Abstract

Motivated by online advertisement and exchange settings, greedy randomized algorithms for the maximum matching problem have been studied, in which the algorithm makes (random) decisions that are essentially oblivious to the input graph. Any greedy algorithm can achieve performance ratio 0.5, which is the expected number of matched nodes to the number of nodes in a maximum matching.

Since Aronson, Dyer, Frieze and Suen proved that the Modified Randomized Greedy (MRG) algorithm achieves performance ratio $0.5 + \epsilon$ (where $\epsilon = \frac{1}{400000}$) on arbitrary graphs in the mid-nineties, no further attempts in the literature have been made to improve this theoretical ratio for arbitrary graphs until two papers were published in FOCS 2012. Poloczek and Szegedy also analyzed the MRG algorithm to give ratio 0.5039, while Goel and Tripathi used experimental techniques to analyze the Ranking algorithm to give ratio 0.56. However, we could not reproduce the experimental results of Goel and Tripathi.

In this paper, we revisit the Ranking algorithm using the LP framework. Special care is given to analyze the structural properties of the Ranking algorithm in order to derive the LP constraints, of which one known as the boundary constraint requires totally new analysis and is crucial to the success of our LP.

We use continuous LP relaxation to analyze the limiting behavior as the finite LP grows. Of particular interest are new duality and complementary slackness characterizations that can handle the monotone and the boundary constraints in continuous LP. We believe our work achieves the currently best theoretical performance ratio of $\frac{2 \sqrt{7} - 5}{\pi} \approx 0.523$ on arbitrary graphs. Moreover, experiments suggest that Ranking cannot perform better than 0.724 in general.
1 Introduction

Maximum matching [11] in undirected graphs is a classical problem in computer science. However, as motivated by online advertising [5, 1] and exchange settings [13], information about the graphs can be incomplete or unknown. Different online or greedy versions of the problem [3, 12, 6] can be formulated by the following game, in which the algorithm is essentially oblivious to the input graph.

Greedy Matching Game. An adversary commits to a graph $G(V, E)$ and reveals the nodes $V$ (where $n = |V|$) to the (possibly randomized) algorithm, while keeping the edges $E$ secret. The algorithm returns a list $L$ that gives a permutation of the set $\binom{V}{2}$ of unordered pairs of nodes. Each pair of nodes in $G$ is probed according to the order specified by $L$ to form a matching greedily. In the round when a pair $e = \{u, v\}$ is probed, if both nodes are currently unmatched and the edge $e$ is in $E$, then the two nodes will be matched to each other; otherwise, we skip to the next pair in $L$ until all pairs in $L$ are probed. The goal is to maximize the performance ratio of the (expected) number of nodes matched by the algorithm to the number of nodes in a maximum matching in $G$.

Observe that any ordering of the pairs $\binom{V}{2}$ will result in a maximal matching in $G(V, E)$, giving a trivial performance ratio at least 0.5. However, for any deterministic algorithm, the adversary can choose a graph such that ratio 0.5 is attained. The interesting question is: how much better can randomized algorithms perform on arbitrary graphs? (For bipartite graphs, there are theoretical analysis of randomized algorithms [7, 10] achieving ratios better than 0.5.)

The Ranking algorithm (an early version appears in [8]) is simple to describe: a permutation $\sigma$ on $V$ is selected uniformly at random, and naturally induces a lexicographical order on the unordered pairs in $\binom{V}{2}$ used for probing. Although by experiments, the Ranking algorithm and other randomized algorithms seem to achieve performance ratios much larger than 0.5, until very recently, the best theoretical performance ratio $0.5 + \epsilon$ (where $\epsilon = \frac{1}{100000}$) on arbitrary graphs was proved in the mid-nineties by Aronson et al. [3], who analyzed the Modified Randomized Greedy algorithm (MRG), which can be viewed as a modified version of the Ranking algorithm.

After more than a decade of research, two papers were published in FOCS 2012 that attempted to give theoretical ratios significantly better than the $0.5 + \epsilon$ bound. Poloczek and Szegedy [12] also analyzed the MRG algorithm to give ratio $0.5 + \frac{1}{250} \approx 0.5039$, and Goel and Tripathi [6] analyzed the Ranking algorithm to give ratio 0.56; however, we could not reproduce the experimental results in [6]. Both papers used a common framework which has been successful for analyzing bipartite graphs: (i) utilize the structural properties of the matching problem to form a minimization linear program that gives a lower bound on the performance ratio; (ii) analyze the LP theoretically and/or experimentally to give a lower bound.

In this paper, we revisit the Ranking algorithm using the same framework: (i) we use novel techniques to carefully analyze the structural properties of Ranking for producing new LP constraints; (ii) moreover, we develop new primal-dual techniques for continuous LP to analyze the limiting behavior as the finite LP grows. Of particular interest are new duality and complementary slackness results that can handle monotone constraints and boundary conditions in continuous LP. We believe that this paper achieves the currently best theoretical performance ratio of $\frac{2(5-\sqrt{7})}{9} \approx 0.523$ on arbitrary graphs. As a side note, our experiments suggest that Ranking cannot perform better than 0.724 in general.
1.1 Our Contribution and Techniques

**Theorem 1.1** For the *Greedy Matching Game* on arbitrary graphs, the *Ranking* algorithm achieves performance ratio at least \(\frac{2(5 - \sqrt{7})}{9} \approx 0.523\).

Following previous work on the analysis of *Ranking* [8], we consider a set \(\mathcal{U}\) of *instances*, each of which has the form \((\sigma, u)\), where \(\sigma\) is a permutation on \(V\) and \(u\) is a node in \(V\). An instance \((\sigma, u)\) is *good* if the node \(u\) is matched when *Ranking* is run with \(\sigma\), and *bad* otherwise; an event is a subset of instances. As argued in [12, 6], one can assume that \(G\) contains a perfect matching when analyzing the ratio of *Ranking*. Hence, the performance ratio of *Ranking* is the fraction of good instances.

(1) Relating Bad and Good Events to Form LP Constraints. A simple combinatorial argument [8] is often used to relate bad and good instances. For example, if each bad instance relates to at least two good instances, and each good instance is related to at most one bad instance, then the fraction of good instances would be at least \(\frac{2}{3}\). By considering the structural properties of *Ranking*, one can define various relations between different bad and good events, and hence can generate various constraints in an LP, whose optimal value gives a lower bound on the performance ratio.

Despite the simplicity of this combinatorial argument, the analysis of these relations can be elusive for arbitrary graphs. Hence, we define and analyze our relations carefully to derive three type of constraints: *monotone* constraints, *evolving* constraints, and a *boundary* constraint, the last of which involves a novel construction of a sophisticated relation, and is crucial to the success of our LP.

(2) Developing New Primal-Dual Techniques for Continuous LP. As in previous works, the optimal value of LP\(_n\) decreases as \(n\) increases. Hence, to obtain a theoretical proof, one needs to analyze the asymptotic behavior of LP\(_n\). It could be tedious to find the optimal solution of LP\(_n\) and investigate its limiting behavior. One could also use experiments (for example using strongly factor-revealing LP [10]) to give a proof. We instead observe that the LP\(_n\) has a continuous LP\(_\infty\) relaxation (in which normal variables becomes a function variable). However, the monotone constraints in LP\(_n\) require that the function in LP\(_\infty\) be monotonically decreasing. Moreover, the boundary constraint has its counterpart in LP\(_\infty\). To the best of our knowledge, such continuous LPs have not been analyzed in the literature.

We describe our formal notation in Section 2. In Section 3, we relate bad and good events in order to form LP\(_n\). In Section 4, we prove a lower bound on the performance ratio; in particular, we develop new primal-dual and complementary slackness characterization of a general class of continuous LP, and solve the continuous LP\(_\infty\) relaxation (and its dual). In Appendix B, we describe a hard instance and our experiments show that *Ranking* performs no better than 0.724.

1.2 Related Work

We describe and compare the most relevant related work. Please refer to the references in [12, 6] for a more comprehensive background of the problem. We describe *Greedy Matching Game* general enough so that we can compare different works that are studied under different names and settings. Dyer and Frieze [4] showed that picking a permutation of unordered pairs uniformly at random cannot produce a constant ratio that is strictly greater than 0.5. On the other hand, this framework also includes the MRG algorithm, which was analyzed by by Aronson et al. [3] to prove the first non-trivial constant performance ratio crossing the 0.5 barrier. One can also consider *adaptive*
algorithms in which the algorithm is allowed to change the order in the remaining list after seeing the probing results; although hardness results have been proved for adaptive algorithms [6], no algorithm in the literature seems to utilize this feature yet.

**On Bipartite Graphs.** Running Ranking on bipartite graphs for the Greedy Matching Game is equivalent to running ranking [8] for the Online Bipartite Matching problem with random arrival order [7]. From Karande, Mehta and Tripathi [7], one can conclude that Ranking achieves ratio 0.653 on bipartite graphs. Moreover, they constructed a hard instance in which Ranking performs no better than 0.727; we modify their hard instance and improve the hardness to 0.724.

On a high level, most works on analyzing ranking or similar randomized algorithms on matching are based on variations of the framework by Karp et al. [8]. The basic idea is to relate different bad and good events to form constraints in an LP, whose asymptotic behavior is analyzed when \( n \) is large. For Online Bipartite Matching, Karp et al. [8] showed that ranking achieves performance ratio \( 1 - \frac{1}{e} \); similarly, Aggarwal et al. [1] also showed that a modified version of ranking achieves the same ratio for the node-weighted version of the problem. Sometimes very sophisticated mappings are used to relate different events, and produce LPs whose asymptotic behavior is difficult to analyze. Mahdian and Yan [10] developed the technique of strongly factor-revealing LP. The idea is to consider another family of LPs whose optimal values are all below the asymptotic value of the original LP. Hence, the optimal value of any LP (usually a large enough instance) in the new family can be a lower bound on the performance ratio. The results of [10] implies that for the Greedy Matching Game on bipartite graphs, Ranking achieves performance ratio 0.696.

**Recent Attempts.** No attempts have been made in the literature to theoretically improve the \( 0.5 + \epsilon \) ratio for arbitrary graphs until two recent papers appeared in FOCS 2012. Poloczek and Szegedy [12] used a technique known as contrast analysis to analyze the MRG algorithm and gave ratio \( \frac{1}{2} + \frac{1}{256} \approx 0.5039 \).

Goel and Tripathi [6] showed a hardness result of 0.7916 for any algorithm and 0.75 for adaptive vertex-iterative algorithms. They also analyzed the Ranking algorithm for a better performance ratio. Moreover, they used strongly factor-revealing LP to analyze the asymptotic behavior of their LP; we ran experiment on the LP described in their paper and could not reproduce the ratio 0.56. On the contrary, we discovered that the optimal value of their original LP drops to 0.5001 when \( n = 400 \). Hence, we do not believe strongly factor-revealing LPs can be used to analyze their original LP to give a ratio larger than 0.5001. We describe the details in Appendix A, which includes a link to source codes if the reader would like to verify our experimental results.

**Continuous LP.** Duality and complementary slackness properties of continuous LP were investigated by Tyndall [14] and Levinson [9]. Anand et al. [2] used continuous LP relaxation to analyze online scheduling.

## 2 Preliminaries

Let \([n] := \{1, 2, \ldots, n\}\), \([a..b] := \{a, a+1, \ldots, b\}\) for \(1 \leq a \leq b\), and \(\Omega\) be the set of all permutations of the nodes in \(V\), where each permutation is a bijection \(\sigma : V \to [n]\). The rank of node \(u\) in \(\sigma\) is \(\sigma(u)\), where smaller rank means higher priority.

**The Ranking algorithm.** For the Greedy Matching Game, the algorithm selects a permutation \(\sigma \in \Omega\) uniformly at random, and returns a list \(L\) of unordered pairs according to the lexicographical order induced by \(\sigma\). Specifically, given two pairs \(e_1 = \{u, v\}\) and \(e_2 = \{u_1, v_1\}\) and
Lemma 2.1), without loss of generality, we can assume that the graph is a perfect matching. According to Corollary 2 of [12] (and also implied by our Lemma 2.1 (Removing One Node.) The symmetric difference $M(L) \oplus M(L_u)$ is an alternating path, which contains at least one edge if $u$ is matched in $L$.

Proof: Observe that probing $G$ with $L_u$ is equivalent to probing $G_u$ with $L$, where $G_u$ is exactly the same as $G$ except that the node $u$ is labeled unavailable and will not be matched in any case. Hence, we will use the same $L$ to probe $G$ and $G_u$, and compare what happens in each round to the corresponding matchings $M = M(L)$ and $M_u = M(L_u)$. For the sake of this proof, “unavailable” and “matched” are the same availability status, while “unmatched” is a different availability status.

We apply induction on the number of rounds of probing. Observe that the following invariants hold initially. (i) There is exactly one node known as the crucial node (which is initially $u$) that has different availability in $G$ and $G_u$. (ii) The symmetric difference $M(L) \oplus M(L_u)$ is an alternating path connecting $u$ to the crucial node; initially, this path is degenerate.

Consider the inductive step. Observe that the crucial node and $M(L) \oplus M(L_u)$ do not change in a round except for the case when the pair being probed is an edge in $G$ (and $G_u$), involving the crucial node $w$ with another currently unmatched node $v$ in $G$, and hence $v$ is also unmatched in $G_u$, as the induction hypothesis states that every other node apart from the crucial node has the same availability in both graphs. In this case, this edge is added to exactly one of $M$ and $M_u$. Therefore, $w$ is matched in both graphs (so no longer crucial), and $v$ becomes the new crucial node; moreover, the edge $\{w, v\}$ is added to $M(L) \oplus M(L_u)$, which now is a path connecting $u$ to $v$. This completes the inductive step.

Observe that $u$ is matched in $M$ in the end, iff in some round an edge involving $u$ must be added to $M$ but not to $M_u$, which is equivalent to the case when $M \oplus M_u$ contains at least one edge. ■

The performance ratio $r$ of Ranking on $G$ is the expected number of nodes matched by the algorithm to the number of nodes in a maximum matching in $G$, where the randomness comes from the random permutation in $\Omega$. We consider the set $U := \Omega \times V$ of instances; an event is a subset of instances. An instance $(\sigma, u) \in U$ is good if $u$ is matched in $\sigma$, and bad otherwise.

Perfect Matching Assumption. According to Corollary 2 of [12] (and also implied by our Lemma 2.1), without loss of generality, we can assume that the graph $G(V, E)$ has a perfect matching $M^* \subseteq E$ that matches all nodes in $V$. For a node $u$, we denote by $u^*$ the partner of $u$ in $M^*$ and we call $u^*$ the perfect partner of $u$. From now on, we consider Ranking on such a graph $G$ without mentioning it explicitly again. Observe that for all $\sigma \in \Omega$, $(\sigma, \sigma^{-1}(1))$ is always good; moreover, the performance ratio is the fraction of good instances.

Definition 2.1 $(\sigma_u, \sigma_u^*)$ For a permutation $\sigma$, let $\sigma_u$ be the permutation obtained by removing $u$ from $\sigma$ while keeping the relative order of other nodes unchanged; running Ranking with $\sigma_u$ means
running $\sigma$ while keeping $u$ always unavailable (or simply deleting $u$ in $G$). Let $\sigma_u^t$ be the permutation obtained by inserting $u$ into $\sigma_u$ at rank $i$ and keeping the relative order of other nodes unchanged.

**Fact 2.1 (Ranking is Greedy)** Suppose Ranking is run with permutation $\sigma$. If $u$ is unmatched in $\sigma$, then each neighbor $w$ of $u$ (in $G$) is matched to some node $v$ in $\sigma$ with $\sigma(v) < \sigma(u)$.

Similar to [12, Lemma 3], the following Fact is an easy corollary of Lemma 2.1, by observing that if $(\sigma, u)$ is bad, then $M(\sigma) = M(\sigma_u)$.

**Fact 2.2 (Symmetric Difference)** Suppose $(\sigma, u)$ is bad, and $(\sigma_u^t, u)$ is good for some $i$. Then, the symmetric difference $M(\sigma) \oplus M(\sigma_u^t)$ is an alternating path $P$ with at least one edge, where except for the endpoints of $P$ (of which $u$ is one), every other node in $G$ is either matched in both $\sigma$ and $\sigma_u^t$, or unmatched in both.

**Definition 2.2 ($Q_t$, $R_t$ and $S_t$)** For each $t \in [n]$, let $Q_t$ be the good event that the node at rank $t$ is matched, where $Q_t := \{ (\sigma, u) : \sigma \in \Omega, u = \sigma^{-1}(t) \text{ is matched in } \sigma \}$; similarly, let $R_t$ be the bad event that the node at rank $t$ is unmatched, where $R_t := \{ (\sigma, u) : \sigma \in \Omega, u = \sigma^{-1}(t) \text{ is unmatched in } \sigma \}$.

Moreover, we define the marginally bad event $S_t$ at rank $t \in [2..n]$ by $S_t := \{ (\sigma, u) \in R_t : (\sigma_u^{t-1}, u) \notin R_{t-1} \}$; observe that $S_1 = R_1 = \emptyset$.

Given any $(\sigma, u) \in \mathcal{U}$, the marginal position of $u$ with respect to $\sigma$ is the (unique) rank $t$ such that $(\sigma_u^t, u) \in S_t$, and is null if no such $t$ exists.

Note that for each $t \in [n]$, $Q_t$ and $R_t$ are disjoint and $|Q_t \cup R_t| = n!$.

**Definition 2.3 ($x_t$, $\alpha_t$)** For each $t \in [n]$, let $x_t = \frac{|Q_t|}{n!}$ be the probability that a node at rank $t$ is matched, over the random choice of permutation $\sigma$. Similarly, let $\alpha_t = \frac{|S_t|}{n!}$; observe that $1 - x_t = \frac{|R_t|}{n!}$.

Note that the performance ratio is $\frac{1}{n} \sum_{t=1}^n x_t$, which will be the objective function of our minimization LP. Observe that all $x_t$’s and $\alpha_t$’s are between 0 and 1, and $x_1 = 1$ and $\alpha_1 = 0$. We derive constraints for the variables in the next section.

### 3 Relating Bad and Good Events to Form LP Constraints

In this section we define some relations between bad and good events to form LP constraints. The high level idea is as follows. Suppose $f$ is a relation between $A$ and $B$, where $f(a)$ is the set of elements in $B$ related to $a \in A$, and $f^{-1}(b)$ is the set of elements in $A$ related to $b \in B$. The injectivity of $f$ is the minimum integer $q$ such that for all $b \in B$, $|f^{-1}(b)| \leq q$. If $f$ has injectivity $q$, we have the inequality $\sum_{a \in A} |f(a)| \leq q|B|$, which follows from counting the number of edges in the bipartite graph induced by $f$ on $A$ and $B$. In our constructions, usually calculating $|f(a)|$ is straightforward, but sometimes special attention is required to bound the injectivity.

#### 3.1 Monotone Constraints: $x_{t-1} \geq x_t$, $t \in [2..n]$

These constraints follow from Lemma 3.1 as the $\alpha_t$’s are non-negative.

**Lemma 3.1 (Bad-to-Marginally Bad)** For all $t \in [n]$, we have $1 - x_t = \sum_{i=1}^t \alpha_i$; this implies that for $t \in [2..n]$, $x_{t-1} - x_t = \alpha_t$. 

5
3.3 Boundary Constraint:

The monotone constraints require that the $x_i$’s do not increase. We next derive the evolving constraints that prevent the $x_i$’s from dropping too fast. Fix $t \in [2..n]$. We shall define a relation $f$ between $\bigcup_{i=1}^{t} S_i$ and $\bigcup_{i=1}^{t-1} Q_i$ such that $f$ has injectivity 1, and for $(σ, u) \in S_i$, $|f(σ, u)| = n - i + 1$. This implies Lemma 3.2; from Lemma 3.1, we can express $α_i = x_{i-1} - x_i$ (recall $α_1 = 0$), and rearrange the terms to obtain the required constraint.

Lemma 3.2 (1-to-$(n-i+1)$ Mapping) For all $t \in [2..n]$, we have $\sum_{i=1}^{t}(n-i+1)α_i \leq \sum_{i=1}^{t-1} x_i$.

Proof: We define a relation $f$ between $A := \bigcup_{i=1}^{t} S_i$ and $B := \bigcup_{i=1}^{t-1} Q_i$. Let $(σ, u) \in A$ be a marginally bad instance. Then, there exists a unique $i \in [2..t]$ such that $(σ, u) \in S_i$. If we move $u$ to any position $j \in [i..n]$, $(σ_u^j, u)$ is still bad, because $i$ is the marginal position of $u$ with respect to $σ$. Moreover, observe that $M(σ_u) = M(σ) = M(σ_u)$ for all $j \in [i..n]$. Hence, it follows that for all $j \in [i..n]$, node $u$’s perfect partner $u^*$ is matched in $σ_u$ to the same node $v$ such that $σ(v) = σ_u^j(v) \leq i - 1 \leq t - 1$, where the first inequality follows from Fact 2.1. In this case, we define $f(σ, u) := \{(σ_u^j, v) : j \in [i..n]\} \subset B$, and it is immediate that $|f(σ, u)| = n - i + 1$.

Injectivity. Suppose $(ρ, v) \in B$ is related to some $(σ, u) \in A$. It follows that $v$ must be matched to $u^*$ in $ρ$; hence, $u$ is uniquely determined by $(ρ, v)$. Moreover, $(ρ, u)$ must be bad, and suppose the marginal position of $u$ with respect to $ρ$ is $i$, which is also uniquely determined. Then, it follows that $σ$ must be $ρ_u^i$. Hence, $(ρ, v)$ can be related to at most one element in $A$.

Observing that $S_1 = ∅$, the result follows from $\sum_{i=1}^{t}(n-i+1)|S_i| = \sum_{a \in A} |f(a)| \leq |B| = \sum_{i=1}^{t-1} |Q_i|$, since $|S_i| = n|α_i|$ and $|Q_i| = n|x_i|$.

3.3 Boundary Constraint: $x_n + \frac{3}{2n} \sum_{i=1}^{n} x_i \geq 1$

According to experiments, the monotone and the evolving constraints alone cannot give ratio better than 0.5. The boundary constraint is crucial to the success of our LP, and hence we analyze our construction carefully. The high level idea is that we define a relation $f$ between $R_n$ and $Q := \bigcup_{i=1}^{n} Q_i$. As we shall see, it will be straightforward to show that $|f(a)| = 2n$ for each $a \in R_n$, but it will require some work to show that the injectivity is at most 3. Once we have established these results, the boundary constraint follows immediately from $\sum_{a \in R_n} |f(a)| \leq 3|Q|$, because $\frac{|R_n|}{n} = 1 - x_n$ and $\frac{|Q_i|}{n} = x_i$.

Defining relation $f$ between $R_n$ and $Q$. Consider a bad instance $(σ, u) \in R_n$. We define $f(σ, u)$ such that for each $i \in [n]$, $(σ, u)$ produces exactly two good instances of the form $(σ_u^i, *)$. 
For each $i \in [n]$, we consider $\sigma^i_u$:

1. if $u$ is unmatched in $\sigma^i_u$: ($u$ and $u^*$ cannot be both unmatched)
   - **R(1)**: produce $(\sigma^i_u, u^*)$ and include it in $f(\sigma, u)$;
   - **R(2)**: let $v$ be the partner of $u^*$ in $\sigma^i_u$; produce $(\sigma^i_u, v)$ and include it in $f(\sigma, u)$.

2. if $u$ is matched in $\sigma^i_u$:
   - **R(3)**: produce $(\sigma^i_u, u)$ and include it in $f(\sigma, u)$;
     - (a) if $u^*$ is matched to $u$ in $\sigma^i_u$:
       - **R(4)**: produce $(\sigma^i_u, u^*)$ and include it in $f(\sigma, u)$;
     - (b) if $u^*$ is matched to $v \neq u$ in $\sigma^i_u$:
       - **R(5)**: produce $(\sigma^i_u, v)$ and include it in $f(\sigma, u)$;
     - (c) if $u^*$ is unmatched in $\sigma^i_u$: (all neighbors of $u^*$ in $G$ must be matched)
       - **R(6)**: let $v_o$ be the partner of $u^*$ in $\sigma$, produce $(\sigma^i_u, v_o)$ and include it in $f(\sigma, u)$.

Observe that for $i \in [6]$, applying each rule $R(i)$ produces exactly one good instance. Moreover, for each $i \in [n]$, when we consider $\sigma^i_u$, exactly 2 rules will be applied: if $u$ is unmatched in $\sigma^i_u$, then $R(1)$ and $R(2)$ will be applied; if $u$ is matched in $\sigma^i_u$, then $R(3)$ and one of $\{R(4), R(5), R(6)\}$ will be applied.

**Observation 3.1** For each $(\sigma, u) \in R_n$, we have $|f(\sigma, u)| = 2n$.

**Observation 3.2** If $(\rho, x) \in f(\sigma, u)$, then $\sigma = \rho^n_u$ and exactly one rule can be applied to $(\sigma, u)$ to produce $(\rho, x)$.

**Bounding Injectivity.** We first show that different bad instances in $R_n$ cannot produce the same good instance using the same rule.

**Lemma 3.3 (Rule Disjunction)** For each $i \in [6]$, any $(\rho, x) \in Q$ can be produced by at most one $(\sigma, u) \in R_n$ using $R(i)$.

**Proof:** Suppose $(\rho, x) \in Q$ is produced using a particular rule $R(i)$ by some $(\sigma, u) \in R_n$. We wish to show that in each case $i \in [6]$, we can recover $u$ uniquely, in which case $\sigma$ must be $\rho^n_u$.

The first 5 cases are simple. Let $y$ be the partner of $x$ in $\rho$. If $i = 1$ or $i = 4$, we know that $x = u^*$ and hence we can recover $u = x^*$; if $i = 2$ or $i = 5$, we know that $y = u^*$ and hence we can recover $u = y^*$; if $i = 3$, we know that $u = x$.

For the case when $i = 6$, we need to do a more careful analysis. Suppose $R(6)$ is applied to $(\sigma, u) \in R_n$ to produce $(\rho, x)$. Then, we can conclude the following: (i) in $\sigma = \rho^n_u$, $u$ is unmatched, and $u^*$ is matched to $x$; (ii) in $\rho$, $u$ is matched, $u^*$ is unmatched, and $x$ is matched.

For contradiction’s sake, assume that $u$ is not unique and there are two $u_1 \neq u_2$ that satisfy the above properties. It follows that $u_1^* \neq u_2^*$ and according to property (ii), in $\rho$, both $u_1$ and $u_2$ are matched, and both $u_1^*$ and $u_2^*$ are unmatched; hence, all 4 nodes are distinct. Without loss of generality, we assume that $\rho(u_1^*) < \rho(u_2^*)$. Let $\sigma_2 := \rho^n_{u_2}$, and observe that $\sigma_2(u_1^*) < \sigma_2(u_2^*)$.

Now, suppose we start with $\sigma_2$, and consider what happens when $u_2$ is promoted in $\sigma_2$ resulting in $\rho$. Observe that $u_2$ changes from unmatched in $\sigma_2$ to matched in $\rho$, and by property (i), $u_2^*$ changes
from matched in $\sigma_2$ to unmatched in $\rho$. From Fact 2.2, every other node must remain matched or unmatched in both $\sigma_2$ and $\rho$; in particular, $u_1^*$ is unmatched in $\sigma_2$. However, $x$ is a neighbor of both $u_1^*$ and $u_2^*$ (in $G$), and $\sigma_2(u_1^*) < \sigma_2(u_2^*)$, but $x$ is matched to $u_2^*$ in $\sigma_2$; this contradicts Fact 2.1. ■

Lemma 3.3 immediately implies that the injectivity of $f$ is at most 6. However, to show a better bound of 3, we need to show that some of the rules cannot be simultaneously applied to produce the same good instance $(\rho, x)$. We consider two cases for the remaining analysis.

Case (1): $x$ is matched to $x^*$ in $\rho$

**Lemma 3.4** For $(\rho, x) \in Q$, if $x$ is matched to $x^*$ in $\rho$, then we have $|f^{-1}(\rho, x)| \leq 3$.

**Proof:** If $(\rho, x)$ is produced using R(1), then $x^*$ must be unmatched in $\rho$; if $(\rho, x)$ is produced by $(\sigma, u)$ using R(2), then $x$ must be matched to $u^* (\neq x^*)$ in $\rho$ since $x \neq u$; similarly, if $(\rho, x)$ is produced by $(\sigma, u)$ using R(5), then $x (\neq u)$ must be matched to $u^* (\neq x^*)$ in $\rho$.

Hence, $(\rho, x)$ cannot be produced by R(1), R(2) or R(5), and at most three remaining rules can produce it. It follows from Lemma 3.3 that $|f^{-1}(\rho, x)| \leq 3$. ■

Case (2): $x$ is not matched to $x^*$ in $\rho$

**Observation 3.3 (Unused Rule)** For $(\rho, x) \in Q$, if $x$ is not matched to $x^*$ in $\rho$, then $(\rho, x)$ cannot be produced by applying R(4).

Out of the remaining 5 rules, we show that $(\rho, x)$ can be produced from at most one of $\{R(2), R(5)\}$, and at most two of $\{R(1), R(3), R(6)\}$. After we show these two lemmas, we can immediately conclude from Lemma 3.3 that $|f^{-1}(\rho, x)| \leq 3$ and complete the case analysis.

**Lemma 3.5 (One in $\{R(2), R(5)\}$)** Each $(\rho, x) \in Q$ cannot be produced from both R(2) and R(5).

**Proof:** Suppose the opposite is true: $(\sigma_1, u_1)$ produces $(\rho, x)$ according to R(2), and $(\sigma_2, u_2)$ produces $(\rho, x)$ according to R(5). This implies that in $\rho$, $x$ is matched to both $u_1^*$ and $u_2^*$, which means $u_1 = u_2$. By Observation 3.2, this means $\sigma_1 = \sigma_2$, which contradicts the fact that the same $(\sigma, u) \in R_n$ cannot use two different rules to produce the same $(\rho, x) \in Q$. ■

**Lemma 3.6 (Two in $\{R(1), R(3), R(6)\}$)** Each $(\rho, x) \in Q$ cannot be produced from all three of R(1), R(3) and R(6).

**Proof:** Assume the opposite is true. Suppose $(\sigma_1, u_1)$ produces $(\rho, x)$ using R(1); then, $x = u_1^*$ (hence, $x$ is a neighbor of $u_1$ in $G$) and $u_1$ is unmatched in $\rho$. Suppose $(\sigma_2, u_2)$ produces $(\rho, x)$ using R(3); then, $x = u_2$ is unmatched in $\sigma_2$, and matched in $\rho$. Suppose $(\sigma_3, u_3)$ produces $(\rho, x)$ using R(6); then, $u_3$ is matched in $\rho$, $u_3^*$ is unmatched in $\rho$ and $x$ is a neighbor (in $G$) of $u_3^*$.

By Observation 3.2, all of $u_1$, $u_2$ and $u_3$ are distinct. In particular, observe that $u_1 = x^* = u_2^* \neq u_3^*$; hence, all of $u_1$, $u_2$ and $u_3^*$ are distinct (since $u_2$ is matched in $\rho$, but the other two are not).

Now, suppose we start from $\sigma_2 = \rho^n_x$ and promote $x = u_2$ resulting in $\rho$. Observe that $u_2$ changes from unmatched in $\sigma_2$ to matched in $\rho$, and both $u_1$ and $u_3^*$ are unmatched in $\rho$. By Fact 2.2, at least one of $u_1$ and $u_3^*$ is unmatched in $\sigma_2$; however, both $u_1$ and $u_3^*$ are neighbors of $x = u_2$ (in $G$), which is unmatched in $\sigma_2$. This contradicts that fact that in any permutation, two unmatched nodes cannot be neighbors in $G$.

We have finally finished the case analysis, and can conclude the $f$ has injectivity at most 3, thereby achieving the boundary constraint.
3.4 Lower Bound the Performance Ratio by LP Formulation

Combining all the proved constraints, the following $LP_n$ gives a lower bound on the performance ratio when Ranking is run on a graph with $n$ nodes. It is not surprising that the optimal value of $LP_n$ decreases as $n$ increases (although our proof does not rely on this). In Section 4, we analyze the continuous relaxation $LP_\infty$ in order to give a lower bound for all finite $LP_n$, thereby proving a lower bound on the performance ratio of Ranking.

\[
LP_n \quad \min \quad \frac{1}{\pi} \sum_{i=1}^{n} x_t \\
\text{s.t.} \quad x_1 = 1, \\
\quad x_{t-1} - x_t \geq 0, \quad t \in [2..n] \\
\quad (1 - \frac{t-1}{n}) x_t + \frac{2}{n} \sum_{i=1}^{t-1} x_i \geq 1, \quad t \in [2..n] \\
\quad x_n + \frac{3}{2n} \sum_{i=1}^{n} x_t \geq 1, \\
\quad x_t \geq 0, \quad t \in [n].
\]

4 Analyzing $LP_n$ via Continuous $LP_\infty$ Relaxation

In this section, we analyze the limiting behavior of $LP_n$ by solving its continuous $LP_\infty$ relaxation, which contains both monotone and boundary condition constraints. We develop new duality and complementary slackness characterizations to solve for the optimal value of $LP_\infty$, thereby giving a lower bound on the performance ratio of Ranking.

4.1 Continuous LP Relaxation

To form a continuous linear program $LP_\infty$ from $LP_n$, we replace the variables $x_i$'s with a function variable $z$ that is differentiable almost everywhere in $[0, 1]$. The dual $LD_\infty$ contains a real variable $\gamma$, and function variables $w$ and $y$, where $y$ is differentiable almost everywhere in $[0, 1]$. In the rest of this paper, we use "$\forall \theta$" to denote "for almost all $\theta$", which means for all but a measure zero set.

It is not hard to see that $x_i$ corresponds to $z(\frac{i}{n})$, but perhaps it is less obvious how $LD_\infty$ is formed. We remark that one could consider the limiting behavior of the dual of $LP_n$ to conclude that $LD_\infty$ is the resulting program. We show in Section 4.2 that the pair $(LP_\infty, LD_\infty)$ is actually a special case of a more general class of primal-dual continuous LP. However, we first show in Lemma 4.1 that $LP_\infty$ is a relaxation of $LP_n$.

\[
LP_\infty \quad \min \int_0^1 z(\theta)d\theta \\
\text{s.t.} \quad z(0) = 1 \\
\quad z'(\theta) \leq 0, \quad \forall \theta \in [0, 1] \\
\quad (1 - \theta)z(\theta) + 2 \int_0^\theta z(\lambda)d\lambda \geq 1, \quad \forall \theta \in [0, 1] \\
\quad z(1) + \frac{3}{2} \int_0^1 z(\theta)d\theta \geq 1 \\
\quad z(\theta) \geq 0, \quad \forall \theta \in [0, 1].
\]

\[
LD_\infty \quad \max \int_0^1 w(\theta)d\theta + \gamma - y(0) \\
\text{s.t.} \quad (1 - \theta)w(\theta) + 2 \int_0^\theta w(\lambda)d\lambda + \frac{3\gamma}{2} + y'(\theta) \leq 1, \quad \forall \theta \in [0, 1] \\
\quad \gamma - y(1) \leq 0 \\
\quad \gamma, y(\theta), w(\theta) \geq 0, \quad \forall \theta \in [0, 1].
\]

Lemma 4.1 (Continuous LP Relaxation) For every feasible solution $x$ in $LP_n$, there exists a feasible solution $z$ in $LP_\infty$ such that $\int_0^1 z(\theta)d\theta = \frac{1}{\pi} \sum_{i=1}^{n} x_t$. In particular, the optimal value of $LP_n$ is at least the optimal value of $LP_\infty$. 

9
Proof: Suppose $x$ is a feasible solution to $\text{LP}_n$. Define a step function $z$ in interval $[0,1]$ as follows: $z(0) := 1$ and $z(\theta) := x_t$ for $\theta \in \left(\frac{t-1}{n}, \frac{t}{n}\right]$ and $t \in \{1, \ldots, n\}$. It follows that

$$\int_0^1 z(\theta) d\theta = \sum_{t=1}^n \int_{\theta}^{\frac{t}{n}} z(\theta) d\theta = \frac{1}{n} \sum_{t=1}^n x_t.$$  

We now prove that $z$ is feasible in $\text{LP}_n$. Clearly $z(0) = 1$ and $z'(\theta) = 0$ for $\theta \in [0,1] \setminus \left\{ \frac{t}{n} : 0 \leq t \leq n, t \in \mathbb{Z} \right\}$. For every $\theta \in (0,1]$, suppose $\theta \in \left(\frac{t-1}{n}, \frac{t}{n}\right]$, and we have

$$(1 - \theta)z(\theta) + 2 \int_0^\theta z(\lambda) d\lambda = (1 - \theta)x_t + 2 \sum_{i=1}^{t-1} \int_{\theta}^{\frac{i}{n}} z(\lambda) d\lambda + 2 \int_{\frac{t-1}{n}}^{\frac{t}{n}} z(\theta) d\theta$$

$$= (1 - \theta)x_t + 2 \sum_{i=1}^{t-1} \frac{x_i}{n} + 2 \left( \theta - \frac{t-1}{n} \right) x_t$$

$$= (1 - \frac{t-1}{n} + (\theta - \frac{t-1}{n}))x_t + 2 \frac{t-1}{n} \sum_{i=1}^{t-1} x_i$$

$$\geq (1 - \frac{t-1}{n}) x_t + 2 \frac{t-1}{n} \sum_{i=1}^{t-1} x_i$$

$$\geq 1,$$

where the last inequality follows from the feasibility of $x$ in $\text{LP}_n$. The above inequality holds trivially at $\theta = 0$. For the last constraint, using the fact that $\int_0^1 z(\theta) d\theta = \frac{1}{n} \sum_{t=1}^n x_t$ we have

$$z(1) + \frac{3}{2} \int_0^1 z(\theta) d\theta = x_n + \frac{3}{2n} \sum_{t=1}^n x_t \geq 1,$$

where the last inequality follows from the feasibility of $x$ in $\text{LP}_n$. \hfill \blacksquare

4.2 Primal-Dual for a General Class of Continuous LP

We study a class of continuous linear program $\text{CP}$ that includes $\text{LP}_\infty$ as a special case. In particular, $\text{CP}$ contains monotone and boundary conditions as constraints. Let $K, L > 0$ be two real constants. Let $A, B, C, F$ be measurable functions on $[0,1]$. Let $D$ be a non-negative measurable function on $[0,1]^2$. We describe $\text{CP}$ and its dual $\text{CD}$, following which we present weak duality and complementary slackness conditions. In $\text{CP}$, the variable is a function $z$ that is differentiable almost everywhere in $[0,1]$; in $\text{CD}$, the variables are a real number $\gamma$, and measurable functions $w$ and $y$, where $y$ is differentiable almost everywhere in $[0,1]$.

\[
\begin{align*}
\text{CP} & \quad \min & p(z) = \int_0^1 A(\theta)z(\theta) d\theta & \max & d(w, y, \gamma) = \int_0^1 C(\theta)w(\theta) d\theta + L\gamma - Ky(0) \\
\text{s.t.} & & z(0) = K & \quad \text{s.t.} & B(\theta)w(\theta) + \int_0^\theta D(\lambda, \theta)w(\lambda) d\lambda \\
& & z'(\theta) \leq 0, \forall \theta \in [0,1] & & \quad + F(\theta)\gamma + y'(\theta) \leq A(\theta), \forall \theta \in [0,1] \\
& & B(\theta)z(\theta) + \int_0^\theta D(\theta, \lambda)z(\lambda) d\lambda & & \quad \gamma - y(1) \leq 0 \\
& & \geq C(\theta), \forall \theta \in [0,1] & & \quad \gamma, y(\theta), w(\theta) \geq 0, \forall \theta \in [0,1]. \\
& & z(1) + \int_0^1 F(\theta)z(\theta) d\theta \geq L & & \quad (4.3) \\
& & z(\theta) \geq 0, \forall \theta \in [0,1]. & & \quad (4.4)
\end{align*}
\]

Lemma 4.2 (Weak Duality and Complementary Slackness) Suppose $z$ and $(w, y, \gamma)$ are feasible solutions to $\text{CP}$ and $\text{CD}$ respectively. Then, $d(w, y, \gamma) \leq p(z)$. Moreover, suppose $z$ and
(w, y, γ) satisfy the following complementary slackness conditions:

\[ z'(\theta)g(\theta) = 0, \quad \forall \theta \in [0, 1] \]  

\[ B(\theta)z(\theta) + \int_0^\theta D(\theta, \lambda)z(\lambda)d\lambda - C(\theta) \]

\[ w(\theta) = 0, \quad \forall \theta \in [0, 1] \]  

\[ z(1) + \int_0^1 F(\theta)z(\theta)d\theta - L = 0 \]

\[ (\gamma - y(1))z(1) = 0. \]

Then, \( z \) and \((w, y, \gamma)\) are optimal to CP and CD, respectively, and achieve the same optimal value.

**Proof:** Using the primal and dual constraints, we obtain

\[
d(w, y, \gamma) = \int_0^1 C(\theta)w(\theta)d\theta + L\gamma - Ky(0)
\]

\[
\leq \int_0^1 \left[ B(\theta)z(\theta) + \int_0^\theta D(\theta, \lambda)z(\lambda)d\lambda \right] w(\theta)d\theta + L\gamma - Ky(0) \quad \text{by (4.3)}
\]

\[
= \int_0^1 \left[ B(\theta)w(\theta) + \int_0^\theta D(\lambda, \theta)w(\lambda)d\lambda \right] z(\theta)d\theta + L\gamma - Ky(0) \quad \text{(\*)}
\]

\[
\leq \int_0^1 \left[ A(\theta) - F(\theta)\gamma - y'(\theta) \right] z(\theta)d\theta + L\gamma - Ky(0) \quad \text{by (4.5)}
\]

\[
= \int_0^1 A(\theta)z(\theta)d\theta - \int_0^1 y'(\theta)z(\theta)d\theta + \left[ L - \int_0^1 F(\theta)z(\theta)d\theta \right] \gamma - Ky(0)
\]

\[
\leq \int_0^1 A(\theta)z(\theta)d\theta - \int_0^1 y'(\theta)z(\theta)d\theta + z(1)\gamma - Ky(0) \quad \text{by (4.4)}
\]

\[
= \int_0^1 A(\theta)z(\theta)d\theta - y(1)z(1) + y(0)z(0) + \int_0^1 z'(\theta)y(\theta)d\theta + z(1)\gamma - Ky(0) \quad \text{(**)}
\]

\[
\leq \int_0^1 A(\theta)z(\theta)d\theta + (\gamma - y(1))z(1) \quad \text{by (4.1), (4.2)}
\]

\[
\leq \int_0^1 A(\theta)z(\theta)d\theta \quad \text{by (4.6)}
\]

\[
= p(z),
\]

where in \(\text{(*)}\) we change the order of integration by using Tonelli’s Theorem on non-negative measurable function \(g: \int_0^1 \int_0^\theta g(\theta, \lambda)d\lambda d\theta = \int_0^1 \int_\theta^1 g(\lambda, \theta)d\lambda d\theta\); and in \(\text{(**)}\) we use integration by parts. Moreover, if \(z\) and \((w, y, \gamma)\) satisfy conditions (4.7) – (4.11), then all the inequalities above hold with equality. Hence, \(d(w, y, \gamma) = p(z)\); so \(z\) and \((w, y, \gamma)\) are optimal to CP and CD, respectively.

**4.3 Lower Bound for the Performance Ratio**

The performance ratio of Ranking is lower bounded by the optimal value of \(\text{LP}_\infty\). We analyze this optimal value by applying the primal-dual method to \(\text{LP}_\infty\). In particular, we construct a primal feasible solution \(z\) and a dual feasible solution \((w, y, \gamma)\) that satisfy the complementary slackness conditions presented in Lemma 4.2. Note that \(\text{LP}_\infty\) and \(\text{LD}_\infty\) are achieved from CP and CD by setting \(K := 1\), \(L := 1\), \(A(\theta) := 1\), \(B(\theta) := 1 - \theta\), \(C(\theta) := 1\), \(D(\theta) := 2\), \(F(\theta) := 3\).

We give some intuition on how \(z\) is constructed. An optimal solution to \(\text{LP}_\infty\) should satisfy the primal constraints with equality for some \(\theta\). Setting the constraint \((1 - \theta)z(\theta) + 2\int_0^\theta z(\lambda)d\lambda \geq 1\) to equality we get \(z(\theta) = 1 - \theta\). However this function violates the last constraint \(z(1) + \frac{3}{2} \int_0^1 z(\theta)d\theta \geq 1\). Since \(z\) is decreasing, we need to balance between \(z(1)\) and \(\int_0^1 z(\theta)d\theta\).

The intuition is that we set \(z(\theta) := 1 - \theta\) for \(\theta \in [0, \mu]\) and allow \(z\) to decrease until \(\theta\) reaches some value \(\mu \in (0, 1)\), and then \(z(\theta) := 1 - \mu\) stays constant for \(\theta \in [\mu, 1]\). To determine the value of \(\mu\),
note that the equation \( z(1) + \frac{3}{2} \int_0^1 z(\theta) d\theta = 1 \) should be satisfied, since otherwise we could construct a feasible solution with smaller objective value by decreasing the value of \( z(\theta) \) for \( \theta \in (\mu, 1) \). It follows that \((1 - \mu) + \frac{3}{2} \left(1 - \mu + \frac{\mu^2}{2} \right) = 1\), that is, the value of \( \mu \in (0, 1) \) is determined by the equation \(3\mu^2 - 10\mu + 6 = 0\).

After setting \( z \), we construct \((w, y, \gamma)\) carefully to fit the complementary slackness conditions. Formally, we set \( z \) and \((w, y, \gamma)\) as follows with their graphs on the right hand side:

\[
\begin{align*}
z(\theta) &= \begin{cases} 1 - \theta, & 0 \leq \theta \leq \mu \\ 1 - \mu, & \mu < \theta \leq 1 \end{cases} \\
w(\theta) &= \begin{cases} \frac{2(1-\mu)^2}{(5-3\mu)(1-\theta)^2}, & 0 \leq \theta \leq \mu \\ 0, & \mu < \theta \leq 1 \end{cases} \\
y(\theta) &= \begin{cases} 0, & 0 \leq \theta \leq \mu \\ \frac{2(\theta - \mu)}{5-3\mu}, & \mu < \theta \leq 1 \end{cases} \\
\gamma &= \frac{2(1-\mu)}{5-3\mu},
\end{align*}
\]

where \( \mu = \frac{5 - \sqrt{7}}{3} \) is a root of the equation \(3\mu^2 - 10\mu + 6 = 0\).

Figure 4.1: Optimal \( z \) and \((w, y, \gamma)\)

**Lemma 4.3 (Optimality of \( z \) and \((w, y, \gamma)\))** The solutions \( z \) and \((w, y, \gamma)\) constructed above are optimal to LP<sub>∞</sub> and LD<sub>∞</sub>, respectively. In particular, the optimal value of LP<sub>∞</sub> is \( \frac{2(5 - \sqrt{7})}{9} \approx 0.523 \).

**Proof:** We list the complementary slackness conditions and check that they are satisfied by \( z \) and \((w, y, \gamma)\). Then Lemma 4.2 gives the optimality of \( z \) and \((w, y, \gamma)\).

\begin{align*}
(4.7) & \quad z'(\theta)y(\theta) = 0: \text{ we have } y(\theta) = 0 \text{ for } \theta \in [0, \mu) \text{ and } z'(\theta) = 0 \text{ for } \theta \in (\mu, 1]. \\
(4.8) & \quad \left[ (1 - \theta)z(\theta) + 2 \int_0^\theta z(\lambda) d\lambda - 1 \right] w(\theta) = 0: \text{ we have } (1 - \theta)z(\theta) + 2 \int_0^\theta z(\lambda) d\lambda - 1 = (1 - \theta)^2 + 2(\theta - \mu) - 1 = 0 \text{ for } \theta \in [0, \mu) \text{ and } w(\theta) = 0 \text{ for } \theta \in (\mu, 1]. \\
(4.9) & \quad \left[ z(1) + \frac{3}{2} \int_0^1 z(\theta) d\theta - 1 \right] \gamma = 0: \text{ we have } z(1) + \frac{3}{2} \int_0^1 z(\theta) d\theta - 1 = (1 - \mu) + \frac{3}{2} \left(1 - \mu + \frac{\mu^2}{2}\right) - 1 = 0 \text{ by the definition of } \mu. \\
(4.10) & \quad \left[ (1 - \theta)w(\theta) + 2 \int_0^\theta w(\lambda) d\lambda + \frac{3\gamma}{2} + y'(\theta) - 1 \right] z(\theta) = 0: \text{ for } \theta \in [0, \mu), \text{ we have } (1 - \theta)w(\theta) + 2 \int_0^\theta w(\lambda) d\lambda + \frac{3\gamma}{2} + y'(\theta) - 1 = \frac{2(1-\mu)^2}{(5-3\mu)(1-\theta)^2} + 2 \int_\theta^1 w(\lambda) d\lambda + \frac{3(1-\mu)}{5-3\mu} + 0 - 1 = 0, \\
& \quad \text{and for } \theta \in (\mu, 1], \text{ we have } (1 - \theta)w(\theta) + 2 \int_\theta^1 w(\lambda) d\lambda + \frac{3\gamma}{2} + y'(\theta) - 1 = \frac{3\gamma}{2} + y'(\theta) - 1 = \frac{3(1-\mu)}{5-3\mu} + \frac{2}{5-3\mu} - 1 = 0.
\end{align*}
(4.11) \((\gamma - y(1))z(1) = 0\): we have \\
\[\gamma - y(1) = \frac{2(1-\mu)}{5-3\mu} - \frac{2(1-\mu)}{5-3\mu} = 0.\]

Moreover, the optimal value of \(LP_\infty\) is \(\int_0^1 z(\theta)d\theta = 1 - \mu + \frac{\mu^2}{2} = \frac{2(5-\sqrt{7})}{9} \approx 0.523.\)

**Proof of Theorem 1.1:** The expected ratio of **Ranking** is lower bounded by the optimal value of \(LP_n\). Hence, the theorem follows from Lemmas 4.1 and 4.3.

**References**

[1] Gagan Aggarwal, Gagan Goel, Chinmay Karande, and Aranyak Mehta. Online vertex-weighted bipartite matching and single-bid budgeted allocations. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA ’11, pages 1253–1264. SIAM, 2011.

[2] S. Anand, Naveen Garg, and Amit Kumar. Resource augmentation for weighted flow-time explained by dual fitting. In *SODA*, pages 1228–1241, 2012.

[3] Jonathan Aronson, Martin Dyer, Alan Frieze, and Stephen Suen. Randomized greedy matching. ii. *Random Struct. Algorithms*, 6(1):55–73, January 1995.

[4] Martin E. Dyer and Alan M. Frieze. Randomized greedy matching. *Random Struct. Algorithms*, 2(1):29–46, 1991.

[5] Gagan Goel and Aranyak Mehta. Online budgeted matching in random input models with applications to adwords. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA ’08, pages 982–991, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[6] Gagan Goel and Pushkar Tripathi. Matching with our eyes closed. In *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, FOCS ’12, pages 718–727, Washington, DC, USA, 2012. IEEE Computer Society.

[7] Chinmay Karande, Aranyak Mehta, and Pushkar Tripathi. Online bipartite matching with unknown distributions. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC ’11, pages 587–596, New York, NY, USA, 2011. ACM.

[8] R. M. Karp, U. V. Vazirani, and V. V. Vazirani. An optimal algorithm for on-line bipartite matching. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing*, STOC ’90, pages 352–358, New York, NY, USA, 1990. ACM.

[9] N. Levinson. A class of continuous linear programming problems. *Journal of Mathematical Analysis and Applications*, 16:73–83, 1966.

[10] Mohammad Mahdian and QiQi Yan. Online bipartite matching with random arrivals: an approach based on strongly factor-revealing lps. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC ’11, pages 597–606, New York, NY, USA, 2011. ACM.

[11] Silvio Micali and Vijay V. Vazirani. An \(O(\sqrt{V E})\) algorithm for finding maximum matching in general graphs. In *FOCS*, pages 17–27. IEEE Computer Society, 1980.
[12] Matthias Poloczek and Mario Szegedy. Randomized greedy algorithms for the maximum matching problem with new analysis. *Foundations of Computer Science, IEEE Annual Symposium on*, 0:708–717, 2012.

[13] Alvin E. Roth, Tayfun Sonmez, and M. Utku Unver. Pairwise kidney exchange. Working Paper 10698, National Bureau of Economic Research, August 2004.

[14] William F. Tyndall. A duality theorem for a class of continuous linear programming problems. *Journal of the Society for Industrial and Applied Mathematics*, 13(3):pp. 644–666, 1965.
Appendix A: Issues with the Experimental Results on $LP(n)$

We ran experiments on the LP described in Section III.B of [6] and obtained the following results. The source code (in MathProg format) is available at:

http://i.cs.hku.hk/~algth/project/online_matching/issue.html.

| $n$  | Ratio   |
|------|---------|
| 20   | 0.5024  |
| 50   | 0.5010  |
| 100  | 0.5005  |
| 200  | 0.5003  |
| 300  | 0.5002  |
| 400  | 0.5001  |

Table 1: Our Experimental Results on $LP(n)$ in [6]

Hence, it is impossible to use $LP(n)$ to show that the performance ratio is larger than 0.5002.

Appendix B: Hardness Result

In this section, we show that we can slightly improve the hardness result in [7] by adjusting the parameter. An example of the graph is shown in B.1. We define the graph as follows:

Let $G$ be a bipartite graph over $2(3 + \epsilon)n$ vertices ($u_i$’s and $v_i$’s). Define the edges by adjacency matrix $A$. ($A[i][j] = 1$ if there is an edge between $u_i$ and $v_j$.)

$$
A[i][j] = \begin{cases} 
1 & \text{if } i = j \\
1 & \text{if } i \in [1, n] \text{ and } j \in (n, (2 + \epsilon)n] \\
1 & \text{if } i \in (n, (2 + \epsilon)n] \text{ and } j \in ((2 + \epsilon)n, (3 + \epsilon)n] \\
0 & \text{otherwise}
\end{cases}
$$

We run experiments on different $n$’s and $\epsilon$’s and get the following result.

| $n$  | $\epsilon = 0.63$  | $n$  | $\epsilon = 0.7314$ | $n$  | $\epsilon = 0.7267$ | $n$  | $\epsilon = 0.7253$ | $n$  | $\epsilon = 0.7244$ | $n$  | $\epsilon = 0.7240$ |
|------|------------------|------|-------------------|------|-------------------|------|-------------------|------|-------------------|------|-------------------|
| 20   | 0.7314           | 50   | 0.7267            | 100  | 0.7253            | 200  | 0.7244            | 500  | 0.7240            |

We observe that when $\epsilon = 1 - 1/e$ the ratio is minimized for this kind of graph. It is close to 0.724 in this case. We leave it as future work to analyze it theoretically.