leq j \leq k} |\{\sigma \in S_n : \sigma(t_j) = c, (t_1 \cdots t_k)^{\sigma} \in \mathcal{M}_{\text{COL}}^{\sigma, f}\}|. \quad (15)$$

Substituting their formulas (respectively in (15) and (12)) for $R_{\text{COL}}^f(t_1 \cdots t_k)$ and $\hat{R}_{\text{COL}}^{\hat{f}}(t_1 \cdots t_k \ominus t_j)$ in equation (14), we find that the holding of equation (14) is implied by the following equation for $1 \leq j \leq k$ and $(f, \hat{f}) \in \Phi_c$:

$$|\{\sigma \in S_n : \sigma(t_j) = c, (t_1 \cdots t_k)^{\sigma} \in \mathcal{M}_{\text{COL}}^{\sigma, f}\}| = |\{\hat{\sigma} \in S_{n-1} : (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \in \hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \hat{f}}\}|. \quad (16)$$

Now we can check the similarity between (16) and (4); nevertheless, a self-contained argument for equation (16) is necessary, though it looks somewhat redundant.

To prove equation (16), we first need a bijection $\Psi_{j,c} : \{\sigma \in S_n : \sigma(t_j) = c\} \rightarrow S_{n-1}$ defined by

$$\sigma \mapsto \delta_c(\sigma(1)) \cdots \delta_c(\sigma(t_j - 1)) \delta_c(\sigma(t_j + 1)) \cdots \delta_c(n),$$

which can be viewed as a generalization of bijection Δ_j from the special case $c = n$ to the general case. Letting $(\sigma, \hat{\sigma}) \in \Psi_{j,c}$, and $r_1 \cdots r_k$ be a k -subset with some $r_i = t_j$, we can see bijection $\Psi_{j,c}$ holds that

$$\begin{aligned} (r_1 \cdots r_k)^\sigma \ominus c &= ((r_1 \cdots r_{i-1} t_j r_{i+1} \cdots r_k)^\sigma \setminus \{\sigma(t_j)\})^{\delta_c} \\ &= \{\delta_c(\sigma(r_1)), \dots, \delta_c(\sigma(r_{i-1})), \delta_c(\sigma(r_{i+1})), \dots, \delta_c(\sigma(r_k))\} \\ &= \{\hat{\sigma}(r_1), \dots, \hat{\sigma}(r_{i-1}), \hat{\sigma}(r_{i+1} - 1), \dots, \hat{\sigma}(r_k - 1)\} \\ &= (r_1 \cdots r_{i-1} (r_{i+1} - 1) \cdots (r_k - 1))^{\hat{\sigma}} \\ &= (r_1 \cdots r_k \ominus t_j)^{\hat{\sigma}}. \end{aligned} \tag{17}$$

Another bijection we need is the bijection

$$\Lambda'_j : \{s_1 \cdots s_k \in \mathcal{X} : t_j \in s_1 \cdots s_k, s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k\} \rightarrow \{r_1 \cdots r_{k-1} \in \hat{\mathcal{X}} : r_1 \cdots r_{k-1} <_{\text{col}} t_1 \cdots t_k \ominus t_j\},$$

which is the same as the one in Section 2.2.

With bijections $\Psi_{j,c}$ and Λ'_j , as well as surjection Φ_c , we are ready for proving equation (16). Combining equation (17) and assertion (11), we have the following assertion for $1 \leq j \leq k$, $(f, \hat{f}) \in \Phi_c$, $(\sigma, \hat{\sigma}) \in \Psi_{j,c}$, and $(s_1 \cdots s_k, s_1 \cdots s_k \ominus t_j) \in \Lambda'_j$:

$$\begin{aligned} (s_1 \cdots s_k)^\sigma <_f (t_1 \cdots t_k)^\sigma &\iff (s_1 \cdots s_k)^\sigma \ominus c <_{\hat{f}} (t_1 \cdots t_k)^\sigma \ominus c && \text{(by assertion (11))} \\ &\iff (s_1 \cdots s_k \ominus t_j)^{\hat{\sigma}} <_{\hat{f}} (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} && \text{(by Eq. (17)).} \end{aligned} \tag{18}$$

Assertion (18) directly implies

$$(t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, f} \implies (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \in \hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \hat{f}},$$

since $|\{s_1 \cdots s_k \in \mathcal{X} : s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k\}| > |\{r_1 \cdots r_{k-1} \in \hat{\mathcal{X}} : r_1 \cdots r_{k-1} <_{\text{col}} t_1 \cdots t_k \ominus t_j\}|$. How about such $s_1 \cdots s_k$ with $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$ and $t_j \notin s_1 \cdots s_k$? We can find it inherently holds that $(s_1 \cdots s_k)^\sigma <_f (t_1 \cdots t_k)^\sigma$, because $c \notin (s_1 \cdots s_k)^\sigma$ and $c \in (t_1 \cdots t_k)^\sigma$. Therefore, it also holds that

$$(t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, f} \Leftarrow (t_1 \cdots t_k \ominus t_j)^{\hat{\sigma}} \in \hat{\mathcal{M}}_{\text{COL}}^{\hat{\sigma}, \hat{f}}.$$

This completes the proof of equation (16), and completes the argument for case 1.1 as well.

Case 1.2: $\bigcap_{\binom{n-1}{k} < l \leq \binom{n}{k}} Y_l = \emptyset$. Let v be the smallest number such that $\bigcap_{v \leq l \leq \binom{n}{k}} Y_l \neq \emptyset$, and for convenience, let $\mathcal{X}_{\text{COL}} = \langle C_1, C_2, \dots, C_{\binom{n}{k}} \rangle$ and $t_1 \cdots t_k = C_u$ with $\binom{n-1}{k} < u \leq \binom{n}{k}$.

We divide the representation of $R_{\text{COL}}^f(C_u)$ in terms of contributions into two parts:

$$R_{\text{COL}}^f(C_u) = (1/n!) \sum_{u \leq l < v} T_{\text{COL}}^f(C_u, Y_l) + (1/n!) \sum_{v \leq l \leq \binom{n}{k}} T_{\text{COL}}^f(C_u, Y_l).$$

First consider $(1/n!) \sum_{u \leq l < v} T_{\text{COL}}^f(C_u, Y_l)$. It is null if $v \leq u$; otherwise we claim that it is equal to 0. Why? From $\bigcap_{u \leq i \leq \binom{n}{k}} C_i \neq \emptyset$ (all including n), we know that $\bigcap_{u \leq i \leq \binom{n}{k}} C_i^\sigma \neq \emptyset$ for any $\sigma \in S_n$. However, for any l with $u \leq l < v$, $\bigcap_{l \leq i \leq \binom{n}{k}} Y_i = \emptyset$, so anyway we have $\{Y_l, Y_{l+1}, \dots, Y_{\binom{n}{k}}\} \not\subseteq \{C_u^\sigma, C_{u+1}^\sigma, \dots, C_{\binom{n}{k}}^\sigma\}$. That is, when $C_u^\sigma = Y_l$ with $u \leq l < v$, C_u^σ cannot be a left-to-right maximum in $\mathcal{X}_{\text{COL}}^\sigma$, under f .

Next, consider $(1/n!) \sum_{v \leq l \leq \binom{n}{k}} T_{\text{COL}}^f(C_u, Y_l)$. We claim that it cannot exceed $P(C_u)$. Indeed, let's construct a valid ranking function $g = \langle Z_1, Z_2, \dots, Z_{\binom{n}{k}} \rangle$ from f such that it satisfies

$$\langle Z_v, Z_{v+1}, \dots, Z_{\binom{n}{k}} \rangle = \langle Y_v, Y_{v+1}, \dots, Y_{\binom{n}{k}} \rangle \text{ and } \bigcap_{\binom{n-1}{k} < l \leq \binom{n}{k}} Z_l \neq \emptyset.$$

That is, g satisfies the condition of case 1.1, and keeps the last $\binom{n}{k} - v + 1$ k -subsets unchanged. Such construction is possible because $\bigcap_{v \leq l \leq \binom{n}{k}} Y_l \neq \emptyset$. Then we have

$$\begin{aligned} (1/n!) \sum_{v \leq l \leq \binom{n}{k}} T_{\text{COL}}^f(C_u, Y_l) &= (1/n!) \sum_{v \leq l \leq \binom{n}{k}} T_{\text{COL}}^g(C_u, Z_l) && \text{(by Obs. 1(ii))} \\ &\leq (1/n!) \sum_{u \leq l \leq \binom{n}{k}} T_{\text{COL}}^g(C_u, Z_l) && \text{(being = if } v \leq u) \\ &= R_{\text{COL}}^g(C_u) \\ &\leq P(C_u) && \text{(from case 1.1).} \end{aligned}$$

Summarizing these two parts results in the inequality $R_{\text{COL}}^f(C_u) \leq P(C_u)$ (i.e., inequality (10)).

Case 2: $t_k < n$. In this case, we know that inequality (10) has already held for (t_k, k) -THP, by the conclusion of case 1. We still use induction to prove that it holds for (n, k) -THP. That is, under the assumption that inequality (10) holds for $(n-1, k)$ -THP, we argue its holding for (n, k) -THP. To distinguish the notations of $(n-1, k)$ -THP from (n, k) -THP, we add \sim to the relevant symbols of $(n-1, k)$ -THP.

Still let $f = \langle Y_1, Y_2, \dots, Y_{\binom{n}{k}} \rangle$. For $1 \leq d \leq n$, we note that there is a surjection $\Omega_d : \mathcal{F} \rightarrow \tilde{\mathcal{F}}$, in which $\tilde{f} = \langle \tilde{Y}_1, \tilde{Y}_2, \dots, \tilde{Y}_{\binom{n-1}{k}} \rangle$ is obtained from f as follows: remove all the k -subsets including d and conduct δ_d to the remaining $\binom{n-1}{k}$ k -subsets. Precisely, letting $\langle Y_{i_1}, Y_{i_2}, \dots, Y_{i_{\binom{n-1}{k}}} \rangle$ with $i_1 < i_2 < \dots < i_{\binom{n-1}{k}}$ be the subsequence of f in which no k -subset includes d , we construct \tilde{f} by the rule

$$\tilde{Y}_j := (Y_{i_j})^{\delta_d}$$

for $1 \leq j \leq \binom{n-1}{k}$. For instance, in Ω_3 of $(5, 2)$ -THP,

$$\langle \underline{15}, \underline{34}, \underline{35}, \underline{23}, 25, 24, \underline{13}, 12, 45, 14 \rangle \mapsto \langle 14, 24, 23, 12, 34, 13 \rangle,$$

where the underlined subsets are neglected. By the mapping rule of Ω_d , we have the following assertion for $d \notin r_1 \cdots r_k \cup s_1 \cdots s_k$ and $(f, \tilde{f}) \in \Omega_d$:

$$r_1 \cdots r_k <_f s_1 \cdots s_k \iff (r_1 \cdots r_k)^{\delta_d} <_{\tilde{f}} (s_1 \cdots s_k)^{\delta_d}. \quad (19)$$

Now we claim that the holding of inequality (10) for case 2 is implied by the following equation for $1 \leq d \leq n$ and $(f, \tilde{f}) \in \Omega_d$:

$$\left| \{ \sigma \in S_n : \sigma(n) = d, (t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, f} \} \right| = \left| \{ \tilde{\sigma} \in S_{n-1} : (t_1 \cdots t_k)^{\tilde{\sigma}} \in \tilde{\mathcal{M}}_{\text{COL}}^{\tilde{\sigma}, \tilde{f}} \} \right|, \quad (20)$$

which is quite similar to equation (8). Indeed, if equation (20) is true, we immediately have

$$\begin{aligned} R_{\text{COL}}^f(t_1 \cdots t_k) &= (1/n!) \sum_{1 \leq d \leq n} (\text{the left-hand side of (20)}) \\ &\leq (1/(n-1)!)\left| \{ \tilde{\sigma} \in S_{n-1} : (t_1 \cdots t_k)^{\tilde{\sigma}} \in \tilde{\mathcal{M}}_{\text{COL}}^{\tilde{\sigma}, \text{col}} \} \right| \quad (\text{by the assumption}) \\ &= \tilde{R}_{\text{COL}}^{\text{col}}(t_1 \cdots t_k) \\ &= P(t_1 \cdots t_k) \quad (\text{from case 2 of Sec. 2.2}). \end{aligned}$$

To prove equation (20), we can use the same bijection $\mathcal{Y}_d : \{ \sigma \in S_n : \sigma(n) = d \} \rightarrow S_{n-1}$ for $1 \leq d \leq n$ as the one in Section 2.2. Recall that for $(\sigma, \tilde{\sigma}) \in \mathcal{Y}_d$ and a k -subset $r_1 \cdots r_k$ with $r_k < n$, bijection \mathcal{Y}_d holds equation (9), i.e., $((r_1 \cdots r_k)^\sigma)^{\delta_d} = (r_1 \cdots r_k)^{\tilde{\sigma}}$, as being restated. Then, combining this equation and assertion (19), we have the following assertion for $1 \leq d \leq n$, $(\sigma, \tilde{\sigma}) \in \mathcal{Y}_d$, $(f, \tilde{f}) \in \Omega_d$, and any $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$:

$$\begin{aligned} (s_1 \cdots s_k)^\sigma <_f (t_1 \cdots t_k)^\sigma &\iff ((s_1 \cdots s_k)^\sigma)^{\delta_d} <_{\tilde{f}} ((t_1 \cdots t_k)^\sigma)^{\delta_d} \quad (\text{by assertion (19)}) \\ &\iff (s_1 \cdots s_k)^{\tilde{\sigma}} <_{\tilde{f}} (t_1 \cdots t_k)^{\tilde{\sigma}} \quad (\text{by Eq. (9)}). \end{aligned} \quad (21)$$

To explain it further, we should see that when $\sigma(n) = d$, every $(s_1 \cdots s_k)^\sigma$ with $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$ does't include d , so the k -subsets in f including d are as ineffective as nothing.

Since assertion (21) directly implies equation (20), we accomplish the argument for case 2, and so accomplish the proof of (II) as well.

2.4. Proof of (III)

Obviously, the equation $E_{\text{LEX}}^{\text{lex}}(n, k) = E_{\text{COL}}^{\text{col}}(n, k)$ we want to prove is implied by the following stronger assertion for any $\sigma \in S_n$ and any $t_1 \cdots t_k \in \mathcal{X}$:

$$(t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{LEX}}^{\sigma, \text{lex}} \iff (t_1 \cdots t_k)^\sigma \in \mathcal{M}_{\text{COL}}^{\sigma, \text{col}}. \quad (22)$$

We now prove this assertion by contradiction.

\Leftarrow : We have the assumption that $(s_1 \cdots s_k)^\sigma <_{\text{col}} (t_1 \cdots t_k)^\sigma$ for all $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$. We now suppose for the purpose of contradiction that there is a $r_1 \cdots r_k$ with $r_1 \cdots r_k <_{\text{lex}} t_1 \cdots t_k$ and $(t_1 \cdots t_k)^\sigma <_{\text{lex}} (r_1 \cdots r_k)^\sigma$. Our objective is to find a $s_1 \cdots s_k$ with $s_1 \cdots s_k <_{\text{col}} t_1 \cdots t_k$ and $(t_1 \cdots t_k)^\sigma <_{\text{col}} (s_1 \cdots s_k)^\sigma$, which contradicts the assumption.

To do this, we let $c_1 \cdots c_l = r_1 \cdots r_k \cap t_1 \cdots t_k$, and let $r'_1 \cdots r'_j = r_1 \cdots r_k \setminus c_1 \cdots c_l$, $t'_1 \cdots t'_j = t_1 \cdots t_k \setminus c_1 \cdots c_l$. Clearly, $j \geq 1$ and $j + l = k$. Then, by the property of lexical ordering, we have that $r'_1 < t'_1$, and $\min_{1 \leq i \leq j} \sigma(t'_i) < \min_{1 \leq i \leq j} \sigma(r'_i)$. Now, let $(t'_1 \cdots t'_j)^\sigma = \sigma(t''_1) \sigma(t''_2) \cdots \sigma(t''_j)$ (with $\sigma(t''_1) < \sigma(t''_2) < \cdots < \sigma(t''_j)$). Then, we claim that the k -subset $\{r'_1, t''_2, \dots, t''_j, c_1, \dots, c_l\}$ is just what we want to find.

In fact, since $r'_1 < t'_1$ and $\{t''_2, \dots, t''_j\} \subset t'_1 \cdots t'_j$, we have

$$\{r'_1, t''_2, \dots, t''_j, c_1, \dots, c_l\} <_{\text{lex}(\text{col})} \{t'_1, t'_2, \dots, t'_j, c_1, \dots, c_l\} = t_1 \cdots t_k. \quad (23)$$

In inequality (23), as well as (24),(25), and (26) hereafter, the reason why both $<_{\text{lex}}$ and $<_{\text{col}}$ hold is that there is only one distinct element between those two k -subsets.

On the other hand, since $\sigma(t''_1) = \min_{1 \leq i \leq j} \sigma(t''_i) < \min_{1 \leq i \leq j} \sigma(r'_i) \leq \sigma(r'_1)$, we have

$$(t_1 \cdots t_k)^\sigma = \{t'_1, t'_2, \dots, t'_j, c_1, \dots, c_l\}^\sigma <_{\text{lex}(\text{col})} \{r'_1, t''_2, \dots, t''_j, c_1, \dots, c_l\}^\sigma. \quad (24)$$

Inequalities (23) and (24) together form a contradiction to the assumption of the \Leftarrow direction.

\Rightarrow : This time we have the assumption that $(s_1 \cdots s_k)^\sigma <_{\text{lex}} (t_1 \cdots t_k)^\sigma$ for all $s_1 \cdots s_k <_{\text{lex}} t_1 \cdots t_k$. We now suppose for the purpose of contradiction that there is a $r_1 \cdots r_k$ with $r_1 \cdots r_k <_{\text{col}} t_1 \cdots t_k$ and $(t_1 \cdots t_k)^\sigma <_{\text{col}} (r_1 \cdots r_k)^\sigma$. We still let $c_1 \cdots c_l = r_1 \cdots r_k \cap t_1 \cdots t_k$, and let $r'_1 \cdots r'_j = r_1 \cdots r_k \setminus c_1 \cdots c_l$, $t'_1 \cdots t'_j = t_1 \cdots t_k \setminus c_1 \cdots c_l$. Then, by the property of colex ordering, we have that $r'_j < t'_j$, and $\max_{1 \leq i \leq j} \sigma(t''_i) < \max_{1 \leq i \leq j} \sigma(r'_i)$. Now, let $(r'_1 \cdots r'_j)^\sigma = \sigma(r''_1) \sigma(r''_2) \cdots \sigma(r''_j)$ (with $\sigma(r''_1) < \sigma(r''_2) < \cdots < \sigma(r''_j)$). Then in this time, we rely on the k -subset $\{t'_1, \dots, t'_{j-1}, r''_j, c_1, \dots, c_l\}$ to derive contradiction.

Indeed, noting that $r''_j < t'_j$ (because $r''_j \in r'_1 \cdots r'_j$ and $r'_j < t'_j$), we have

$$\{t'_1, \dots, t'_{j-1}, r''_j, c_1, \dots, c_l\} <_{\text{col}(\text{lex})} \{t'_1, \dots, t'_j, c_1, \dots, c_l\} = t_1 \cdots t_k. \quad (25)$$

On the other hand, since $\max_{1 \leq i \leq j} \sigma(t'_i) < \max_{1 \leq i \leq j} \sigma(r'_i) = \sigma(r''_j)$, we have

$$(t_1 \cdots t_k)^\sigma = \{t'_1, \dots, t'_j, c_1, \dots, c_l\}^\sigma <_{\text{col}(\text{lex})} \{t'_1, \dots, t'_{j-1}, r''_j, c_1, \dots, c_l\}^\sigma. \quad (26)$$

Inequalities (25) and (26) together form a contradiction to the assumption of the \Rightarrow direction.

2.5. Proof of (IV)

We start by presenting two properties of $E_{ord}^f(n, k)$ for general ord and f . Then the assertion we want to prove here can be viewed as a corollary of assertion (II) plus these two properties.

Let $\mathcal{X}_{ord} = \langle X_1, X_2, \dots, X_{\binom{n}{k}} \rangle$ and $f = \langle Y_1, Y_2, \dots, Y_{\binom{n}{k}} \rangle$. We define a series of “complements” as follows:

- $\overline{X_i} := [n] \setminus X_i$;
- $\overline{\mathcal{X}_{ord}} := \langle \overline{X_1}, \overline{X_2}, \dots, \overline{X_{\binom{n}{k}}} \rangle$, the ordering of all the $(n-k)$ -subsets of $[n]$ in the way of \overline{ord} ;
- $\overline{f} := \langle \overline{Y_1}, \overline{Y_2}, \dots, \overline{Y_{\binom{n}{k}}} \rangle$, a ranking function of $(n, n-k)$ -THP satisfying that $\overline{f}(\overline{X_i}) = f(X_i)$ for $1 \leq i \leq \binom{n}{k}$.

We claim that $E_{ord}^f(n, k)$ holds the following complementary-equivalence property:

$$E_{ord}^f(n, k) = E_{ord}^{\bar{f}}(n, n - k). \quad (27)$$

Indeed, for any $\sigma \in S_n$, we observe that $\overline{X_i^\sigma} = [n] \setminus X_i^\sigma = [n]^\sigma \setminus X_i^\sigma = ([n] \setminus X_i)^\sigma = \overline{X_i}^\sigma$, so we have $\langle f(X_i^\sigma) \rangle_{1 \leq i \leq \binom{n}{k}} = \langle \bar{f}(\overline{X_i^\sigma}) \rangle_{1 \leq i \leq \binom{n}{k}} = \langle \bar{f}(\overline{X_i}^\sigma) \rangle_{1 \leq i \leq \binom{n}{k}}$, from which equation (27) follows.

Let $\pi, \omega \in S_n$. We define ord^π and f^ω as follows:

- $\mathcal{X}_{ord^\pi} := \langle X_1^\pi, X_2^\pi, \dots, X_{\binom{n}{k}}^\pi \rangle$;
- $f^\omega := \langle Y_1^\omega, Y_2^\omega, \dots, Y_{\binom{n}{k}}^\omega \rangle$, a ranking function of (n, k) -THP satisfying that $f^\omega(X_i^\omega) = f(X_i)$ for $1 \leq i \leq \binom{n}{k}$.

We claim that $E_{ord}^f(n, k)$ holds the following permuting-equivalence property:

$$E_{ord}^f(n, k) = E_{ord^\pi}^{f^\omega}(n, k). \quad (28)$$

Indeed, for any $\sigma \in S_n$, we observe that $f(X_i^\sigma) = f((X_i^\pi)^{\pi^{-1} \cdot \sigma}) = f^\omega((X_i^\pi)^{\pi^{-1} \cdot \sigma \cdot \omega})$. That is, whenever there is a $\sigma \in S_n$ incurring the sequence $\langle f(X_i^\sigma) \rangle_{1 \leq i \leq \binom{n}{k}}$, there is only one $\sigma' = \pi^{-1} \cdot \sigma \cdot \omega$ such that $\langle f^\omega((X_i^\pi)^{\sigma'}) \rangle_{1 \leq i \leq \binom{n}{k}} = \langle f(X_i^\sigma) \rangle_{1 \leq i \leq \binom{n}{k}}$. Therefore, equation (28) follows.

It is a common knowledge that the lexical and colex orderings are complementary to each other. If we only use their normal forms, letting $\pi = n(n-1) \cdots 1$, we have that in symbols, $\text{COL} = \overline{\text{LEX}}^\pi$, $\text{col} = \overline{\text{lex}}^\pi$, and vice versa. Therefore for any f , we have

$$\begin{aligned} E_{\text{LEX}}^f(n, k) &= E_{\overline{\text{LEX}}^\pi}^{\bar{f}}(n, n - k) = E_{\text{COL}}^{\bar{f}}(n, n - k) \\ &\leq E_{\text{COL}}^{\text{col}}(n, n - k) = E_{\overline{\text{COL}}^\pi}^{\overline{\text{col}}^\pi}(n, k) = E_{\text{LEX}}^{\text{lex}}(n, k), \end{aligned}$$

where the inequality comes from assertion (II). This completes the proof of assertion (IV).

3. Proof of Corollary 2

Let us first prove the symmetry of $E_{\text{LC}}^*(n, k)$ in k and $n - k$. Indeed, it has been implied by the results of the previous section: By assertions (I)–(IV), we know that $E_{\text{LC}}^*(n, k) = E_{\text{LEX}}^{\text{lex}}(n, k) = E_{\text{COL}}^{\text{col}}(n, k)$, and from Section 2.5, we know that $E_{\text{LEX}}^{\text{lex}}(n, k) = E_{\text{COL}}^{\text{col}}(n, n - k)$; consequently, combining these two results gives the equation $E_{\text{LC}}^*(n, k) = E_{\text{LC}}^*(n, n - k)$.

As a by-product of the above proof, we have established, and given a probabilistic interpretation to, the following identity involving $P(\cdot)$ defined by recurrence (1):

$$P(t_1 \cdots t_k) = P([n] \setminus t_1 \cdots t_k)^\pi, \quad \text{where } \pi = n(n-1) \cdots 1, 1 \leq t_1 < \cdots < t_k \leq n. \quad (29)$$

For example, by applying identity (29) to the subset 245 with $n = 5, 6, 7, 8, \dots$, we have

$$P(245) = P(35) = P(146) = P(1257) = P(12368) = \cdots = \frac{7}{60}.$$

A numerical interpretation of this mysterious identity will lead to an alternative proof of the equation $E_{\text{LC}}^*(n, k) = E_{\text{LC}}^*(n, n - k)$. It is left to readers as a challenging exercise.

Now, let us prove the bounds of $E_{\text{LC}}^*(n, k)$ presented in Corollary 2, directly from the formulas in Theorem 1. We start by claiming that $P(\cdot)$ holds the following inequality:

$$P(s_1 \cdots s_k) \geq P(t_1 \cdots t_k), \quad \text{for } s_1 \leq t_1, s_2 \leq t_2, \dots, s_k \leq t_k. \quad (30)$$

Indeed, inequality (30) can be easily proved by induction. Since $1/s_1 \geq 1/t_1$ in the case $k = 1$, inequality (30) holds for the basic $(n, 1)$ -THP. We now assume that inequality (30) holds for $(n-1, k-1)$ -THP with $1 < k \leq n$. Under this assumption, we observe that

$$\begin{aligned} P(s_1 \cdots s_k) &= (1/s_k) \sum_{1 \leq j \leq k} P(s_1 \cdots s_{j-1} (s_{j+1} - 1) \cdots (s_k - 1)) \\ &\geq (1/t_k) \sum_{1 \leq j \leq k} P(t_1 \cdots t_{j-1} (t_{j+1} - 1) \cdots (t_k - 1)) \\ &= P(t_1 \cdots t_k). \end{aligned}$$

That is, inequality (30) holds for (n, k) -THP.

Then, looking at the right-hand side of recurrence (1), we find that

$$P(t_1 \cdots t_{k-1}) = \max_{1 \leq j \leq k} P(t_1 \cdots t_{j-1} (t_{j+1} - 1) \cdots (t_k - 1)). \quad (31)$$

Therefore, by (31), we have

$$P(t_1 \cdots t_k) \leq \frac{k}{t_k} P(t_1 \cdots t_{k-1}) \leq \frac{k(k-1)}{t_k t_{k-1}} P(t_1 \cdots t_{k-2}) \leq \cdots \leq \frac{k!}{t_k t_{k-1} \cdots t_1},$$

and at the same time, also have

$$P(t_1 \cdots t_k) \geq \frac{1}{t_k} P(t_1 \cdots t_{k-1}) \geq \frac{1}{t_k t_{k-1}} P(t_1 \cdots t_{k-2}) \geq \cdots \geq \frac{1}{t_k t_{k-1} \cdots t_1}.$$

So far, we have obtained the following bounds for $E_{\text{LC}}^*(n, k)$:

$$\sum_{1 \leq t_1 < \cdots < t_k \leq n} \frac{1}{t_1 t_2 \cdots t_k} \leq E_{\text{LC}}^*(n, k) \leq k! \sum_{1 \leq t_1 < \cdots < t_k \leq n} \frac{1}{t_1 t_2 \cdots t_k}.$$

At last, all that remains for the inequality in Corollary 2 is the identity

$$\sum_{1 \leq t_1 < \cdots < t_k \leq n} \frac{1}{t_1 t_2 \cdots t_k} = \frac{1}{n!} \binom{n+1}{k+1}, \quad (32)$$

which can be found in [1, pp.120-121] (without interpretation), or other references.

Indeed, recalling that $\binom{n+1}{k+1}$ is the coefficient of x^{k+1} in the polynomial $x(x+1) \cdots (x+n)$, we have $\binom{n+1}{k+1} = \sum_{1 \leq s_1 < \cdots < s_{n-k} \leq n} s_1 s_2 \cdots s_{n-k}$. Whenever there is a product $s_1 s_2 \cdots s_{n-k}$ with $1 \leq s_1 < \cdots < s_{n-k} \leq n$, there is a fraction $n!/(t_1 t_2 \cdots t_k)$ with $1 \leq t_1 < \cdots < t_k \leq n$, which is both one-to-one corresponding to and equal to it. Hence, identity (32) follows.

4. Proof of Corollary 3

Suppose that $2 < k < n$. We start by separating the summation formula in Theorem 1 for $E_{\text{LC}}^*(n, k)$:

$$\begin{aligned}
E_{\text{LC}}^*(n, k) &= \sum_{1 \leq t_1 < \dots < t_k < n} P(t_1 \dots t_k) + \sum_{1 \leq t_1 < \dots < t_k = n} P(t_1 \dots t_k) \\
&= E_{\text{LC}}^*(n-1, k) + \frac{1}{n} \sum_{1 \leq t_1 < \dots < t_{k-1} < n} \left(P(t_1 \dots t_{k-1}) + \sum_{1 \leq j < k} P(t_1 \dots t_{k-1} n \ominus t_j) \right) \\
&= E_{\text{LC}}^*(n-1, k) + \frac{1}{n} E_{\text{LC}}^*(n-1, k-1) + \frac{1}{n} \sum_{1 \leq t_1 < \dots < t_{k-1} < n} \sum_{1 \leq j < k} P(t_1 \dots t_{k-1} n \ominus t_j),
\end{aligned}$$

where we recall that $t_1 \dots t_{k-1} n \ominus t_j$ can be written as $t_1 \dots t_{j-1} (t_{j+1} - 1) \dots (t_{k-1} - 1) (n - 1)$. For convenience, let $tail(n, k)$ denote the last term of the above formula. Then by inequality (30), we have

$$\begin{aligned}
tail(n, k) &\leq \frac{1}{n} \sum_{1 \leq t_1 < \dots < t_{k-1} < n} (k-1) P(t_1 \dots t_{k-2} (n-1)) && \text{(by choosing the max } P(\cdot)\text{)} \\
&\leq \frac{k-1}{n} \sum_{1 \leq t_1 < \dots < t_{k-1} < n} \frac{k-1}{n-1} P(t_1 \dots t_{k-2}) && \text{(by choosing the max } P(\cdot)\text{)} \\
&= \frac{(k-1)^2}{n(n-1)} \sum_{k-2 \leq j \leq n-2} \sum_{1 \leq t_1 < \dots < t_{k-2} \leq j} P(t_1 \dots t_{k-2}) && \text{(by letting } j = t_{k-1} - 1\text{)} \\
&= \frac{(k-1)^2}{n(n-1)} \sum_{k-2 \leq j \leq n-2} E_{\text{LC}}^*(j, k-2) \\
&\leq \frac{(k-1)^2}{n} E_{\text{LC}}^*(n-2, k-2) && \text{(since } k > 2\text{).}
\end{aligned}$$

So far, we have obtained the following recursive formula for $E_{\text{LC}}^*(n, k)$:

$$\begin{aligned}
E_{\text{LC}}^*(n, k) &= E_{\text{LC}}^*(n-1, k) + \frac{1}{n} E_{\text{LC}}^*(n-1, k-1) + tail(n, k), \\
\text{with } tail(n, k) &\leq \frac{(k-1)^2}{n} E_{\text{LC}}^*(n-2, k-2) \quad (2 < k < n). \quad (33)
\end{aligned}$$

Now, we claim that $E_{\text{LC}}^*(n, k)$ holds the bound

$$E_{\text{LC}}^*(n, k) \leq \frac{1}{k!} H_n^k + c H_n^{k-1} \quad (\exists c > 0, n \rightarrow \infty), \quad (34)$$

with c being independent of n . To prove this bound by induction, we first note that $E_{\text{LC}}^*(n, 1)$, $E_{\text{LC}}^*(n, 2)$, and $E_{\text{LC}}^*(k, k)$ hold it for any $c > 1$. For $2 < k < n$, we assume that $E_{\text{LC}}^*(n-1, k)$,

$E_{\text{LC}}^*(n-1, k-1)$, and $E_{\text{LC}}^*(n-2, k-2)$ all hold this bound. Then, substituting the assumed bounds into (33), we have

$$\begin{aligned}
E_{\text{LC}}^*(n, k) &\leq \frac{1}{k!} H_{n-1}^k + c H_{n-1}^{k-1} + \frac{1}{n} \left(\frac{1}{(k-1)!} H_{n-1}^{k-1} + c H_{n-1}^{k-2} \right) + \text{tail}(n, k) \\
&= \frac{1}{k!} \left(H_{n-1}^k + \frac{k}{n} H_{n-1}^{k-1} \right) + c \left(H_{n-1}^{k-1} + \frac{k-1}{n} H_{n-1}^{k-2} \right) + \text{tail}(n, k) - \frac{c(k-2)}{n} H_{n-1}^{k-2} \\
&\leq \frac{1}{k!} \left(H_{n-1} + \frac{1}{n} \right)^k + c \left(H_{n-1} + \frac{1}{n} \right)^{k-1} + \text{tail}(n, k) - \frac{c(k-2)}{n} H_{n-1}^{k-2} \\
&\leq \frac{1}{k!} H_n^k + c H_n^{k-1} + \frac{(k-1)^2}{n} \left(\frac{1}{(k-2)!} H_{n-2}^{k-2} + c H_{n-2}^{k-3} \right) - \frac{c(k-2)}{n} H_{n-1}^{k-2} \\
&\leq \frac{1}{k!} H_n^k + c H_n^{k-1} + \frac{c(k-1)^2}{n} H_{n-1}^{k-3} + \frac{(k-1)^2/(k-2)! - c(k-2)}{n} H_{n-1}^{k-2} \\
&\leq \frac{1}{k!} H_n^k + c H_n^{k-1},
\end{aligned}$$

where the last step holds as long as $c > \frac{(k-1)^2}{(k-2)(k-2)!}$ and $n \rightarrow \infty$. Combining the base cases and the inductive case accomplishes the argument for the bound in (34).

Since $H_n = \ln n + O(1)$, the bound in (34) implies that

$$E_{\text{LC}}^*(n, k) \leq \frac{1}{k!} (\ln n)^k + b (\ln n)^{k-1} \quad (\exists b > 0, k = O(1), n \rightarrow \infty), \quad (35)$$

with b being independent of n . From Corollary 2, we also know that

$$E_{\text{LC}}^*(n, k) \geq \frac{1}{n!} \binom{n+1}{k+1} \sim \frac{1}{k!} (\ln n)^k \quad (k = O(1), n \rightarrow \infty). \quad (36)$$

The bounds in (35) and (36) together implies that

$$E_{\text{LC}}^*(n, k) = \frac{1}{k!} (\ln n)^k + O((\ln n)^{k-1}) \quad (k = O(1), n \rightarrow \infty).$$

Finally, by considering the symmetry of $E_{\text{LC}}^*(n, k)$ in n and $n-k$, we accomplish the proof of Corollary 3.

5. Concluding remarks

Letting $\mathcal{X}_{ord} = \langle X_i \rangle_{1 \leq i \leq \binom{n}{k}}$, we denote the sequence $\langle f(X_i^\sigma) \rangle_{1 \leq i \leq \binom{n}{k}}$ by $\mathcal{X}_{ord}^{\sigma, f}$. Then, under a uniformly random $\sigma \in S_n$, and for a fixed ord and f , $\mathcal{X}_{ord}^{\sigma, f}$ forms a “partly random” permutation of $[\binom{n}{k}]$ when $k > 1$. In the previous sections, we only quantify the expected-value measure for one statistic of $\mathcal{X}_{ord}^{\sigma, f}$ —the number of left-to-right maxima—for two special pairs of ord and f : (LEX, lex) and (COL, col). It should be of interest to explore other statistics on $\mathcal{X}_{ord}^{\sigma, f}$, under general or special ord and f . Among all the other statistics discussed in the textbooks [8], [19], and [22] for permutations, what we can immediately answer are the expected numbers of *descents* (consecutive descending pairs) and *inversions* (any descending pairs): In terms of the expected-value measure of these two statistics, $\mathcal{X}_{ord}^{\sigma, f}$ is as “random” as a uniformly random permutation of $[\binom{n}{k}]$.

Proposition 2. For any fixed ord and f , the expected number of descents in $\mathcal{X}_{ord}^{\sigma,f}$ is $\binom{n}{k} - 1)/2$, and the expected number of inversions in $\mathcal{X}_{ord}^{\sigma,f}$ is $\binom{n}{k}(\binom{n}{k} - 1)/4$.

Proof. Let $X_1, X_2 \in \mathcal{X}$ with $X_1 \neq X_2$, and let $Y_1, Y_2 \in \mathcal{X}$ with $Y_1 \neq Y_2$. We claim that

$$|\{\sigma \in S_n : X_1^\sigma = Y_1, X_2^\sigma = Y_2\}| = |\{\sigma \in S_n : X_1^\sigma = Y_2, X_2^\sigma = Y_1\}|.$$

Why? Let $|X_1 \cap X_2| = k - j$ with $1 \leq j \leq k$. Then, these two cardinalities are both equal to 0 if $|Y_1 \cap Y_2| \neq k - j$, and both equal to $(k - j)!(j!)^2(n - k - j)!$ otherwise. We have shown that for any pair of k -subsets X_1 and X_2 in \mathcal{X}_{ord} , the probability that $f(X_1^\sigma) > f(X_2^\sigma)$ is always $1/2$. \square

For further analysis of the team-hiring problem, perhaps the next concern should be of the variances of the left-to-right maxima in $\mathcal{X}_{ord}^{\sigma,f}$ under a random σ . To obtain a probabilistic analysis, as dealing with expectations, we should focus on the worst variance over all f . Precisely, for a fixed ordering, what we are concerned with should be

$$V_{ord}^*(n, k) := \max_{f \in \mathcal{F}} V_{ord}^f(n, k) := (1/n!) \max_{f \in \mathcal{F}} \sum_{\sigma \in S_n} (|\mathcal{M}_{ord}^{\sigma,f}| - E_{ord}^f(n, k))^2.$$

Under this definition, even $V_{LEX}^*(n, k)$ and $V_{COL}^*(n, k)$ are difficult to solve. We don't know whether or not $V_{LEX}^*(n, k) = V_{LEX}^{lex}(n, k)$ and $V_{COL}^*(n, k) = V_{COL}^{col}(n, k)$, let alone their formulas. What can be sure now is only that $V_{LEX}^{lex}(n, k) = V_{COL}^{col}(n, k)$, according to assertion (22) in Section 2.4.

Apart from the lexical and colex orderings for listing all k -subsets of $[n]$, there is another well-known sequential ordering, called the *revolving-door* or *minimal-change* ordering (see [14, Section 2.3.3]). The sequence of k -subsets in accordance with this ordering can also be generated efficiently in both time and space, and it looks somewhat more "disorderly" than the lexical and colex orderings. Let "REV" be the tag of this revolving-door ordering. Is $E_{REV}^*(n, k)$ asymptotically smaller than $E_{LC}^*(n, k)$? That is the final question we are mostly concerned with.

References

- [1] V. Adamchik. On Stirling numbers and Euler sums, *Journal of Computational and Applied Mathematics*, 79:119-130, 1997.
- [2] Z.-D. Bai, C.-C. Chao, H.-K. Hwang, and W.-Q. Liang. On the variance of the number of maxima in random vectors and its applications. *The Annals of Applied Probability*, 8:886-895, 1998.
- [3] J. L. Bentley, T. H. Kung, M. Schkolnick, and C. D. Thompson. On the average number of maxima in a set of vectors and applications. *Journal of ACM*, 25:536-543, 1978.
- [4] C. Buchta. On the average number of maxima in a set of vectors. *Information Processing Letters*, 33:63-65, 1989.
- [5] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. MIT Press, 2nd edition, 1998, 3rd edition, 2009.
- [6] T. Ferguson. Who solved the secretaty problem? *Statistical Science*, 4:282-296, 1989.
- [7] N. Glick. Breaking records and breaking boards. *American Mathematical Monthly*, 85:2-26, 1978.
- [8] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, 2nd edition, 1994.

- [9] A. Helmi and A. Panholzer. Analysis of the “hiring above the median” selection strategy for the hiring problem. *Algorithmica*, 66:762-803, 2013.
- [10] E. S. Key. On the number of records on an iid discrete sequence. *Journal of Theoretical Probability*, 18:99-107, 2005.
- [11] A. Knopfmacher and T. Mansour. Record statistics in a random decomposition. *Discrete Applied Mathematics*, 160:593-603, 2012.
- [12] D. E. Knuth. *The Art of Computer Programming. Volume 2: Seminumerical Algorithms*, Addison-Wesley, 3rd edition, 1997.
- [13] D. E. Knuth, *The Art of Computer Programming. Volume 4A: Combinatorial Algorithms, Part 1*, Addison-Wesley, 2011.
- [14] D. L. Kreher and D. R. Stinson. *Combinatorial Algorithms: Generation, Enumeration, and Search*. CRC Press, 1998.
- [15] J. H. Mueller, C. Sanchet-Sanchet, L. F. Simões, and D. Izzo, Optimal orderings of k -subsets for star identification, In *IEEE Symposium Series on Computational Intelligence (SSCI)*, 2016, pp.1-8.
- [16] A. N. Myers and H. S. Wilf. Left-to-right maxima in words and multiset permutations. *Israel Journal of Mathematics*, 166:167-183, 2008.
- [17] A. Nijenhuis and H. S. Wilf. *Combinatorial Algorithms for Computers and Calculators*. Academic Press, 2nd edition, 1978.
- [18] H. Prodinger. Records in geometrically distributed words: sum of positions. *Applicable Analysis and Discrete Mathematics*, 2:234-240, 2008.
- [19] R. Sedgewick and P. Flajolet. *An Introduction to the Analysis of Algorithms*. Addison-Wesley, 2nd edition, 2013.
- [20] I. Semba. An efficient algorithm for generating all k -subsets ($1 \leq k \leq m \leq n$) of the set $\{1, 2, \dots, n\}$ in lexicographical order. *Journal of Algorithms*, 5:281-283, 1984.
- [21] N. J. A. Sloane. editor, *The On-Line Encyclopedia of Integer Sequences*. Published electronically at <http://oeis.org>.
- [22] R. P. Stanley. *Enumerative Combinatorics, Volume 1*. Cambridge University Press, 2nd edition, 1997.
- [23] H. S. Wilf. On the outstanding elements of permutations, Jan. 1995. Available at <http://www.math.upenn.edu/~wilf/website/outstelmts.pdf>.