Parallel matching for ranking all teams in a tournament

Kei Kobayashi*, Hideki Kawasaki† and Akimichi Takemura‡
Department of Mathematical Informatics
Graduate School of Information Science and Technology
University of Tokyo

May, 2003

Abstract

We propose a simple and efficient scheme for ranking all teams in a tournament, where matches can be played simultaneously. We show that the distribution of the number of rounds of the proposed scheme can be derived by lattice path counting techniques used in ballot problems. We also discuss our method from the viewpoint of parallel sorting algorithms.

Key words: ballot problem, Bradley-Terry model, extreme value distribution, Hasse diagram, Kolmogorov-Smirnov statistic, lattice path counting, odd-even merge, parallel sorting, partially ordered set.

1 Introduction

Consider a tournament style matching of teams or players in sports, such as football or tennis. The winner of the final match is naturally considered to be the strongest. Usually the loser of the final match is considered to be the second strongest. However this might not be true, considering the possibility that the true second strongest team might have been defeated by the strongest team at an early stage of the tournament. In this case, the true second strongest team could not proceed further. In fact any team defeated by the strongest team at some stage may be the second strongest team. In order to determine the true second strongest team, we have to arrange further matches between the teams defeated by the strongest team. If we want to determine the third strongest

*kkoba@stat.t.u-tokyo.ac.jp
†kawasaki@expm.t.u-tokyo.ac.jp
‡takemura@stat.t.u-tokyo.ac.jp

team, the problem becomes more complicated. In this paper we propose a simple and efficient scheme for determining ranks of all teams in a tournament.

The tournament above refers to single elimination or knockout tournament, where a team exits from the tournament once it is defeated by another team. The problem of a strong team being eliminated too early can be alleviated by double elimination tournaments ([6], [16]), where a team exits from the tournament after two losses. However double elimination does not directly address the problem of ranking.

Initial motivation for this work was the above simple question of how to determine the true second (third, fourth, etc.) strongest player in a tournament. However we found that this problem dates back to Lewis Carroll [4], that it has close connections with lattice path counting in ballot problems, extreme value distribution, partially ordered sets and the parallel sorting algorithms. We believe that, in addition to proposing a new scheme and deriving its properties, the present paper has merits in bringing together these rather separate fields in a problem of historical interest.

In many sports games, matches can be played simultaneously. We call each set of simultaneous matches a *round*. We use the number of rounds as a measure of the efficiency of a ranking scheme while there have been several researches evaluating the number of matches for some tournaments (for example, [5]). We assume some probabilistic models of the result of each game and investigate the distribution of the number of rounds for determining the whole ranking. It is combinatorially very difficult to obtain the optimal ranking scheme in the sense of minimizing the expected number of rounds. Although our scheme is not optimal, we will check numerically that the expected number of rounds of our scheme is close to optimal for small number of teams.

We adopt the following notations and assumptions in this paper. "Match" and "game" are used synonymously.

1. The set of teams is denoted by $\mathcal{T} = \{t_1, \ldots, t_N\}$, where N is the number of teams. For example $\mathcal{T} = \{\text{Yankees, Blue Jays}, \ldots\}$. The ranking function is defined to be an injection $r : \mathcal{T} \to \{1, \ldots, N\}$, such that $r(t_i) < r(t_j)$ if and only if t_i is (considered to be) stronger than t_j. The strongest team t^* satisfies $r(t^*) = 1$. Similarly for a subset $\mathcal{S} \subset \mathcal{T}$, the relative rank of $t_i \in \mathcal{S}$ in $\mathcal{S}$ is denoted by $r_{\mathcal{S}}(t_i)$ $(1 \le r_{\mathcal{S}}(t_i) \le \#\mathcal{S})$.

2. Each match is played between two teams. The result of each game is either win or lose. We assume no tie. Each team plays at most one game in a round and each game between two teams is played at most once. We denote the match between t_i and t_j by $[t_i \text{ vs } t_j]$ or simply by $[i \text{ vs } j]$ and the set of matches by $\mathcal{G} = \{[t_i \text{ vs } t_j] \mid 1 \le j < i \le N\}$. Furthermore let

$$w_{ij} = \begin{cases} 1, & \text{if } t_i \text{ wins against } t_j \ , \\ -1, & \text{if } t_j \text{ wins against } t_i \ . \end{cases}$$

Thus $w_{ij} = -w_{ji}$.

3. Let $\mathcal{G}_k \subset \mathcal{G}$ denote the set of matches belonging to the k-th round. We do not impose any upper bound on the size of $\mathcal{G}_k$ and $\lfloor N/2 \rfloor$ matches can be played simultaneously. This corresponds to the situation, where there is no restriction on play fields.

4. $\mathcal{G}_k$ can depend on the results of the past rounds $\cup_{i=1}^{k-1} \mathcal{G}_i$.

5. If team t_i wins against t_j in a match, the ranking of t_i must be higher than t_j. In other words, $w_{ij} = 1 \Rightarrow r(t_i) < r(t_j)$.

6. Two teams t_i and t_j should not play a match, if their relative strength is already determined by results of past rounds, i.e., if there exist teams $t_{l_1}, \dots, t_{l_m}$ such that $w_{il_1} = w_{l_1 l_2} = \cdots = w_{l_m j} = 1$ (or -1) in past rounds.

The last three assumptions are special characteristics of our problem, which are different from the problem of parallel sorting algorithms. Since the outcome of a match may be random, a ranking scheme can lead to loops or contradictions. For example if $w_{ij} = w_{kl} = 1$ in past rounds, we should avoid matches [i vs l] and [j vs k] simultaneously at the current round, because if $w_{il} = w_{kj} = -1$, then we have the contradiction

$$r(t_i) < r(t_j) < r(t_k) < r(t_l) < r(t_i).$$

Therefore in designing a ranking scheme we have to consider not only the result of the past rounds but also all possible results of the future rounds. Our goal is to find a non-contradictory game scheduling scheme which needs small number of rounds for determining the whole ranking.

Before proposing our ranking scheme, we consider the following subproblem:

Parallel merge problem

Suppose that there are two sets of teams $A = \{a_1, \dots, a_n\}$ and $B = \{b_1, \dots, b_m\}$ with $n \le m$ and assume that relative ranks within each set are already determined as $r_A(a_i) = i$, $i = 1, \dots, n$ and $r_B(b_i) = i$, $i = 1, \dots, m$. Determine the ranks of the teams in the union $A \cup B$.

We propose *equivalent rank matching(ERM)* as a solution to this problem. For simplicity we first assume that the sizes of the sets A and B are equal, i. e. $n = m$.

- Step 1: At the first round, a_i matches against b_i for each $i = 1, \dots, n$, simultaneously. As a result a partial order is introduced into $A \cup B$. The partial order can be conveniently displayed by a diagram, where we put a stronger team immediately above the teams next in the partial order. This diagram is called the Hasse diagram representing the partial order (see the next section for precise definition of the Hasse diagram). As the result of the first round, the Hasse diagram becomes some combination of (single or connected) *diamonds* as Figure 1 and *line segments* as Figure 2. Connected diamonds correspond to the winning streaks of teams from A or teams from B. The rankings of the parts of line segments are already determined at this stage. Thus, if the Hasse diagram becomes a single line, the whole ranking is determined.

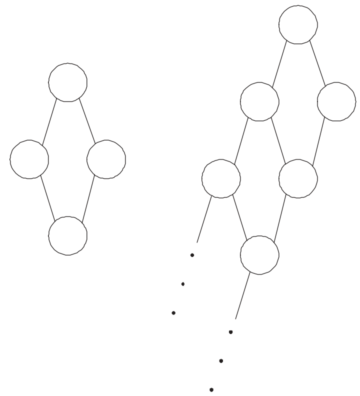


Figure 1: A diamond (left) and connected diamonds (right).

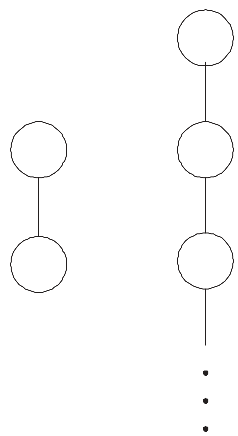


Figure 2: Line segments.

- Step 2: At the next round, only the parts of connected diamonds have to be considered. Next we match the teams corresponding to the horizontal diagonal nodes of each diamond in the Hasse diagram. Figure 3 shows an example of this process. In Figure 3 “W” indicates the winner of the match between horizontal diagonal nodes. It is easy to see that the result of this round is also represented by a Hasse diagram consisting of line segments and diamonds. The diamonds correspond to further winning streaks from the same set of teams as in the previous round.

- Step 3: If the Hasse diagram becomes a line, the whole ranking is determined. Otherwise go back to step 2.

This method is finished in at most n rounds because the first round creates at most n diamonds and the number of diamonds decreases at every round.

So far we have treated the case $n = m$ for simplicity. If $n < m$ we add $m - n$ dummy teams $a_{n+1}, \ldots, a_m$, which are defined to be weaker than b_m. Then we apply the above

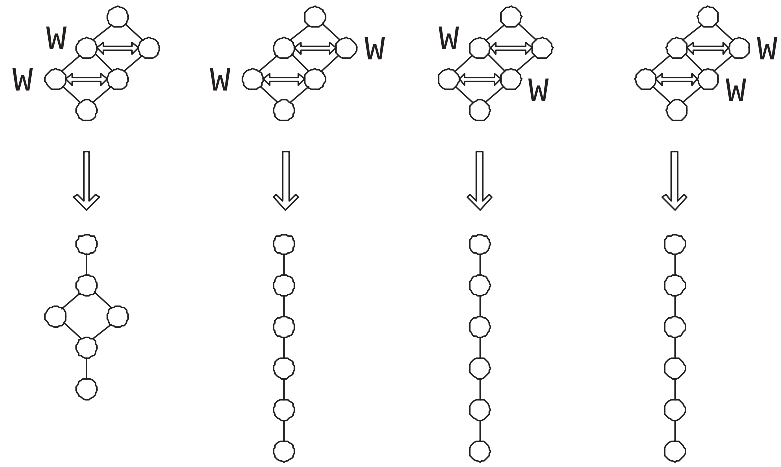


Figure 3: The process of step 2 of ERM.

merging scheme for the case $n = m$.

With ERM, we can determine the whole ranking of $\mathcal{T}$ recursively as follows:

- Step 1: Divide randomly the set of all teams $\mathcal{T}$ into two sets of teams such that the difference of the sizes is at most 1. Repeat these divisions until each set consists of just one team. This process is represented as a binary tree whose leaf corresponds to each team (Figure 4). We call the binary tree *merging tree*.

- Step 2: Merge the sets of the teams in the reverse order of Step 1. Use ERM for each merging (Figure 5). Sets at the same horizontal level of the merging tree are merged in parallel and each merging requires a random number of rounds. We call the set of rounds needed to merge sets at the same horizontal level a *stage*. For example the rounds at the bottom of the binary tree constitute the first stage of rounds and the rounds of the final merging of two sets constitute the last stage.

We call this method *parallel merge sort by equivalent rank matching (PMS)*. If we use the PMS, it is easy to visualize and grasp the schedule and progress of the games. This is one of the benefits of the PMS.

The rest of the paper is organized as follows. In section 2, basic notions of partially ordered sets and their Hasse diagrams are stated and a class of scheduling schemes including PMS is defined. In section 3, the sure winner probabilistic model is investigated. In section 4, the totally random probabilistic model is investigated, where each match is a fair coin tossing. These two models are simple probabilistic models and the distribution of the number of rounds can be evaluated by lattice path counting techniques used in ballot problems. We prove that there exists a stochastic order between these two models (theorem 4.6). In section 5, we discuss our scheme in view of existing literature on parallel sorting algorithms. We also evaluate by simulation the distribution of the number of rounds in the case of a one-parameter Bradley-Terry model connecting the sure winner model and the totally random model.

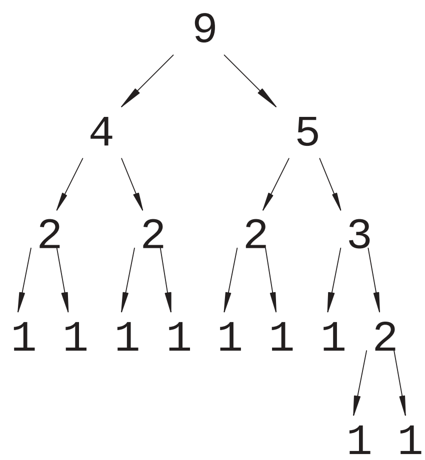


Figure 4: Step 1 of PMS (N=9).

2 Preliminaries

In this section we prepare some basic tools for investigating the behavior of PMS. First, we review basic notions of partially ordered sets and their Hasse diagrams. In the later part of this section, we propose a class of schemes for merging two sets of teams called *rectangle merge scheme*.

2.1 Partially ordered set and Hasse diagrams

Here we quote basic notions of partially ordered set from Chapter 3 of [15]. Partially ordered set P is a set together with a binary relation $\leq$, satisfying the following three axioms: reflexivity ($\forall x \in P, x \leq x$), antisymmetry ($x \leq y$ and $y \leq x \Rightarrow x = y$) and transitivity ($x \leq y$ and $y \leq z \Rightarrow x \leq z$). The notation $x < y$ means $x \leq y$ and $x \neq y$. We say two elements x and y of P are comparable if $x \leq y$ or $y \leq x$, otherwise x and y are incomparable. If every pair of elements of P is comparable, P is called a totally ordered set. If $x, y \in P$, then we say y covers x if and only if $x < y$ and there is no element $z \in P$ satisfying $x < z < y$. The Hasse diagram of a finite partially ordered set P is the graph whose vertices are the elements of P, whose edges are the cover relations, and such that if $x < y$ then y is drawn with a higher horizontal coordinate than x.

In our case $P = \mathcal{T}$ and partial order is induced by matches between the teams. Note that in the Hasse diagram of the teams, we put stronger teams higher. Since a stronger team has a smaller rank, the order in the above definition is conversely related to the ranks. Note that the teams t_i and t_j can be comparable, i.e., their relative strength has been already determined, even if they did not play each other in a match.

If team t_i covers t_j, then it follows that t_i has defeated t_j in a match, but the converse is not necessarily true. Suppose that in past rounds both t_i and t_j have defeated t_k and both cover t_k before the current round. If the match [i vs j] is played at the current round

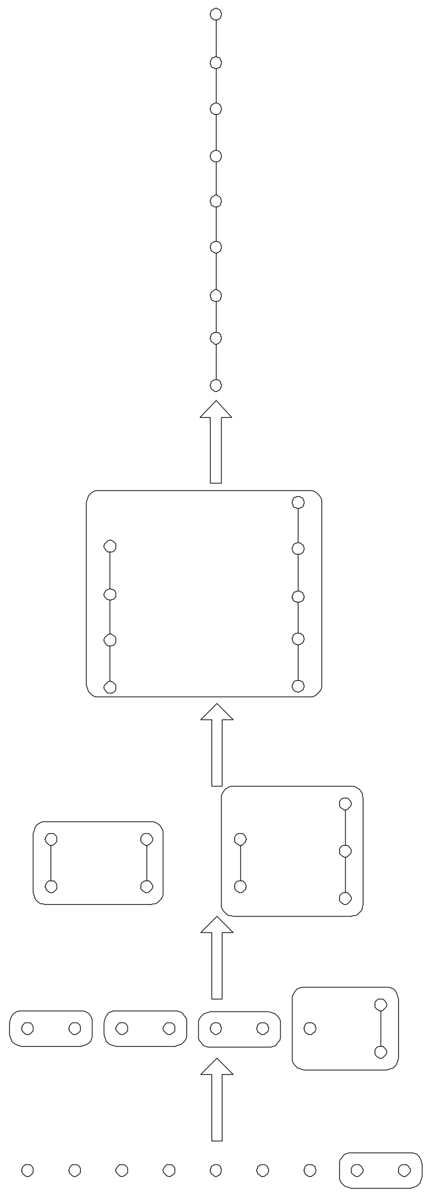


Figure 5: Step 2 of PMS (N=9).

and if $w_{ij} = 1$, then t_i no longer covers t_k.

As the rounds progress, the partial order becomes finer and finer until a whole ranking is determined and $\mathcal{T}$ becomes a totally ordered set.

The Hasse diagram is one of the effective ways to display the game process. However, for most of the game scheduling schemes, the Hasse diagram becomes much more complicated than PMS. If the PMS is used, all Hasse diagrams during the game consist of line segments and diamonds only.

2.2 Probabilistic models

In next two sections, we investigate the following two probabilistic models:

- *sure winner case*
  The "true" ranking is determined before the tournament. If team t_i is truly stronger than t_j, then t_i wins against t_j with probability 1. The randomness comes from the random assignment of N teams to the leaves of the merging tree. All permutations of the teams are equally likely.

- *totally random case*
  The result of each match is independently and identically sampled from Bernoulli distribution with success probability 1/2.

With the sure winner model, of course the final ranking "determined" by the results of the games is equal to the true ranking. On the other hand, in the totally random case, the final ranking is determined only by the random result of the matches. The Bradley-Terry model, which is an intermediate model between the sure winner case and the totally random case, seems to be more realistic than these two models. However theoretical investigation of the Bradley-Terry model in our case seems to be difficult. We study the Bradley-Terry model by simulation in section 5.

2.3 Rectangle merge scheme for parallel merge problem

In this subsection, we consider a parallel merge problem of teams $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_m\}$. We assume that these two sets are already ordered, i.e., $r_A(a_i) = i,\ i = 1, \ldots, n$ and $r_B(b_i) = i,\ i = 1, \ldots, m$. Let $\mathbb{I}_{n,m} = \{(i, j) \mid 1 \le i \le n,\ 1 \le j \le m\}$ be the index set of pairs of a team in A and a team in B. We construct an $n \times m$ matrix $X = (x_{ij})$ whose element x_{ij} represents the round of the match between a_i and b_j. For example if $x_{12} = 2$, the match between a_1 and b_2 is scheduled in the second round. Note that the match between a_i and b_j might or might not be played. Therefore x_{ij} stands for the round in the case that the match between a_i and b_j has to be played. One of our strategies is to determine all elements of X before the matches. We call this class of game scheduling scheme *static rectangle scheme*. We have to impose some restriction on X to cause no contradiction. If we select all matches [i vs j] for round k such that $x_{ij} = k$ and a_i and b_j are incomparable, then any set of results of the round k must make no contradiction. Matrix X satisfying this requirement is called *adequate scheduling matrix*.

Conditions of adequacy of X are different for the sure winner case and for the totally random case. If X is adequate for the totally random case, then it is adequate for the sure winner case. However the converse does not hold. The necessary and sufficient conditions for the two cases are as follows:

X is an adequate scheduling matrix for sure winner model $\Leftrightarrow$

$$
\begin{aligned}
&\forall (i_1, j),\ (i_2, j) \in \mathbb{I}_{n,m} \text{ s.t. } i_1 < i_2 \text{ and } x_{i_1 j} = x_{i_2 j};\\
&\exists i \text{ s.t. } i_1 < i < i_2 \text{ and } x_{ij} < x_{i_1 j},\\
&\qquad\qquad\qquad \text{and}\\
&\forall (i, j_1),\ (i, j_2) \in \mathbb{I}_{n,m} \text{ s.t. } j_1 < j_2 \text{ and } x_{ij_1} = x_{ij_2};\\
&\exists j \text{ s.t. } j_1 < j < j_2 \text{ and } x_{ij} < x_{ij_1}.
\end{aligned} \tag{1}
$$

X is an adequate scheduling matrix for totally random model $\Leftrightarrow$

$$
\begin{aligned}
&\forall (i_1, j_1),\ (i_2, j_2) \in \mathbb{I}_{n,m} \text{ s.t. } i_1 \le i_2,\ j_2 \le j_1 \text{ and } x_{i_1 j_1} = x_{i_2 j_2};\\
&\exists (i, j) \in \mathbb{I}_{n,m} \text{ s.t. } i_1 \le i \le i_2,\ j_2 \le j \le j_1 \text{ and } x_{ij} < x_{i_1 j_1}.
\end{aligned} \tag{2}
$$

Proofs of (1) and (2) are given in Appendix A. The bitonic merge discussed in section 5.1 satisfies (1) but not (2).

Static rectangle schemes determine all elements of X before the matches. A more general class of scheduling methods is to determine the matches of the k-th round depending on the result of the past rounds. We call this class *dynamic rectangle scheme*. Every non-contradictory merging scheme can be specified by dynamic rectangle scheme satisfying the above conditions at each round. However dynamic rectangle scheme is not represented as a single matrix while the static rectangle scheme is. Therefore, it is very difficult to understand visually the process of games with dynamic rectangle schemes.

We define $\breve{X} = (\breve{x}_{ij})$ where $\breve{x}_{ij} = x_{j,m-i+1}$, i.e.,

$$
\breve{X} = \begin{bmatrix}
x_{1m} & x_{2m} & \dots & x_{nm}\\
\vdots & \vdots & & \vdots\\
x_{12} & x_{22} & \dots & x_{n2}\\
x_{11} & x_{21} & \dots & x_{n1}
\end{bmatrix}.
$$

$\breve{X}$ is used for notational consistency with lattice path counting techniques.

We consider a class of scheduling method whose matrix $\breve{X}$ has the same values in the elements at each diagonal line of 45 degrees. We arrange the diagonal lines such that between any two lines whose values are same there is at least one line which has a smaller value (for example Table 1). It is evident that the methods in this class satisfy the conditions (1) and (2). This class is called *45 degrees method*.

Furthermore, matrixes $\breve{X}$ of ERM are like Table 2. Thus ERM is a scheme of 45 degrees method and satisfies the conditions (1) and (2).

Table 1: An example of matrix $\breve{X}$ of 45 degrees method for the $n \times m$ merge problem.

5	2	4	3	1	4
2	4	3	1	4	2
4	3	1	4	2	3
3	1	4	2	3	4
1	4	2	3	4	5

Table 2: The matrix $\breve{X}$ of equivalent rank matching for the $n \times m$ merge problem.

5	4	3	2	1	2
4	3	2	1	2	3
3	2	1	2	3	4
2	1	2	3	4	5
1	2	3	4	5	6

2.4 Notations for merging trees and stages

Finally we set up notations for merging tree and its stages. Let T be a merging tree. Let J denote the number of stages of T. $J+1$ is the depth of the rooted tree T. At the j-th stage ($1 \le j \le J$), $2 \times k_j$ sets of teams are merged into k_j sets in parallel. Let $Y_{jl}, l = 1, \ldots, k_j$, denote the number of rounds needed to merge the sets. Then the total number of rounds Y_j of the j-th stage is given by

$$Y_j = \max(Y_{j1}, \ldots, Y_{jk_j}).$$

The total number of rounds is given by

$$Y = Y_1 + \cdots + Y_J. \tag{3}$$

Note that in the above definition of Y_j, all teams wait until all of k_j mergings are finished at the j-th stage. We call this *synchronous* tree merging. In terms of reducing the total number of rounds, synchronous tree merging is clearly not optimal. For example, as soon as A_{j1} and A_{j2} are merged and A_{j3} and A_{j4} are merged at the j-th stage, we can begin merging $A_{j1} \cup A_{j2}$ and $A_{j3} \cup A_{j4}$ at the $(j+1)$-th stage. We call this scheme *asynchronous* tree merging. In the asynchronous case, the notion of stages lose simultaneity and it becomes harder to grasp the progress of the games. Furthermore the total number of rounds is only defined recursively. For example, with $N = 8$ teams the total number of rounds for the asynchronous case is written as

$$Y = Y_{31} + \max(Y_{21} + \max(Y_{11}, Y_{12}), Y_{22} + \max(Y_{13}, Y_{14})).$$

In this paper we adopt synchronous tree merging for simplicity.

3 Sure winner case

In this section, we study distribution of number of rounds in the sure winner probabilistic model. First we establish the basic independence of number of rounds for different mergings in theorem 3.1. Then in theorem 3.2 we derive recursion formula for evaluating the distribution function of the number of rounds in ERM for merging two sets of teams.

Let $Y_{j1}, \dots, Y_{jk_j},\ j = 1, \dots, J$, be defined as in section 2.4. Consider any merging scheme of two sets of teams A, B. We call the merging scheme *local*, if the resulting ranking of $A \cup B$ only depends on the outcomes of matches between a team of A and a team of B. Then we have the following basic independence of number of rounds for the sure winner probabilistic model.

Theorem 3.1 *Assume the sure winner model. For any local merging scheme of two sets of teams, Y_{jl}, $l = 1, \dots, k_j,\ j = 1, \dots, J$, are mutually independently distributed.*

Proof. We argue recursively from the last stage. At the last stage we merge two sets of teams of A and B of sizes n and m, $n + m = N$. Because the assignment of the teams at the leaves of the binary tree is random, the set of ranks of teams of A are equally likely and each set has the probability $1/\binom{N}{n}$. Given the set of ranks of A, the merging process up to A depends only on relative ranks of teams within A. Therefore the (conditional) joint distribution of the rounds to form A is the same as in the original problem with N replaced by n and the joint distribution does not depend on the ranks of teams of A within $\mathcal{T} = A \cup B$. The same thing holds for B. Therefore by induction the number of rounds are all mutually independent. □

By this independence the distribution function of $Y_j = \max(Y_{j1}, \dots, Y_{jk_j})$ is evaluated as

$$P(Y_j \le y) = \prod_{l=1}^{k_j} P(Y_{jl} \le y).$$

Furthermore $Y_1, \dots, Y_J$ are independent. Therefore the distribution function of $Y = Y_1 + \cdots + Y_J$ can be evaluated if the distribution of Y_{jl} is evaluated.

Now we consider the distribution of number of rounds Y_{jl} in a parallel merge problem. First, we show a necessary and sufficient condition of identification of the ranking from the viewpoint of rectangle scheme. We use the same notation of $\mathcal{T}, A, B, t_i, a_i, b_j$ as in the previous sections. Here $\mathcal{T} = A \cup B$ and the size of $\mathcal{T}$ is $n + m$. Consider an $n \times m$ square lattice whose nodes are denoted as $\{(i, j) \mid i = 0, 1, \dots, n, j = 0, 1, \dots, m\}$. We consider a path which starts from the origin and proceeds at the i-th step to the right by the vector $\boldsymbol{e}_1 = (1, 0)$ if $r^{-1}(i) \in A$, i.e. the i-th strongest team in the true ranking belongs to A, and upward by the vector $\boldsymbol{e}_2 = (0, 1)$ if $r^{-1}(i) \in B$. Thus, each true ranking order corresponds to a path of length $n + m$ from the origin to (m, n) along the edges of the square lattice. If we compare the square lattice and the matrix $\breve{X}$, each square is regarded as a match. It is evident that the ranking is identified if and only if the matches at every "inner corner" of the corresponding path have been held. Here the inner corner is defined to be the lower

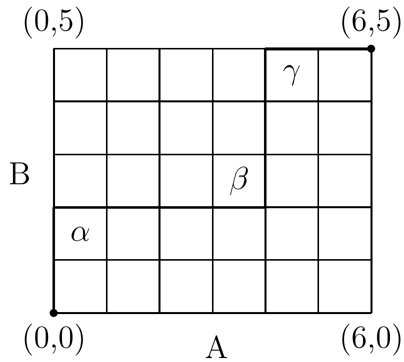


Figure 6: A path corresponding to a "true" ranking order in the sure winner case.

right square when the path turns to the right ( ) and the upper left square when the path turn to the left ( ). For example, the path denoted by the thick line segments in Figure 6 is identified by the matches α (x_{12}), β (x_{43}) and γ (x_{55}). Comparing Figure 6 and matrix $\check{X}$ (Table 2), $x_{12} = x_{43} = 2$ and $x_{55} = 1$. Therefore, this path is identified at the second round.

Let $Y(n, m)$ denote the number of rounds for the $n \times m$ merge problem and let

$$Q_{n,m}(k) = P(Y(n, m) \leq k \mid \text{sure winner model})$$

denote the distribution function of $Y(n, m)$ under the sure winner model. For the case of $n = 0$ or $m = 0$, $Q_{n,m}$ is formally defined as $Q_{n,m}(k) = 1$ for all $k \geq 0$. The following theorem presents a recurrence formula of $Q_{n,m}(k)$.

Theorem 3.2 *For all $n, m \geq 1$,*

$$\binom{n+m}{n} Q_{n,m}(k) = \begin{cases} \binom{2m+k-2}{m} Q_{m+k-2,m}(k), & \text{if } n \geq m+k, & \text{(4a)} \\ \binom{2n+k-2}{n} Q_{n,n+k-2}(k), & \text{if } m \geq n+k, & \text{(4b)} \\ \binom{n+m-1}{n} Q_{n,m-1}(k) + \binom{n+m-1}{m} Q_{n-1,m}(k), & \text{otherwise.} & \text{(4c)} \end{cases}$$

Proof. We first prove the case of $n \geq m + k$. Figure 7 is an example of square lattice of this case ($n = 6,\ m = 4,\ k = 2$). The squares with the diagonal ( ) correspond to the matches held until the k-th round. Because $n \geq m + k$ in this example, the top and rightmost square of the lattice (corresponding to x_{nm}) does not have a diagonal. The top and rightmost square which has the diagonal is the one corresponding to $x_{m+k-1,m}$. Thus, all paths from the origin to (n, m) pass through the point $(m + k - 2, m)$. Conversely, all paths from the origin to $(m + k - 2, m)$ can be extended to (n, m). The number of such paths is $\binom{2m+k-2}{m} Q_{m+k-2,m}(k)$. Therefore, $\binom{n+m}{n} Q_{n,m}(k) = \binom{2m+k-2}{m} Q_{m+k-2,m}(k)$ if $n \geq m + k$. When $m \geq n + k$, the equation is proved by the symmetry. Otherwise, the square corresponding x_{nm} has the diagonal. In this situation, all paths from the origin to $(n, m - 1)$ or $(n - 1, m)$ can be extended to (n, m). The number of such paths is

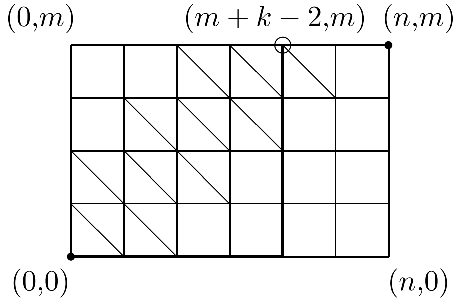


Figure 7: All paths must pass $(m+k-2,m)$. (In this example, $n=6,\ m=4,\ k=2$.)

$\binom{n+m-1}{n}Q_{n,m-1}(k)+\binom{n+m-1}{m}Q_{n-1,m}(k)$. Therefore, $\binom{n+m}{n}Q_{n,m}(k)=\binom{n+m-1}{n}Q_{n,m-1}(k)+\binom{n+m-1}{m}Q_{n-1,m}(k)$. □

For understanding the distribution $Q_{n,m}(k)$, it is convenient to consider a slack problem such that paths can pass through all edges of squares with the diagonal. We call this slack problem *SP*. The slack problem is directly related to lattice path counting for two-sample Kolmogorov Smirnov statistic. Let $\tilde{Y}(n,m)$ denote the number of rounds for SP. $\tilde{Y}(n,m)$ corresponds to the largest (horizontal or vertical) distance from the 45 degrees line through the origin to the path. The following inequality holds between $Y(n,m)$ and $\tilde{Y}(n,m)$:

$$\tilde{Y}(n,m)\leq Y(n,m)\leq \tilde{Y}(n,m)+1. \tag{5}$$

Let $\tilde{Q}_{n,m}(k)$ denote the distribution function of $\tilde{Y}(n,m)$ and let $E^{\tilde{Q}}_{n,m}=E[\tilde{Y}(n,m)]$ denote the expectation. The following facts (see e.g. [10], [7], [8] and [13]) on the two-sample Kolmogorov-Smirnov statistic are well known.

Lemma 3.3

$$\tilde{Q}_{n,m}(k)=\binom{n+m}{n}^{-1}\sum_{j=-\infty}^{\infty}(-1)^j\binom{n+m}{n+j(k+1)},$$

$$E^{\tilde{Q}}_{n,m}=\binom{n+m}{n}^{-1}\sum_{k=0}^{\infty}\sum_{j\neq 0}(-1)^{j+1}\binom{n+m}{n+j(k+1)}. \tag{6}$$

If $m=n$ or $n+1$, for $z>0$

$$\lim_{n\to\infty}\tilde{Q}_{n,m}(\sqrt{n}z)=1-2\sum_{j=1}^{\infty}(-1)^{j-1}\exp(-j^2z^2).$$

From these results we can prove the following results.

Lemma 3.4 (a) *For any odd natural number J_1 and even natural number J_2,*

$$\binom{n+m}{n}^{-1} \sum_{j=1}^{J_2} (-1)^{j+1} \alpha_j \leq E^{\tilde{Q}}_{n,m} \leq \binom{n+m}{n}^{-1} \sum_{j=1}^{J_1} (-1)^{j+1} \alpha_j \tag{7}$$

where

$$\alpha_j = j^{-1} \sum_{l=0}^{j-1} \omega_j^{-ln} (1+\omega_j^l)^{n+m} - \binom{n+m}{n}$$

and ω_j is a primitive j-th root of unity. In particular,

$$\binom{n+m}{n}^{-1} 2^{n+m-1} \leq E^{\tilde{Q}}_{n,m} \leq \binom{n+m}{n}^{-1} 2^{n+m} - 1. \tag{8}$$

(b) *If $m = n$ or $n+1$,*

$$\lim_{n\to\infty} E^{\tilde{Q}}_{n,m}/\sqrt{n} = \sqrt{\pi} \log 2.$$

Proof. Let $a_i = \sum_{l=0}^{\infty} \binom{n+m}{i+jl-1}$ for $i = 1, \ldots, j$. Thus

$$\begin{bmatrix} (1+1)^{n+m} \\ (1+\omega_j)^{n+m} \\ (1+\omega_j^2)^{n+m} \\ \vdots \\ \\ (1+\omega_j^{j-1})^{n+m} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega_j & \omega_j^2 & \omega_j^3 & \ldots & \omega_j^{j-1} \\ 1 & \omega_j^2 & \omega_j^4 & \omega_j^6 & & \\ 1 & \omega_j^3 & \omega_j^6 & \omega_j^9 & & \vdots \\ \vdots & \vdots & & & \ddots & \\ 1 & \omega_j^{j-1} & & \ldots & & \omega_j^{(j-1)^2} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \\ \vdots \\ \\ a_j \end{bmatrix}.$$

The inverse of the matrix is simple and

$$\begin{bmatrix} a_1 \\ a_2 \\ \\ \vdots \\ \\ a_j \end{bmatrix} = j^{-1} \begin{bmatrix} 1 & 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega_j^{-1} & \omega_j^{-2} & \omega_j^{-3} & \ldots & \omega_j^{-(j-1)} \\ 1 & \omega_j^{-2} & \omega_j^{-4} & \omega_j^{-6} & & \\ 1 & \omega_j^{-3} & \omega_j^{-6} & \omega_j^{-9} & & \vdots \\ \vdots & \vdots & & & \ddots & \\ 1 & \omega_j^{-(j-1)} & & \ldots & & \omega_j^{-(j-1)^2} \end{bmatrix} \begin{bmatrix} (1+1)^{n+m} \\ (1+\omega_j)^{n+m} \\ (1+\omega_j^2)^{n+m} \\ \vdots \\ \\ (1+\omega_j^{j-1})^{n+m} \end{bmatrix}.$$

From (6),

$$E^{\tilde{Q}}_{n,m} = \binom{n+m}{n}^{-1} \sum_{j=1}^{\infty} (-1)^{j+1} \alpha_j$$

where

$$\begin{aligned}\alpha_j &= a_{(n+1 \bmod j)} - \binom{n+m}{n} \\ &= j^{-1}\sum_{l=0}^{j-1} \omega_j^{-ln}(1+\omega_j^l)^{n+m} - \binom{n+m}{n}.\end{aligned}$$

From the definition, α_j is nonnegative and monotonically nonincreasing. Thus inequality (7) is valid. Let $J_1 = 1$ and $J_2 = 2$, (8) is derived.

Next we prove (b). The asymptotic expectation under $\tilde{Q}$ is evaluated as

$$\begin{aligned}\lim_{n\to\infty} E_{n,n}^{\tilde{Q}}/\sqrt{n} &= \lim_{n\to\infty}\sum_{k=0}^{\infty}(1-\tilde{Q}_{n,n}(k))/\sqrt{n} \\ &= 2\sum_{j=1}^{\infty}(-1)^{j-1}\int_0^\infty \exp(-j^2z^2)dz \\ &= \sqrt{\pi}\sum_{j=1}^{\infty}(-1)^{j-1}j^{-1} \\ &= \sqrt{\pi}\log 2.\end{aligned}$$

Here the interchanges of integrals and limits can be easily justified by dominated convergence theorem and Fubini's theorem. The result is the same for $m = n+1$. □

Let $E_{n,m}^{Q} = E[Y(n,m)]$ denote the expectation of $Y(n,m)$. By (5), inequalities for $E_{n,m}^{Q}$, the asymptotic distribution of the number of rounds for ERM and the asymptotic expectation $E_{n,n}^{Q}$ and $E_{n,n+1}^{Q}$ are evaluated as follows:

Theorem 3.5 (a)

$$\binom{n+m}{n}^{-1} 2^{n+m-1} \le E_{n,m}^{Q} \le \binom{n+m}{n}^{-1} 2^{n+m}.$$

(b) *If* $m = n$ *or* $n+1$,

$$\begin{aligned}\lim_{n\to\infty} Q_{n,m}(\sqrt{n}z) &= 1 - 2\sum_{j=1}^{\infty}(-1)^{j-1}\exp(-j^2z^2), \qquad z > 0, \\ \lim_{n\to\infty} E_{n,m}^{Q}/\sqrt{n} &= \sqrt{\pi}\log 2.\end{aligned}$$

4 Totally random case

In this section, we evaluate the number of rounds in totally random model with ERM. The result of each match is independently and identically sampled from the Bernoulli trial with success probability 1/2.

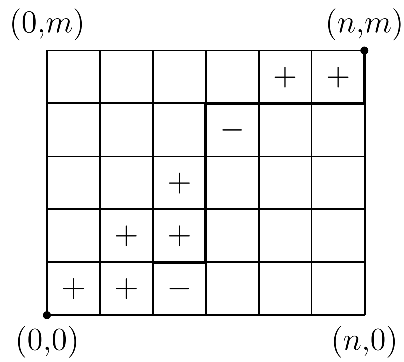


Figure 8: The games held until a path is identified. (+: A wins, −: B wins.)

First, as in the sure winner model, we state the following basic independence of number of rounds for the totally random model.

Theorem 4.1 *Assume the totally random model. For any local merging scheme of two sets of teams,* Y_{jl}*,* $l = 1, \ldots, k_j$*,* $j = 1, \ldots, J$*, are mutually independently distributed.*

Proof. In totally random model, result of each match is sampled independently from Bernoulli distribution. Therefore the theorem is evident. □

Now we concentrate on square lattices as in the previous section. In the sure winner case, the uniform probability is assigned to each path. However, in the totally random case, the probabilities are different. If k matches are necessary to identify the path, the probability assigned to the path is 2^{-k} because the other matches are not held by ERM. In Figure 8, 9 games are needed to identify the path denoted by the thick line segments. By Table 2, the number of rounds k equals 3 in the example.

Let $P_{n,m}(k)$ denote the distribution of the number of rounds with ERM in the totally random case.

$$P_{n,m}(k) = P(Y(n,m) \leq k \mid \text{ totally random model}).$$

The following theorem presents the recurrence formula of $P_{n,m}(k)$.

Theorem 4.2 *For all* $n, m \geq 1$*,*

$$P_{n,m}(k) = \begin{cases} P_{m+k-1,m}(k) - 2^{-k} P_{m+k-1,m-1}(k), & \text{if } m+k \leq n, \quad (9a) \\ P_{n-1,m}(k) - 2^{-n+m-1} P_{n-1,m-1}(k) + 2^{-n+m-1} P_{n,m-1}(k), & \\ & \text{if } m < n < m+k, \quad (9b) \\ 2^{-1} P_{n-1,m}(k) + 2^{-1} P_{n,m-1}(k), & \text{if } n = m, \quad (9c) \\ P_{n,m-1}(k) - 2^{n-m-1} P_{n-1,m-1}(k) + 2^{n-m-1} P_{n-1,m}(k), & \\ & \text{if } n < m < n+k, \quad (9d) \\ P_{n,n+k-1}(k) - 2^{-k} P_{n-1,n+k-1}(k), & \text{if } n+k \leq m. \quad (9e) \end{cases}$$

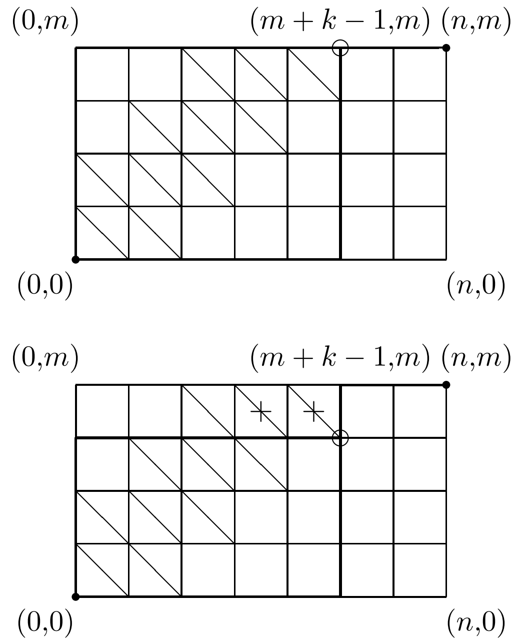


Figure 9: Paths through $(m+k-1, m)$. (In this example, $n = 7,\ m = 4,\ k = 2$.)

Proof. For the case $n = m$ the equation follows from the fact that the match $[a_n \text{ vs } b_n]$ is played at the first round. Because $P_{n,m}(k) = P_{m,n}(k)$ by symmetry, we have only to prove the case of $m < n$. As in the sure winner case, if $n \geq m+k$ all paths corresponding to the ranking orders identified until the k-th round must pass the point $(m+k-2, m)$. We compute the probability of these paths by subtracting the probability of paths through $(m+k-1, m-1)$ and $(m+k-1, m)$ from the probability of paths through $(m+k-1, m)$ (Figure 9). The probability of the former paths is given by $2^{-k} P_{m+k-1,m-1}(k)$. Here, 2^{-k} is the probability such that b_m loses to all of $a_m, a_{m+1}, \ldots, a_{m+k-1}$. Thus,

$$P_{n,m}(k) = P_{m+k-1,m}(k) - 2^{-k} P_{m+k-1,m-1}(k). \tag{10}$$

If $m < n < m+k$, all paths corresponding to the ranking orders which are identified until the k-th round must pass point $(n-1, m)$ or $(n, m-1)$ (Figure 10). By similar argument to (10), the probability of the paths through $(n-1, m)$ is given by $P_{n-1,m}(k) - 2^{-(n-m+1)} P_{n-1,m-1}(k)$. On the other hand, the probability of the paths through $(n, m-1)$ is $2^{-(m-n+1)} P_{n-1,m}(k)$. Here, $2^{-(n-m+1)}$ is the probability such that b_m loses all of $a_m, a_{m+1}, \ldots, a_n$. Thus $P_{n,m}(k) = P_{n-1,m}(k) - 2^{-(n-m+1)} P_{n-1,m-1}(k) + 2^{-(n-m+1)} P_{n,m-1}(k)$.
$\square$

The following lemma gives some properties of $P_{n,m}(k)$. We use them later in the proof of the theorems which evaluate $P_{n,m}(k)$ and its expected value.

Lemma 4.3

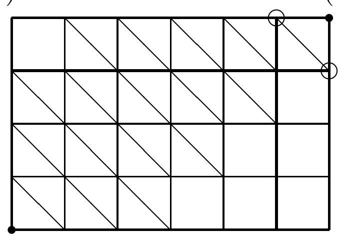


Figure 10: All paths must pass $(n-1, m)$ or $(n, m-1)$. (In this example, $n = 6,\ m = 4,\ k = 3$.)

(a)
$$P_{n,n}(k) = P_{n,n-1}(k) = P_{n-1,n}(k). \tag{11}$$

(b) *If* $m < n \leq m + k$,
$$P_{n,m}(k) \leq P_{n-1,m}(k). \tag{12}$$

The inequality is strict if and only if $n \geq k + 1$.

(c) *If* $m < n < m + k$,
$$P_{n,m}(k) \geq P_{n,m-1}(k). \tag{13}$$

The inequality is strict if and only if $n \geq k + 1$.

(d) *If* $m \leq n$,
$$P_{n,m}(k) \geq 2^{-1} P_{m,m}(k). \tag{14}$$

Proof. (a) is an immediate consequence of (9c) and the symmetry $P_{n,n-1}(k) = P_{n-1,n}(k)$.

Next we prove (b) and (c). If $m < n \leq k$, $P_{n,m}(k) = P_{n-1,m}(k) = P_{n,m-1}(k) = 1$. Therefore we have only to prove for $n \geq k+1$. If $m = n - k$, $P_{n,m}(k) - P_{n-1,m}(k) = -2^{-k} P_{n-1,m-1}(k) < 0$ by (9a). Assume $P_{n,m}(k) < P_{n-1,m}(k)$ for a (m, n) such that $m + 1 < n \leq m + k$, then $P_{n,m+1} - P_{n-1,m+1}(k) = 2^{-n+m}(P_{n,m}(k) - P_{n-1,m}(k)) < 0$ because of (9b). Induction on m proves (b). If $m < n < m + k$,

$$\begin{aligned}
P_{n,m}(k) &= P_{m,m}(k) + \sum_{j=1}^{n-m} (P_{m+j,m}(k) - P_{m+j-1,m}(k)) \\
&\overset{(9b)}{=} P_{m,m}(k) + \sum_{j=1}^{n-m} 2^{-(j+1)} (P_{m+j,m-1}(k) - P_{m+j-1,m-1}(k)) \\
&\overset{(a),(b)}{\geq} P_{m,m-1}(k) + \sum_{j=1}^{n-m} (P_{m+j,m-1}(k) - P_{m+j-1,m-1}(k)) \\
&= P_{n,m-1}(k).
\end{aligned}$$

Thus, (c) is proved.

A proof of (d) is as follows. From (b) and (9a), if $n < m + k < n'$, $P_{n,m}(k) \geq P_{m+k,m}(k) = P_{n',m}(k)$. Therefore, it is enough to prove the case $n = m + k$.

$$\begin{aligned}
P_{m+k,m}(k) &\overset{(9a),(9b)}{=} P_{m,m}(k) + \sum_{j=1}^{k-1} 2^{-(j+1)}(P_{m+j,m-1}(k) - P_{m+j-1,m-1}(k)) \\
&\qquad - 2^{-k} P_{m+k-1,m-1}(k) \\
&\geq P_{m,m}(k) - \sum_{j=1}^{k-1} 2^{-(j+1)} P_{m+j-1,m-1}(k) - 2^{-k} P_{m+k-1,m-1}(k) \\
&\overset{(b)}{\geq} P_{m,m}(k) - P_{m,m-1}(k) \left(\sum_{j=1}^{k-1} 2^{-(j+1)} + 2^{-k} \right) \\
&\overset{(a)}{=} 2^{-1} P_{m,m}(k).
\end{aligned}$$

Thus, (d) is proved. □

Note that from (11) and (12) we have

$$P_{n,n}(k) \leq P_{n-1,n-1}(k),$$

i.e. $P_{n,n}(k)$ is decreasing in n.

Let $E^P_{n,m}$ denote the expectation of $Y(n,m)$ under the totally random model. The following theorem gives the order of $E^P_{n,m}$.

Theorem 4.4 *If $m = n$ or $n + 1$,*

$$\lim_{n\to\infty} \frac{E^P_{n,m}}{(\log n)^{1/2}} = \sqrt{\frac{2}{\log 2}}. \tag{15}$$

Proof. Fix an arbitrarily small $\delta > 0$. At the end of the proof we let $\delta \downarrow 0$.

We first consider an upper bound of $E^P_{n,m}$ for $m = n$ or $n+1$. Let $W^M_n = (w^M_{ij})_{i,j=1,\ldots n}$ be defined such that $w^M_{ij} = 1$ if a_i wins against b_j, $w^M_{ij} = -1$ if b_j wins against a_i and $w^M_{ij} = 0$ if the match is not played. Let $\breve{W}^M_n = (\breve{w}^M_{ij})$, $\breve{w}^M_{ij} = w^M_{j,m-i+1}$, be defined as in section 2.3. At least $k + 1$ rounds are necessary to identify a ranking if and only if $\breve{W}^M_n$ contains at least one $(k+1) \times (k+1)$ submatrix as

$$\begin{bmatrix} w^M_{t,t+k} & \ldots & w^M_{t+k,t+k} \\ \vdots & & \vdots \\ w^M_{tt} & \ldots & w^M_{t+k,t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \ldots & 1 & 1 \\ 1 & 1 & \ldots & 1 & \pm 1 \end{bmatrix} \text{ or } \begin{bmatrix} \pm 1 & -1 & \ldots & -1 & -1 \\ -1 & -1 & \ldots & -1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & \ldots & 0 & 0 \\ -1 & 0 & \ldots & 0 & 0 \end{bmatrix}.$$

Thus by Bonferroni's inequality,

$$1 - P_{n,m}(k) \leq (n-k)2^{-(k+1)(k+2)/2+2} \leq 4n2^{-k^2/2}. \tag{16}$$

Therefore

$$\begin{aligned}
E^P_{n,m} &= 1 + \sum_{k=1}^{m-1}(1 - P_{n,m}(k)) \\
&\leq 1 + \sum_{k=1}^{\lfloor 2(\log_2 n)^{1/2}\rfloor} (1 - P_{n,m}(k)) + \sum_{k=\lfloor 2(\log_2 n)^{1/2}\rfloor+1}^{m-1} (n-k)2^{-(k+1)(k+2)/2+2} \\
&\leq 1 + \sum_{k=1}^{\lfloor 2(\log_2 n)^{1/2}\rfloor} (1 - P_{n,m}(k)) + 4n^2 2^{-2\log_2 n} = \sum_{k=1}^{\lfloor 2(\log_2 n)^{1/2}\rfloor} (1 - P_{n,m}(k)) + 5.
\end{aligned}$$

Now we divide the sum on the right hand side into two parts:

$$\sum_{k=1}^{\lfloor 2(\log_2 n)^{1/2}\rfloor} = \sum_{k=1}^{\lfloor (2(1+\delta)\log_2 n)^{1/2}\rfloor} + \sum_{k=\lfloor (2(1+\delta)\log_2 n)^{1/2}\rfloor+1}^{\lfloor 2(\log_2 n)^{1/2}\rfloor}. \tag{17}$$

By (16) for $k \geq (2(1+\delta)\log_2 n)^{1/2}$

$$1 - P_{n,m}(k) \leq 4n \times n^{-(1+\delta)} = 4n^{-\delta}.$$

Therefore

$$\sum_{k=\lfloor (2(1+\delta)\log_2 n)^{1/2}\rfloor+1}^{\lfloor 2(\log_2 n)^{1/2}\rfloor} (1 - P_{n,m}(k)) \leq 8(\log_2 n)^{1/2} n^{-\delta}.$$

For the first summation on the right hand side of (17) we only use $1 - P_{n,m}(k) \leq 1$. Combining these we obtain

$$E^P_{n,m} \leq 5 + (2(1+\delta)\log_2 n)^{1/2} + 8(\log_2 n)^{1/2} n^{-\delta}.$$

Therefore

$$\limsup_{n\to\infty} \frac{E^P_{n,m}}{(\log n)^{1/2}} \leq \sqrt{\frac{2(1+\delta)}{\log 2}}. \tag{18}$$

We now give a lower bound of $E^P_{n,m}$.

$$\begin{aligned}
P_{n,n}(k) &= P_{n,n-1}(k) \\
&\overset{(9a),(9b)}{\leq} P_{n-1,n-1}(k) - 2^{-2}(P_{n-1,n-2}(k) - P_{n,n-2}(k)) \\
&\leq P_{n-1,n-1}(k) - 2^{-2-3}(P_{n-1,n-3}(k) - P_{n,n-3}(k)) \\
&\leq \dots \\
&\leq P_{n-1,n-1}(k) - 2^{-k(k+1)/2+1}(P_{n-1,n-k}(k) - P_{n,n-k}(k)) \\
&\overset{(9a)}{=} P_{n-1,n-1}(k) - 2^{-(k+1)(k+2)/2+2}P_{n-1,n-k-1}(k) \\
&\overset{(14)}{\leq} P_{n-k-1,n-k-1}(k) - 2^{-(k+1)(k+2)/2+1}P_{n-k-1,n-k-1}(k) \\
&= (1 - 2^{-(k+1)(k+2)/2+1})P_{n-k-1,n-k-1}(k).
\end{aligned}$$

From $P_{k,k}(k) = 1$,

$$P_{n,n}(k) \leq (1 - 2^{-(k+1)(k+2)/2+1})^{\lfloor (n-k)/(k+1) \rfloor} \leq (1 - 2^{-(k+2)^2/2})^{n/(k+2)-2}.$$

$P_{n,n}(k)$ is increasing in k. Thus for $k \leq (2(1-\delta)\log_2 n)^{1/2} - 2$,

$$\begin{aligned}
P_{n,n}(k) &\leq (1 - n^{-(1-\delta)})^{n/(2(1-\delta)\log_2 n)^{1/2}-2} \\
&= \left((1 - n^{-(1-\delta)})^{n^{1-\delta}}\right)^{n^\delta/(2(1-\delta)\log_2 n)^{1/2}-2} \to 0 \qquad (n \to \infty).
\end{aligned}$$

Therefore $\exists N > 0$, $\forall n \geq N$,

$$\begin{aligned}
E^P_{n,m} \geq E^P_{n,n} &\geq 1 + \sum_{k=1}^{n}(1 - P_{n,n}(k)) \\
&\geq 1 + \sum_{k=1}^{\lfloor (2(1-\delta)\log_2 n)^{1/2}-2 \rfloor}(1 - P_{n,n}(k)) \\
&> 1 + (1-\delta)((2(1-\delta)\log_2 n)^{1/2} - 3).
\end{aligned}$$

Therefore

$$\liminf_{n\to\infty} \frac{E^P_{n,m}}{(\log n)^{1/2}} \geq (1-\delta)\sqrt{\frac{2(1-\delta)}{\log 2}}. \tag{19}$$

Letting $\delta \downarrow 0$ in (18) and (19) proves (15). □

In section 3, we derived the asymptotic distribution of $Q_{n,m}$. We now argue that the asymptotic distribution of $P_{n,m}$ does not exist.

Remark 4.5 *There exists no sequence $\{c_n\}$ and a distribution function $F(z)$ such that $P_{n,n}((z - c_n)(\log n)^{1/2}) \to F(z)$.*

This fact is derived from the asymptotic theory of extreme values. Let n be large and consider matching $[a_i \text{ vs } b_i]$, $i = 1, 2, \ldots,$ in sequence. As the first round of ERM progresses, the Hasse diagram grows and consists of blocks of connected diamonds and line segments. The distribution of number of rounds for each block of connected diamonds is independently and identically distributed. By strong law of large numbers, the asymptotics in n and the asymptotics in the number of blocks is essentially the same. We denote the distribution the number of rounds in each block by $P^{CD}(k)$. Then $P_{n,n}(k)$ corresponds to the distribution of the maximum value of i.i.d. samples from $P^{CD}(k)$. Because $1 - P^{CD}(l) \leq \sum_{k=0}^{\infty} 2^{-(l+k)+1}(k+1)(1 - P^{CD}(l-1))$, there exists $c > 0$ such that $1 - P^{CD}(l+1) \leq 2^{-l}c(1 - P^{CD}(l))$ for all sufficiently large l. From theorem 2.4.5 of [9] or theorem 1.7.13 of [11], the asymptotic distribution of $P_{n,n}$ does not exist.

From the theorem 3.5 and 4.4, the asymptotic order of $E^P_{n,m}$ is smaller than that of $E^Q_{n,m}$ when $m = n$ or $n + 1$. Next we prove that $E^P_{n,m}$ is smaller than that of $E^Q_{n,m}$ for all $n, m \geq 1$. Actually we prove a stronger result of stochastic order under two models.

Theorem 4.6 *$Y(n,m)$ is stochastically larger under the sure winner model than under the totally random model, i.e.*

$$P_{n,m}(k) \geq Q_{n,m}(k) \quad \textit{for all } n, m \geq 1, k \geq 1. \tag{20}$$

The inequality is strict if and only if $\max(n,m) \geq k + 1$.

Proof. If $\max(n,m) \leq k$, then $P_{n,m}(k) = Q_{n,m}(k) = 1$. Therefore, we assume that $\max(n,m) \geq k+1$. Because of the symmetry, we assume that $n \geq m$. From theorem 3.2, and by induction we have only to prove

$$\binom{n+m}{m} P_{n,m}(k) > \binom{2m+k-2}{m} P_{m+k-2,m}(k) \quad \text{if } n \geq m+k, \tag{21}$$

$$\binom{n+m}{m} P_{n,m}(k) > \binom{n+m-1}{m} P_{n-1,m}(k) + \binom{n+m-1}{n} P_{n,m-1}(k) \quad \text{if } m < n < m+k \tag{22}$$

and

$$\binom{n+m}{m} P_{n,m}(k) = \binom{n+m-1}{m} P_{n-1,m}(k) + \binom{n+m-1}{n} P_{n,m-1}(k) \quad \text{if } n = m. \tag{23}$$

We first prove (21). If $k = 1$ and $n \geq m + 1$,

$$\begin{aligned} P_{n,m}(1) &\overset{(14)}{\geq} 2^{-1} P_{m,m}(1) \\ &\overset{(11)}{=} 2^{-1} P_{m-1,m}(1) \\ &> \binom{2m+1}{m}^{-1} \binom{2m-1}{m} P_{m-1,m}(1) \\ &\geq \binom{n+m}{m}^{-1} \binom{2m-1}{m} P_{m-1,m}(1). \end{aligned}$$

If $k \geq 2$ and $n \geq m + k$,

$$
\begin{aligned}
P_{n,m}(k) &\overset{(9a)}{=} P_{m+k-1,m}(k) - 2^{-k} P_{m+k-1,m-1}(k) \\
&\overset{(9b)}{=} P_{m+k-2,m}(k) - 2^{-k} P_{m+k-2,m-1}(k) \\
&\overset{(13)}{\geq} (1 - 2^{-k}) P_{m+k-2,m}(k) \\
&> \binom{n+m}{m}^{-1} \binom{2m+k-2}{m} P_{m+k-2,m}(k). \qquad (24)
\end{aligned}
$$

The last inequality is derived as

$$
\begin{aligned}
\binom{n+m}{m}^{-1} \binom{2m+k-2}{m} &\leq \binom{2m+k}{m}^{-1} \binom{2m+k-2}{m} \\
&< \frac{m+k-1}{2m+k-1} \\
&= 1 - (2 + (k-1)/m)^{-1} \\
&\leq 1 - 2^{-k}. \qquad (25)
\end{aligned}
$$

Inequality (22) is proved as follows:

$$
\begin{aligned}
P_{n,m}(k) &\overset{(9b)}{=} P_{n-1,m}(k) - 2^{-n+m-1} P_{n-1,m-1}(k) + 2^{-n+m-1} P_{n,m-1}(k) \\
&\overset{(11),(13)}{\geq} (1 - 2^{-(n-m+1)}) P_{n-1,m}(k) + 2^{-(n-m+1)} P_{n,m-1}(k) \\
&> \binom{n+m}{m}^{-1} \left\{ \binom{n+m-1}{m} P_{n-1,m}(k) + \binom{n+m-1}{n} P_{n,m-1}(k) \right\}.
\end{aligned}
$$

The last inequality is derived from (25), $P_{n-1,m}(k) > P_{n,m-1}(k)$ which follows from lemma 4.3(b) and (c), and the fact

$$
1 = (1 - 2^{-(n-m+1)}) + 2^{-(n-m+1)} = \binom{n+m}{m}^{-1} \left\{ \binom{n+m-1}{m} + \binom{n+m-1}{n} \right\}.
$$

Equation (23) is derived from lemma 4.3(a). Thus, the theorem is proved. □

From theorem 4.6, the following fact is proved.

Corollary 4.7

$$
E^P_{m,n} \leq E^Q_{m,n}. \qquad (26)
$$

The inequality is strict if and only if $\max(m, n) \geq 2$.

Table 5, 6 and Figure 12 in appendix B present $Q_{n,n}(k)$ and $P_{n,n}(k)$ for some small n.

Since stochastic order is preserved under convolution and under taking maximum of independent random variables, we have the following corollary on the stochastic order of the total number of rounds Y in (3) under two models.

Corollary 4.8 *Let Y be the total number of rounds to determine ranking of N teams. Then*

$$P(Y \leq k \mid \text{totally random model}) \geq P(Y \leq k \mid \text{sure winner model}), \quad \forall k \geq 1.$$

The inequality is strict for all $N \geq 3$ and for all k in the support of Y.

Table 7 and Figure 13 in appendix C present the expected value and the standard deviation of the total number of rounds Y by PMS for $N = 2, \ldots, 40$ teams. Figure 14 is the distribution of the number of rounds Y for some N.

5 Some discussions

5.1 Parallel sorting

There is extensive literature on parallel sorting algorithms[14], [12]. EREW P-RAM is exclusive read exclusive write parallel random-access machine which allows no concurrent reads and no concurrent writes. This computing model is the closest to our setting. However there are several differences between the parallel sorting problem by EREW P-RAM and our parallel matching problem. In the parallel sorting problem, only the sure winner stochastic model is assumed, i.e. a "true" ranking is given first and fixed during the whole algorithm. No contradiction of ranking occurs. Thus wider class of algorithms (scheduling methods) are admissible than the parallel matching problem. Furthermore, the parallel sorting algorithms compare already comparable teams. Although this seems to be waste of calculation cost (or the number of necessary rounds), if we consider the cost of the worst ordering case (or the worst "true" ranking case), the waste does not usually affect the order of the cost. Actually, most papers on parallel sorting algorithms evaluate the number of rounds in the worst case. This is one of the largest difference between those papers and the present paper, which evaluates the expectation of the number of rounds.

In spite of these differences, some results on the parallel sorting are useful for the parallel matching problem. Odd-even merge is one of the standard parallel merge algorithms [2]. If we omit matches between comparable teams, odd-even merge corresponds to a rectangle method which has the matrix X as $x_{ij} = \min\{k > 0 \mid n + 1 - i - j \equiv 0 \mod 2^{\lceil \log_2 n \rceil - k + 1}\}$ (see Table 3). It is evident that this algorithm is a 45 degrees method and an adequate rectangle method. The order of the calculation cost of the worst case of this algorithm is easily proved to be $O(\log n)$. Thus the expected number of rounds with the sure winner model is at most $O(\log n)$. This means that the order of expectation $O(\sqrt{n})$ of ERM with sure winner model is not optimal. Another standard parallel merge algorithm is the bitonic merge [2], whose scheduling matrix is upside down of Table 3. In particular the bitonic merge matches teams in the reverse order: [a_1 vs b_n], [a_2 vs b_{n-1}], ..., [a_n vs b_1] at the first round. This is clearly contradictory in the totally random model.

Even if we use the odd even merge or bitonic merge, the whole sorting needs $O((\log n)^2)$ costs in the worst case. The algorithm given by Ajatak, Komlos and Szemeredi [1] and several other algorithms present $O(\log n)$ optimal parallel sorting algorithm. However, all of them are contradictory in the totally random case.

Table 3: An example of scheduling matrix $\check{X}$ of odd-even merge ($n = m = 8$).

4	3	4	2	4	3	4	1
3	4	2	4	3	4	1	4
4	2	4	3	4	1	4	3
2	4	3	4	1	4	3	4
4	3	4	1	4	3	4	2
3	4	1	4	3	4	2	4
4	1	4	3	4	2	4	3
1	4	3	4	2	4	3	4

5.2 Some optimality of PMS for small size problems

This subsection presents some results of sure winner model with PMS. When the number of teams is small, we can count all permutations of "true" ranking and find the optimal algorithm in the sense minimum expected number of rounds. For example, we present the optimal algorithm for $N = 5$ in appendix D. The result of $N = 2, 3, 4, 5$ is given in Table 4. This shows that PMS is optimal if $N \leq 4$. Although the order of the expected number of rounds for $n \times n$ PMS is not optimal, PMS is close to optimal with small N. For larger N, it seems combinatorially formidable to find the optimal algorithm under the sure winner case.

Table 4: The expected number of rounds by optimum scheduling and PMS.

#teams	optimum	PMS
2	1.00	1.00
3	2.67	2.67
4	2.67	2.67
5	4.00	4.77

5.3 Bradley-Terry model

We assume in the true strength order $t_1, t_2, \ldots, t_N \in \mathcal{T}$ satisfy $r(t_1) < r(t_2) < \cdots < r(t_N)$. In general Bradley-Terry model [3], a positive parameter of strength π_i is assigned to each team t_i. The probability that team t_i wins team t_j is calculated as $p_{ij} = \pi_i/(\pi_i + \pi_j)$. One of the most remarkable properties of the model is that it satisfies $p_{ij}p_{jk}p_{ki} = p_{ji}p_{ik}p_{kj}$. In other words, there is no three-cornered deadlock. It is evident that if $\pi_i = \pi_j$ for all i and j, the Bradley-Terry model corresponds to the totally random model. On the other hand, if $\pi_i/\pi_j \to \infty$ for all $j = i+1$, the asymptotic model corresponds to the sure winner

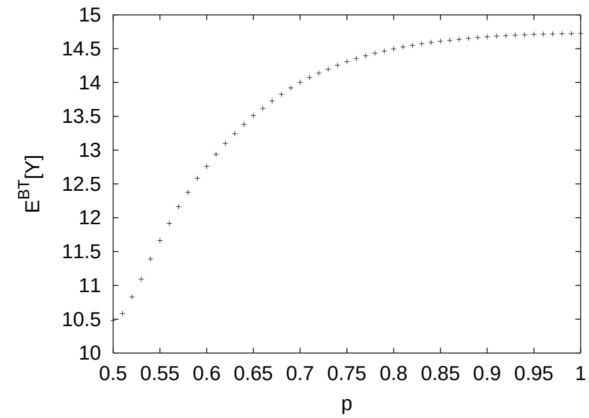


Figure 11: Expected number of rounds by PMS with Bradley-Terry model ($N = 32$).

model. Therefore, the two probabilistic models investigated in the previous sections are two extremes of Bradley-Terry model.

In this sense, Bradley-Terry model is much general than the two probabilistic models and more realistic in the actual games. However, it is very difficult to evaluate the number of rounds of general Bradley-Terry model because we have to consider the ranking of each team in the whole teams even in the recursive subproblems. If we adopt the sure winner or totally random model, only the ranking in the each subset is needed.

We did some simulation studies of PMS with Bradley-Terry model. Figure 11 presents the expected number of rounds in the following simple one-parameter Bradley-Terry model.

$$\frac{\pi_i}{\pi_i + \pi_{i+1}} = p, \qquad i = 1, \ldots, N-1, \quad \frac{1}{2} \leq p \leq 1.$$

Figure 11 shows that the expected number increases monotonically in p.

5.4 Concluding remarks and future works

This paper presented a new game scheduling scheme, parallel merge sorting with equivalent rank matching. The ranking by this scheme does not cause any contradiction for all possible results of simultaneous matches. Furthermore, Hasse diagram of the partial order by this scheme is simple and easy to understand. Two probabilistic models, sure winner model and totally random model, are investigated. For each model, the recurrence formula of the distribution function of the number of rounds for merging is given and the order of the expected number of rounds is evaluated.

In sure winner case, a parallel sorting algorithm for parallel computers gives smaller order of the expected number of rounds than the proposed scheme. The optimality in the sense of the expected number of rounds needs more research. Our investigation of the optimal ranking scheme for the case $N = 5$ suggests that it is combinatorially very difficult to obtain the optimal ranking scheme for $N \geq 6$.

Concerning the totally random case, almost nothing is known on the optimality.

A Proof of (1) and (2)

Proof of (1)
($\Leftarrow$)
At least one of the pairs(a_{i_1}, b_j) and (a_{i_2}, b_j) has become comparable by the match (a_i, b_j). Thus in every round, each b_j has at most one match. For the same reason, each a_i has at most one match in a round. In the sure winner model, this evidently makes no contradiction for any results of the matches.
($\Rightarrow$)
We prove the contraposition. Assume that $\exists (i_1, j), (i_2, j) \in \mathbb{I}_{n,m}$ s.t. $i_1 < i_2$ and $x_{i_1 j} = x_{i_2 j}$; there is no i s.t. $i_1 < i < i_2$ and $x_{ij} < x_{i_1 j}$. Consider paths such that $r(a_{i_1}) = r(a_{i_1+1}) - 1 = \cdots = r(a_{i_2}) - (i_2 - i_1) = r(b_j) - (i_2 - i_1 + 1)$ or $r(b_j) = r(a_{i_1}) - 1 = r(a_{i_1+1}) - 2 = \cdots = r(a_{i_2}) - (i_2 - i_1 + 1)$. Let $k = x_{i_1 j}$. Then by the results until the round $k-1$ these paths are not identified and both of the matches $[a_{i_1}$ vs $b_j]$ and $[a_{i_2}$ vs $b_j]$ have to be played. This is contradiction because b_j can have only one match in a round. Thus X is not an adequate scheduling matrix. $\square$

Proof of (2)
($\Leftarrow$)
At least one of the pairs (a_{i_1}, b_{j_1}) or (a_{i_2}, b_{j_2}) has become comparable by the match (a_i, b_j). In the totally random model, this makes no contradiction for any result of the matches.
($\Rightarrow$)
We prove the contraposition. Assume that $\exists (i_1, j_1), (i_2, j_2) \in \mathbb{I}_{n,m}$ s.t. $i_1 \le i_2, j_2 \le j_1$ and $x_{i_1 j_1} = x_{i_2 j_2}$; there is no $(i, j) \in \mathbb{I}_{n,m}$ s.t. $i_1 \le i \le i_2, j_2 \le j \le j_1$ and $x_{ij} < x_{i_1 j_1}$. Let $k = x_{i_1, j_1}$. Suppose that in the all games until the $(k-1)$-th round, $a_{i'}$ wins $b_{j'}$ if and only if $i' \ge i_1$ and $j' \ge j_1$, then (a_{i_1}, b_{j_1}) and (a_{i_2}, b_{j_2}) have not become comparable until the $(k-1)$-th round. This means that both of the matches $[a_{i_1}$ vs $b_{j_1}]$ and $[a_{i_2}$ vs $b_{j_2}]$ should be done in the k-th round. However, if b_{j_1} wins a_{i_1} and a_{i_2} wins b_{j_2}, we have contradiction

$$r(b_{j_1}) < r(a_{i_1}) < r(a_{i_2}) < r(b_{j_2}) < r(b_{j_1})$$

Thus, X is not an adequate scheduling matrix. $\square$

B $P_{n,n}(k)$ and $Q_{n,n}(k)$

Table 5: $P_{n,n}(k)$ and $Q_{n,n}(k)$ for $n = 2, \ldots, 6$.

k	$P_{2,2}(k)$	$Q_{2,2}(k)$
1	0.5	0.333333333333333
2	1	1
k	$P_{3,3}(k)$	$Q_{3,3}(k)$
1	0.25	0.1
2	0.9375	0.8
3	1	1
k	$P_{4,4}(k)$	$Q_{4,4}(k)$
1	0.125	0.0285714285714286
2	0.890625	0.628571428571429
3	0.99609375	0.942857142857143
4	1	1
k	$P_{5,5}(k)$	$Q_{5,5}(k)$
1	0.0625	0.00793650793650794
2	0.84375	0.476190476190476
3	0.99267578125	0.865079365079365
4	0.9998779296875	0.984126984126984
5	1	1
k	$P_{6,6}(k)$	$Q_{6,6}(k)$
1	0.03125	0.00216450216450216
2	0.7998046875	0.354978354978355
3	0.9892578125	0.779220779220779
4	0.999763488769531	0.954545454545455
5	0.999998092651367	0.995670995670996
6	1	1

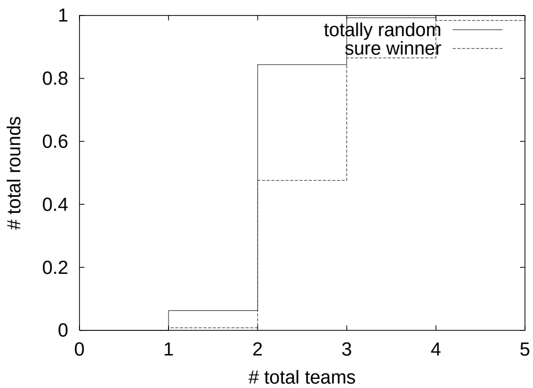
Table 6: $P_{n,n}(k)$ and $Q_{n,n}(k)$ for $n = 7, \ldots, 10$.

k	$P_{7,7}(k)$	$Q_{7,7}(k)$
1	0.015625	0.000582750582750583
2	0.758056640625	0.261072261072261
3	0.985847473144531	0.692890442890443
4	0.999649047851563	0.914918414918415
5	0.999996244907379	0.985431235431235
6	0.999999985098839	0.998834498834499
7	1	1

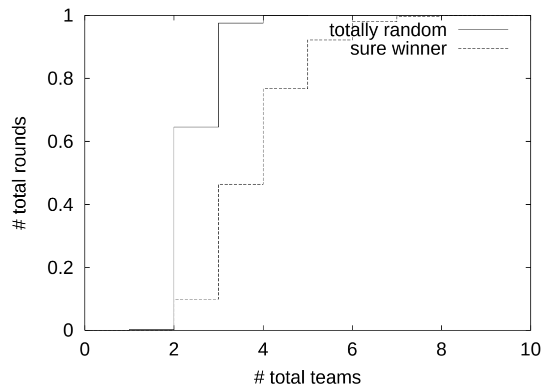
k	$P_{8,8}(k)$	$Q_{8,8}(k)$
1	0.0078125	0.000155400155400155
2	0.718505859375	0.19020979020979
3	0.982449531555176	0.61025641025641
4	0.999534636735916	0.868842268842269
5	0.999994397163391	0.969230769230769
6	0.999999970430508	0.995493395493396
7	0.999999999941792	0.9996891996892
8	1	1

k	$P_{9,9}(k)$	$Q_{9,9}(k)$
1	0.00390625	4.11353352529823e-005
2	0.681015014648438	0.137556561085973
3	0.979063272476196	0.533525298231181
4	0.999420234933496	0.819086795557384
5	0.999992549477611	0.947799259563966
6	0.999999955762178	0.989387083504731
7	0.999999999884039	0.998642533936652
8	0.999999999999886	0.999917729329494
9	1	1

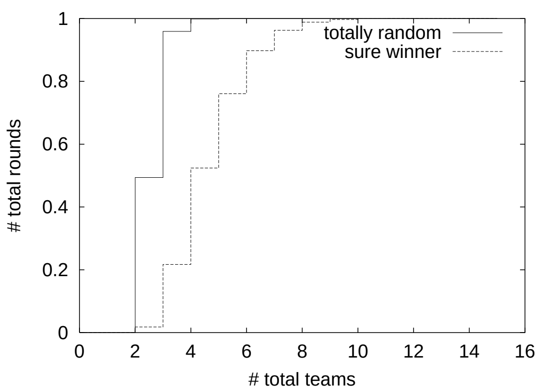
k	$P_{10,10}(k)$	$Q_{10,10}(k)$
1	0.001953125	1.0825088224469e-005
2	0.645481109619141	0.0988980060187491
3	0.975688681006432	0.46371430427158
4	0.999305846402422	0.76768278161467
5	0.999990701800925	0.92212431531317
6	0.999999941093904	0.980319989607915
7	0.999999999826287	0.996471021238823
8	0.999999999999773	0.999599471735695
9	1	0.999978349823551
10	1	1



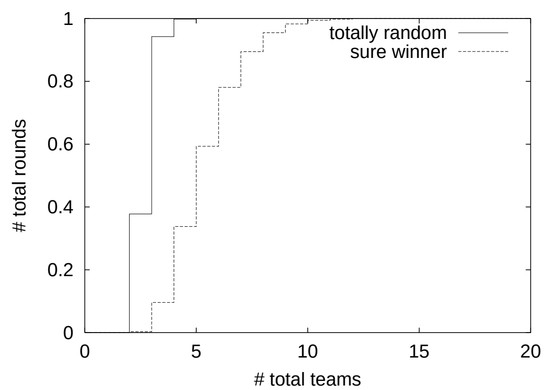
(a) $n = 5$



(b) $n = 10$



(c) $n = 15$



(d) $n = 20$

Figure 12: $P_{n,n}(k)$ and $Q_{n,n}(k)$ for $n = 5, 10, 15, 20$.

C The distribution of the number of rounds for PMS

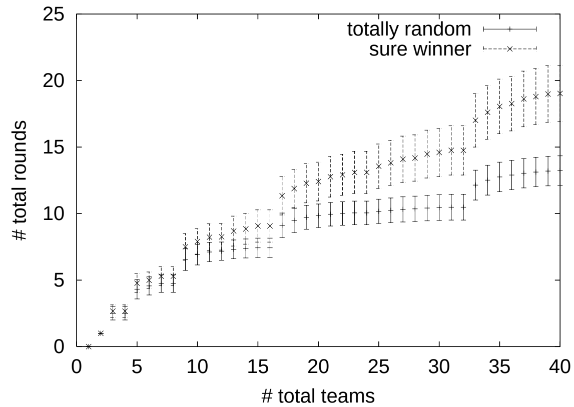


Figure 13: Expectation and standard deviation of the number of rounds by PMS

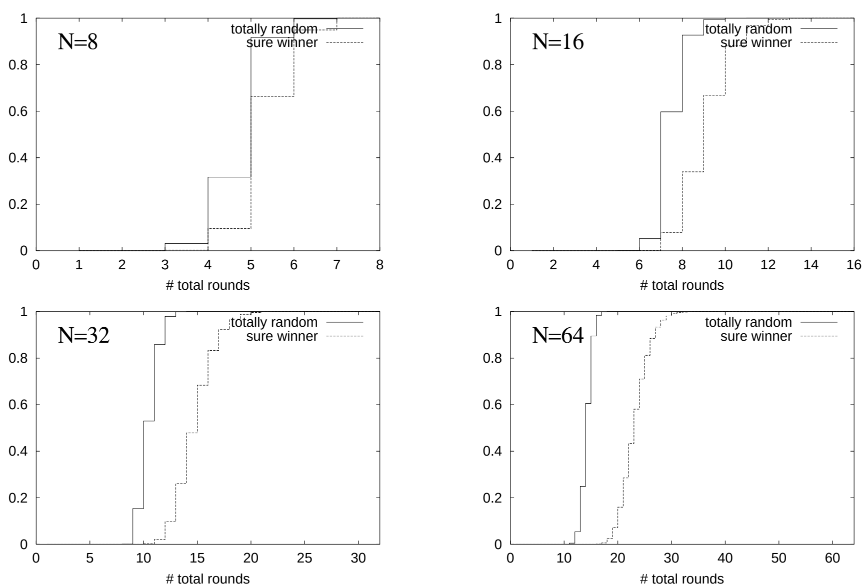


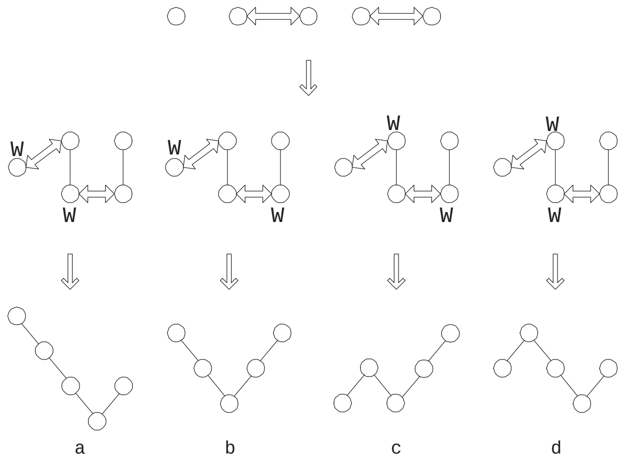
Figure 14: Distribution of the number of rounds by PMS for $N = 8, 16, 32, 64$.

Table 7: Expectation and standard deviation of the number of rounds Y by PMS, P: totally random model, Q: sure winner model.

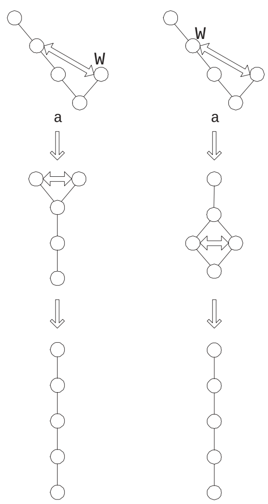
#teams N	$E_P[Y]$	$\sigma_P[Y]$	$E_Q[Y]$	$\sigma_Q[Y]$
2	1	0	1	0
3	2.5	0.25	2.6666667	0.22222222
4	2.5	0.25	2.6666667	0.22222222
5	4.3125	0.52734375	4.7666667	0.51222222
6	4.5625	0.46484375	4.9888889	0.38876543
7	4.7382813	0.43345642	5.2888889	0.51019400
8	4.7382813	0.43345642	5.2888889	0.51019400
9	6.5386963	0.66468607	7.5	0.99920635
10	6.9097900	0.59875284	7.9055556	0.91769400
11	7.1135197	0.52202997	8.2263829	1.0159289
12	7.1760197	0.47124872	8.2510742	0.99245689
13	7.3160209	0.49145171	8.6853533	1.2643305
14	7.3772850	0.50862291	8.8490268	1.3243800
15	7.4281625	0.52404641	9.0688791	1.4676343
16	7.4281625	0.52404641	9.0688791	1.4676343
17	9.1198993	0.84516094	11.337687	2.0752866
18	9.4910901	0.83603157	11.878133	2.0667551
19	9.7198121	0.80287764	12.278142	2.1570787
20	9.8377168	0.77797932	12.405467	2.1251380
21	9.9496437	0.78142229	12.766808	2.3309827
22	10.004891	0.78279367	12.912042	2.3748439
23	10.048581	0.78611977	13.092216	2.5087774
24	10.052488	0.78225930	13.092521	2.5084728
25	10.168297	0.83129325	13.559209	2.7858319
26	10.245227	0.86588579	13.813915	2.8703629
27	10.316693	0.89282647	14.084322	3.0217865
28	10.353913	0.90992329	14.174794	3.0255882
29	10.418159	0.92376163	14.471967	3.2367407
30	10.450092	0.93056406	14.595246	3.2900150
31	10.479308	0.93389641	14.749723	3.4268167
32	10.479308	0.93389641	14.749723	3.4268167
33	12.138754	1.2616910	17.013258	4.0685685
34	12.512224	1.2606555	17.616033	4.1023615
35	12.752424	1.2301378	18.055634	4.2155490
36	12.898304	1.2164817	18.264664	4.1962134
37	13.030912	1.2135887	18.616938	4.3775149
38	13.118209	1.2175961	18.795943	4.4048197
39	13.192656	1.2220949	18.987325	4.5129092
40	13.236359	1.2274871	19.029573	4.4864944

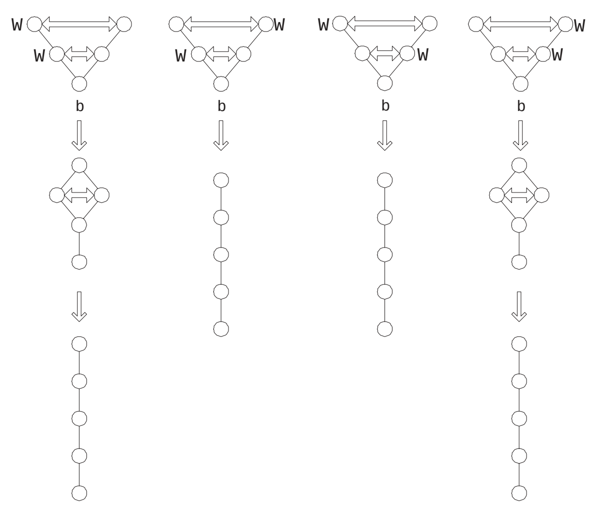
D The optimal algorithm for $N = 5$

The following pictures from Figure 15(a) to Figure 16(d) are Hasse diagrams of the optimal algorithm for $N = 5$ under the sure winner model. It has the minimum expected number of rounds among all adequate dynamic merging schemes.



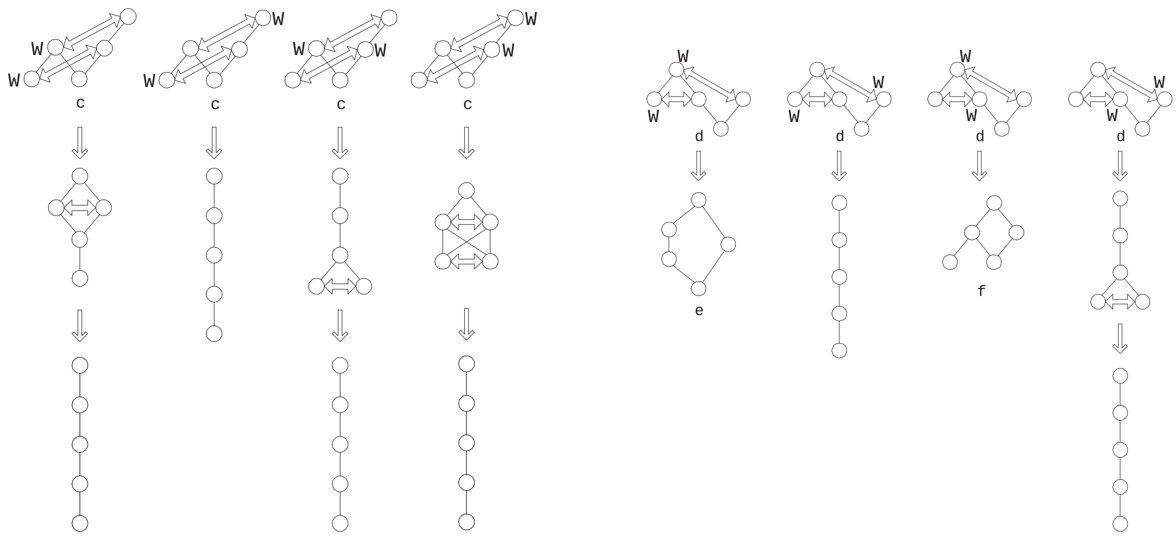
(a) The 1st and 2nd round



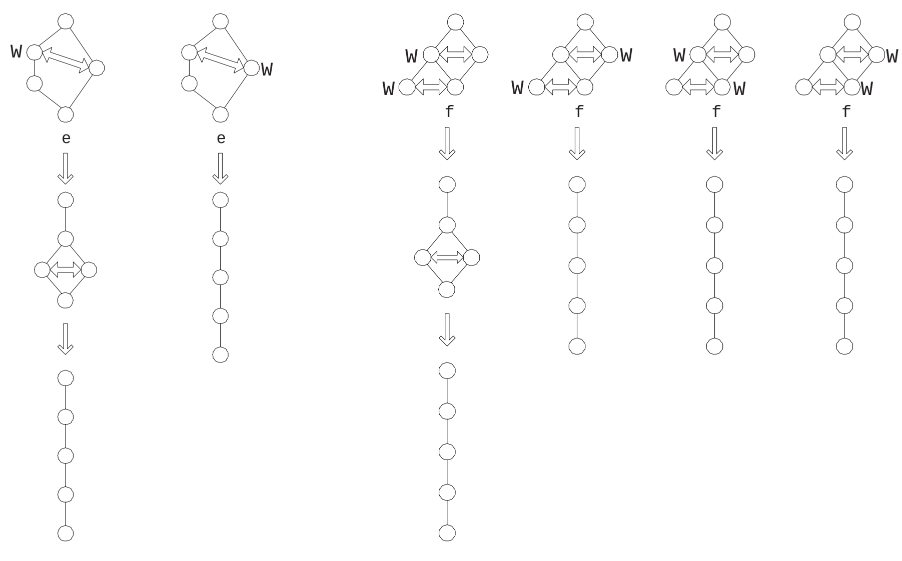


(b) after “a” (c) after “b”

Figure 15: The optimal algorithm for $n = 5$: PART 1.



(a) after “c” (b) after “d”



(c) after “e” (d) after “f”

Figure 16: The optimal algorithm for $n = 5$: PART 2.

Acknowledgement. We are very grateful to Tomomi Matsui for initial suggestion of the idea of ERM.

References

[1] Ajtai, M., Komlos, W. and Szemeredi, E., "Sorting in $c\log(n)$ parallel steps," *Combinatorica*, 3, pp. 1–19, 1983.

[2] Batcher, K., "Sorting networks and their applications," *AFIPS Spring Joint Computing Conference*, 32, pp. 33–39 , 1968.

[3] Bradley, R. A. and Terry, M. E., "Rank analysis of incomplete block designs I. the method of paired comparisons," *Biometrika*, 39, pp. 324–345, 1952.

[4] Carroll, L., "Lawn Tennis Tournaments," *The complete works of Lewis Carroll*, Modern Library, New York, 1947.

[5] David, H. A., *Method of Paired Comparisons*, 2nd edition, Oxford University Press, 1988.

[6] Edwards, C. T., "Double elimination tournaments: counting and calculating", *Amer. Statistician*, 50, pp. 27–33, 1996.

[7] Feller, W., "On the Kolmogorov-Smirnov limit theorems for empirical distributions," *Ann. Math. Statist.*, 19, pp. 177–189, 1948.

[8] Feller, W., *Introduction to Probability Theory and Its Applications* , 3rd edition, John Willy & Suns, 1968.

[9] Galambos, J., *The Asymptotic Theory of Extreme Order Statistics (2nd ed.)*, Robert E. Krieger Publishing Company, Florida, 1987.

[10] Gnedenko, B. V. and Koroljuk, V. S.,"On the maximum discrepancy between two empirical distributions," *Selected Translations in Mathematical Statistics and Probability*, 1, pp. 13–16, 1961.

[11] Leadbetter, M. R., *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, New York, 1983.

[12] Leighton, F. T., *Introduction to Parallel Algorithms and Architectures*, Morgan Kaufmann Publishers, California, 1992.

[13] Mohanty, S. G., *Lattice Path Counting and Applications*, Academic Press, New York, 1979.

[14] Parhami, B., *Introduction to Parallel Processing*, Plenum Press, New York, 1999.

[15] Stanley, R. P., *Enumerative Combinatorics. Volume I*, Cambridge University Press, Cambridge, 1997.

[16] Thompson, P., Jaryszak, E. and Wamil, J., "Reducing the pairing effect in 4-team double-elimination tournaments," *Math. Scientist*, 24, pp. 110–121, 2000.