Analyzing Node-Weighted Oblivious Matching Problem via Continuous LP with Jump Discontinuity

Being around two unmatched people means you will always be matched no matter where you go.

T-H. Hubert Chan* Fei Chen* Xiaowei Wu*

Abstract

We prove the first non-trivial performance ratio strictly above 0.5 for the weighted Ranking algorithm on the oblivious matching problem where nodes in a general graph can have arbitrary weights.

We have discovered a new structural property of the ranking algorithm: if a node has two unmatched neighbors, then it will still be matched even when its rank is demoted to the bottom. This property allows us to form LP constraints for both the weighted and the unweighted versions of the problem.

Using a new class of continuous LP, we prove that the ratio for the weighted case is at least 0.501512, and improve the ratio for the unweighted case to 0.526823 (from the previous best 0.523166 in SODA 2014). Unlike previous continuous LP in which the primal solution must be continuous everywhere, our new continuous LP framework allows the monotone component of the primal function to have jump discontinuities, and the other primal components to take non-conventional forms such as the Dirac delta function.

---

*Department of Computer Science, the University of Hong Kong. {`hubert`,`fchen`,`xwwu`}`@cs.hku.hk`

1 Introduction

We analyze the Ranking algorithm for the (node-weighted) Oblivious Matching Problem on arbitrary graphs [2, 1, 14, 7, 4, 3]. While the classical maximum matching problem [13] is well understood, the oblivious version is motivated by online advertising [6, 1] and exchange settings [15], in which information about the underlying graphs might be unknown. We state the problem formally as follows.

Oblivious Matching Problem. An *adversary* commits to a simple undirected graph $G = (V, E)$, where each node $u \in V$ has non-negative weight w_u. The nodes V (where $n = |V|$) and their weights are revealed to the (randomized) *algorithm*, while the edges E are kept 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) sum of weights of nodes matched by the algorithm to that of a maximum weight matching in G.

Weighted Ranking algorithm. Given the node weights w, the algorithm determines a distribution $\mathcal{D}_w$ on permutations of V. It samples a permutation π from $\mathcal{D}_w$, and returns a list L of unordered pairs according to the lexicographical order induced by π, where nodes appearing earlier in the permutation have higher priority. Specifically, for a permutation $\pi : V \to [n]$, given two pairs e_1 and e_2 (where for each j, $e_j = \{u_j, v_j\}$ and $\pi(u_j) < \pi(v_j)$), the pair e_1 has higher priority than e_2 if (i) $\pi(u_1) < \pi(u_2)$, or (ii) $u_1 = u_2$ and $\pi(v_1) < \pi(v_2)$.

Note that a Ranking algorithm is characterized by how it determines the distribution $\mathcal{D}_w$ on permutations of V. For instance, the (deterministic) greedy algorithm uses the permutation of nodes sorted in non-increasing order of weights; it can be shown that its performance ratio is at least 0.5.

For nodes having uniform weight (also known as the unweighted case), it is known [3] that sampling a permutation on V uniformly at random gives ratio strictly larger than 0.5. The interesting question is whether the result can be extended for the case when the nodes in an arbitrary graph have arbitrary weights.

Background of Problem: Uniform Weight. For uniform weight, Dyer and Frieze [5] showed that the performance ratio is $0.5 + o(1)$ when the permutation of unordered pairs is chosen uniformly at random. In the mid-nineties, Aronson et al. [2] showed that the Modified Randomized Greedy algorithm (MRG) has ratio $0.5 + \epsilon$ (where $\epsilon = \frac{1}{400000}$). For bipartite graphs, a version of the ranking algorithm was first proposed by Karp et al. [10] to solve Online Bipartite Matching with ratio $1 - \frac{1}{e}$, which directly translates to the same ratio for the Oblivious Matching Problem.

Since running Ranking on bipartite graphs for the Oblivious Matching Problem is equivalent to running the ranking algorithm for the Online Bipartite Matching problem with *random arrival order*, the result of Karande et al. [9] implies that the ranking algorithm has a ratio at least 0.653 for the Oblivious Matching Problem on bipartite graphs. Mahdian and Yan [12] improved the ratio to 0.696. Karande et al. [9] also constructed a hard instance in which Ranking performs no better than 0.727

For Oblivious Matching Problem on arbitrary graphs, Poloczek and Szegedy [14] analyzed the MRG algorithm and gave ratio $\frac{1}{2} + \frac{1}{256} \approx 0.5039$. However, from personal communication with the authors, we are told that they are currently bridging some gaps in their proof at the time of writing. Goel and Tripathi [7] 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, but later withdrew the paper [8] when a bug is discovered in their proof. In a recent SODA 2014 paper, Chan et al. [3] proved that Ranking algorithm has performance ratio at least 0.523166. We improve the analysis in this paper.

General Weights: Weighted Ranking. Aggarwal et al. [1] showed that the ranking algorithm can be applied to Online Bipartite Matching when the offline nodes have general weights; they proved that the performance ratio is $1 - \frac{1}{e}$. Devanur et al. [4] gave an alternative proof using randomized primal-dual analysis.

We observe that their analysis can be applied to the node-weighted Oblivious Matching Problem on bipartite graphs. We use Ω to denote the sample space of *configurations* from which the algorithm derives its randomness. Specifically, the weighted Ranking algorithm considers an *adjustment function* $\varphi(t) := 1 - e^{t-1}$ for $t \in [0, 1]$; it samples $\sigma \in \Omega_\infty := [0, 1]^V$ uniformly at random, and uses the nodes sorted in non-increasing order of the adjusted weight $w(\sigma, u) := \varphi(\sigma(u)) \cdot w_u$ as the permutation in our earlier description. We consider a different adjustment function φ in this paper.

Since their analysis assumes that the online nodes arrive in arbitrary order, by exchanging the roles of online and offline nodes for both partition of nodes, it can be shown that weighted Ranking achieves the same ratio of $1 - \frac{1}{e}$ on bipartite graphs.

In this paper, we prove that a weighted version of Ranking can achieve ratio strictly larger than 0.5; as far as we know, there is no such result previously in the literature for node-weighted Oblivious Matching Problem on general graphs.

Our Contribution and Results. We first describe the challenges encountered when previous techniques are applied to the node-weighted version of the problem on general graphs.

- *Why is the problem difficult on general graphs (as opposed to bipartite graphs)?* Bipartite graphs have the following nice property. Suppose in configuration σ, node u is unmatched, while its partner u^* in the optimal matching is matched to some node v. If the rank of u is promoted to form configuration σ', then u^* will be matched to some node v' such that the adjusted weight $w(\sigma', v') \geq w(\sigma, v)$ does not decrease. This naturally gives a way to relate the bad instance (σ, u) to the good instance (σ', v'), but unfortunately this property does not hold in general graphs. In fact, u^* might be unmatched in σ' as a result of u's promotion.
- *Why is the problem difficult when nodes have arbitrary weights (as opposed to uniform weight)?* In previous work [3] on unweighted case, when u^* is matched in σ' in the above scenario, it is argued that the bad instance (σ, u) can be related to the good instance (σ', v), where v is matched in σ' to u^*. However, there is no guarantee that the adjusted weight $w(\sigma', v)$ of the good instance is at least $w(\sigma, u)$, which is needed as in [1, 4] to analyze the ratio for the weighted version.

Exploiting structural properties of Ranking. We analyze how the resulting matching would change if the rank of one node is changed (in Lemma 3.3), and give finer classification of good instances. In particular, the following notions are useful for relating bad instances to good instances in order to form LP constraints.

- **Graceful Instance.** A good instance (σ, u) is *graceful* if u is currently matched to a node v such that its optimal partner v^* is also matched. This is similar to the "Type 2 good event" defined in [7]. This idea for keeping track of when both partners v and v^* in the optimal matching are currently matched is also used in [9, 14].
- **Perpetual Instance.** We discover a new structural property of Ranking that if in a good instance (σ, u), node u has two unmatched neighbors, then (σ, u) is *perpetually good* in the sense that u will still be matched even when its rank is demoted to the bottom.

Weighted vs Unweighted. As in [1], we analyze the discrete sample space $\Omega_m := [m]^V$ (with the adjustment function $\varphi(t) := 1 - \frac{e^{17t}-1}{e^{17}-1}$, $\psi(i) := \varphi(\frac{i}{m})$ and adjusted weight $w(\sigma, u) := \psi(\sigma(u)) \cdot w_u$), and show that the performance ratio of weighted Ranking is at least the optimal value of some finite LP_m^{ψ} with m variables. Using similar techniques, we also derive a new finite LP_n^{U}, that gives a lower bound on the performance of unweighted Ranking running on graphs of size n. An important difference is that LP_m^{ψ} does not have a dependence on the size of G, and hence, computing the optimal value of LP_m^{ψ} for some large enough m is sufficient to prove a lower bound on the ratio of weighted Ranking.

Theorem 1.1 (Weighted Ranking with Finite Sampling) *For $m = 10000$, weighted* Ranking *using sample space Ω_m (with adjustment function $\varphi(t) := 1 - \frac{e^{17t}-1}{e^{17}-1}$) has performance ratio at least* 0.501505*.*

Even though we can prove by computing the value of some finite LP that there exists a weighted Ranking algorithm that achieves ratio strictly larger than 0.5, it will be interesting to investigate the limiting behavior as m tends to infinity, because experiments suggest that the value of LP_m^{ψ} increases as m increases. Moreover, for the unweighted version of the problem, the corresponding LP_n^{U} is actually decreasing as n increases, and the limiting behavior has to be considered to give a proof on the ratio.

New Class of Continuous LP with Jump Discontinuity. We develop a new class of continuous LP that generalizes the framework in [3] and contains a unified constraint that can capture both the weighted and the unweighted cases. The primal-dual framework in [3] requires the primal solution to be continuous everywhere, but experiments on the finite LP_m^{ψ} suggest that the optimal primal solution might not be continuous. Hence, we extend our weak duality and complementary slackness characterization to allow jump discontinuity on the primal component on which monotonicity is imposed. For other primal components with no monotonicity constraint, our framework can incorporate non-conventional functions such as the Dirac delta function, which is sometimes useful in our proofs. We use our continuous LP framework to obtain better analysis for both the weighted and the unweighted cases.

Given an adjustment function φ, our techniques can give a lower bound on the ratio in terms of a continuous $\mathsf{LP}_\infty^{\varphi}$. However, at the moment, we have not tried to obtain the best possible φ yet, since the best φ to optimize $\mathsf{LP}_\infty^{\varphi}$ might not necessarily be the best φ to optimize the ratio. Indeed, we are aware of other adjustment functions that can achieve even slightly better ratios, but we just present here one of simple form that can cross the 0.5 barrier.

Theorem 1.2 (Weighted Ranking with Continuous Sampling) *Using continuous sample space Ω_∞ (with adjustment function $\varphi(t) := 1 - \frac{e^{17t}-1}{e^{17}-1}$), weighted* Ranking *has performance ratio at least* 0.501512*.*

Theorem 1.3 (Unweighted Ranking) *Unweighted* Ranking *has performance ratio at least* 0.526823*.*

Our result for the node-weighted case achieves the first non-trivial performance ratio that is strictly larger than 0.5. Although our new theoretical guarantee for the unweighted case has improvement only at the third decimal place over the previous result (0.523 in [3]), we believe our new combinatorial analysis will shed light on the problem and inspire further research in the community. Moreover, our generalized framework of continuous LP provides a powerful tool to analyze the asymptotic behavior as the size of finite LP grows, and it will be of independent interest to explore further applications.

2 Preliminaries

We denote $[m] = \{1, 2, \ldots, m\}$ for any positive integer m. Suppose an adversary commits to a simple undirected graph $G = (V, E)$ with $n = |V|$ nodes, where each node u has a non-negative weight w_u. We fix some maximum weight matching OPT in G. When the context is clear, we also use OPT to denote the set of nodes covered by the matching. Observe that in general OPT might be a proper subset of V. Let $w(\mathsf{OPT}) = \sum_{u \in \mathsf{OPT}} w_u$ be the total weight of OPT. For any $u \in V$, if u is matched in OPT, then we denote by u^* the partner of u in OPT, and we call u^* the *optimal partner* of u. If $u \notin \mathsf{OPT}$, then we say that u^* does not exist.

Weighted Ranking. As described in the introduction, it suffices to describe how the algorithm samples a permutation of nodes, which induces a lexicographical order on the node pairs that is used for probing. As in [1, 4], the algorithm fixes an *adjustment function* $\varphi : [0, 1] \to [0, 1]$ that is non-increasing. The function $\varphi(t) := 1 - e^{t-1}$ is used in [1, 4]. We shall consider other adjustment functions such that $\varphi(1) = 0$ also holds (which is needed for the limiting case).

Let Ω be the sample space of *configurations* from which the algorithm derives its randomness. Let m be a large enough integer. For ease of description, we will mostly consider the discrete space $\Omega_m := [m]^V$. The algorithm samples $\sigma \in \Omega_m$ uniformly at random, which is equivalent to picking $\sigma(u) \in [m]$ uniformly at random and independently for each $u \in V$. We denote $\psi(i) := \varphi(\frac{i}{m})$. Then, a permutation on V is induced by σ by sorting the nodes in non-increasing order of *adjusted weight* $w(\sigma, u) := \psi(\sigma(u)) \cdot w_u$, where ties are resolved deterministically (for instance by the identities of the nodes). We denote $(\sigma, u) > (\sigma, v)$ when node u comes before v in the permutation induced by σ, in which case u has higher priority than v.

In the limiting case when m tends to infinity, σ is drawn from continuous $\Omega_\infty := [0, 1]^V$, and the adjusted weight is given by $w(\sigma, u) := \varphi(\sigma(u)) \cdot w_u$. We omit the subscript for Ω when the context is clear.

We denote $\mathcal{U} := \Omega \times V$ as the set of *instances*. Let $M(\sigma)$ be the matching obtained when $\mathsf{Ranking}$ is run with configuration σ. If u is matched to some v after running $\mathsf{Ranking}$ with configuration σ, then we say that u is matched in σ and v is the (current) partner of u in σ. An instance (σ, u) is *good* if u is matched in σ, and otherwise *bad*. An event is a subset of instances.

Given $\sigma \in \Omega_m$, let σ_u^j be obtained by setting $\sigma_u^j(u) = j$ and $\sigma_u^j(v) = \sigma(v)$ for all $v \neq u$.

Definition 2.1 (Events) *For each $i \in [m]$, define the following:*

- *Rank-i good event: $Q_i := \{(\sigma, u) | \sigma(u) = i$ and u is matched in $\sigma\}$*
- *Rank-i bad event: $R_i := \{(\sigma, u) | \sigma(u) = i, u$ is not matched in σ and $u \in \mathsf{OPT}\}$*

Let $Q := \cup_{i \in [m]} Q_i$ and $R := \cup_{i \in [m]} R_i$.

Notice that Q_i and R_i are disjoint. While Q_i could involve nodes that are not in OPT, R_i only involves nodes in OPT; this idea also appears in [1] for dealing with the case when OPT is a proper subset of V. Define $x_i := \frac{\sum_{(\sigma,u) \in Q_i} w_u}{w(\mathsf{OPT}) \cdot m^{n-1}}$, which can be interpreted as the conditional expected contribution of the nodes given that they are at rank i. We next derive some properties of the x_i's.

- **Monotonicity.** For $i \geq 2$, we have $x_{i-1} \geq x_i \geq 0$, since if $(\sigma, u) \in Q_i$, then $(\sigma_u^{i-1}, u) \in Q_{i-1}$. However, $1 \geq x_1$ does not necessarily hold since there may exist $u \notin \mathsf{OPT}$ and $(\sigma, u) \in Q_1$.
- **Loss due to unmatched nodes.** Similar to x_i associated with Q_i, we consider an analogous quantity associated with R_i:

$$\overline{x}_i := \frac{\sum_{(\sigma,u) \in R_i} w_u}{w(\mathsf{OPT}) \cdot m^{n-1}} = \frac{\sum_{(\sigma,u) \in Q_i \cup R_i} w_u - \sum_{(\sigma,u) \in Q_i} w_u}{w(\mathsf{OPT}) \cdot m^{n-1}} \geq \frac{w(\mathsf{OPT}) \cdot m^{n-1} - \sum_{(\sigma,u) \in Q_i} w_u}{w(\mathsf{OPT}) \cdot m^{n-1}} = 1 - x_i, \quad (2.1)$$

where the inequality $\sum_{(\sigma,u)\in Q_i\cup R_i} w_u \geq w(\mathsf{OPT}) \cdot m^{n-1}$ could be strict because Q_i might involve nodes not in OPT.

- **Performance Ratio.** The performance ratio is the expected sum of weights of matched nodes divided by $w(\mathsf{OPT})$, which is given by $\frac{\sum_{(\sigma,u)\in Q} w_u}{w(\mathsf{OPT})\cdot m^n} = \frac{1}{m}\sum_{i=1}^m x_i$.

Definition 2.2 (Marginally Bad Event) *For $i \in [m]$, we define rank-i marginally bad event as follows. Let $S_1 := R_1$; for $i \geq 2$, let $S_i := \{(\sigma,u) \in R_i | (\sigma_u^{i-1}, u) \in Q_{i-1}\}$.*

Let $S := \cup_{i\in[m]} S_i$ and $\alpha_i := \frac{\sum_{(\sigma,u)\in S_i} w_u}{w(\mathsf{OPT})\cdot m^{n-1}}$ for all $i \in [m]$.

Observe that for an instance (σ,u) such that (σ_u^m, u) is bad, there exists a unique $j \in [m]$ such that $(\sigma_u^j, u) \in S_j$, and we say that j is the *marginal position* of (σ,u).

Relating x_i's and α_i's. From a marginally bad instance $(\sigma,u) \in S_i$, node u will be matched when its rank is promoted to $i-1$. Hence, for $i \geq 2$, we immediately have

$$\alpha_i \leq \frac{\sum_{(\sigma,u)\in Q_{i-1}} w_u - \sum_{(\sigma,u)\in Q_i} w_u}{w(\mathsf{OPT})\cdot m^{n-1}} = x_{i-1} - x_i. \tag{2.2}$$

Moreover, for $i \in [m]$, any bad instance $(\sigma,u) \in R_i$ has a unique marginal position $j \in [i]$ such that $(\sigma_u^j, u) \in S_j$; for each $(\sigma,u) \in S_j$ such that $j \leq i$, we also have $(\sigma_u^i, u) \in R_i$. Hence, there is a one-one correspondence between R_i and $\cup_{j=1}^i S_j$, and so we have:

$$\sum_{j=1}^i \alpha_j = \frac{\sum_{j=1}^i \sum_{(\sigma,u)\in S_j} w_u}{w(\mathsf{OPT})\cdot m^{n-1}} = \frac{\sum_{(\sigma,u)\in R_i} w_u}{w(\mathsf{OPT})\cdot m^{n-1}} = \overline{x}_i \geq 1 - x_i. \tag{2.3}$$

Remark. Observe that when all nodes in V are covered by OPT, equality holds for both (2.2) and (2.3). In fact, the following lemma allow us to remove the α_i's from the LP constraints.

Lemma 2.1 *Suppose that $\{b_i\}_{i=1}^{m+1}$ is non-negative and non-increasing such that $b_{m+1} = 0$, and $\{c_i\}_{i=1}^{m+1}$ is non-decreasing such that $c_1 = 0$. Then, we have*

(a) $\sum_{i=1}^m b_i\alpha_i \geq b_1 - \sum_{i=1}^m (b_i - b_{i+1})x_i$.
(b) $\sum_{i=1}^m b_i c_i \alpha_i \geq -\sum_{i=1}^m (b_i c_i - b_{i+1}c_{i+1})x_i$.

Proof: Statement (a) follows because

$$\sum_{i=1}^m b_i\alpha_i = \sum_{i=1}^m (b_i - b_{i+1})\sum_{j=1}^i \alpha_j \geq \sum_{i=1}^m (b_i - b_{i+1})(1 - x_i) = b_1 - \sum_{i=1}^m (b_i - b_{i+1})x_i,$$

where the inequality comes from (2.3).

For statement (b), observing that $c_1 = 0$, we can assume that $\alpha_1 = x_0 - x_1$, where $x_0 = 1$. Let $C = \max_i c_i$, and define $d_i := C - c_i \geq 0$. Then, we have

$$\sum_{i=1}^m b_i c_i \alpha_i = \sum_{i=1}^m C b_i \alpha_i - \sum_{i=1}^m b_i d_i \alpha_i \geq C b_1 - C\sum_{i=1}^m (b_i - b_{i+1})x_i - \sum_{i=1}^m b_i d_i (x_{i-1} - x_i) = -\sum_{i=1}^m (b_i c_i - b_{i+1}c_{i+1})x_i,$$

where in the inequality we apply statement (a) to the first term (which is still valid because $\alpha_1 \geq 1 - x_1$ holds), and apply $\alpha_1 = x_0 - x_1$ and (2.2) to the second term. ■

Fact 2.1 ($\mathsf{Ranking}$ is Greedy) *Suppose $\mathsf{Ranking}$ is run with configuration σ. If (σ,u) is bad, then each neighbor of u (in G) is matched in σ to some node v such that $(\sigma,v) > (\sigma,u)$.*

3 Formulating LP Constraints for Weighted Case

In this section we define some relations from (marginally) bad events to good events to formulate our LP constraints. We describe a general principle which is a weighted version of the argument used in [3].

Weighting Principle. Suppose f is a relation from subset A to subset B of instances, where $f(a)$ is the set of elements in B that is related to $a \in A$, and $f^{-1}(b)$ is the set of elements in A that is related to $b \in B$. Recall that each instance $a = (\sigma, u)$ has adjusted weight $w(a) = w(\sigma, u)$. Suppose further that for all $a \in A$, for all $b \in f(a)$, $w(a) \leq w(b)$. Then, by considering the bipartite graph H induced by f on $A \cup B$, and comparing the weights of end-points for each edge in H, it follows that $\sum_{a \in A} |f(a)| \cdot w(a) \leq \sum_{b \in B} |f^{-1}(b)| \cdot w(b)$.

We shall formulate constraints by considering relations between subsets of instances. The injectivity of a relation f is the minimum integer q such that for all $b \in B$, $|f^{-1}(b)| \leq q$. In this case, we have

$$\textstyle\sum_{a \in A} |f(a)| \cdot w(a) \leq q \sum_{b \in B} w(b). \tag{3.1}$$

3.1 Demoting Marginally Bad Instances

Lemma 3.1 *We have:* $\frac{1}{m} \sum_{i=1}^{m} [2\psi(i) + (m-i)(\psi(i) - \psi(i+1))] x_i \geq \psi(1)$.

Proof: We define a relation f from the set S of marginally bad instances to the set Q of good instances. Observe that for a (marginally) bad instance (σ, u), u is unmatched in σ and its optimal partner u^* exists. If we further demote u by setting its rank to $j \geq \sigma(u)$, the resulting matching is unchanged. Therefore, by Fact 2.1, for each $j \geq \sigma(u)$, u^* is matched to the same v such that $w(\sigma, u) \leq w(\sigma, v) = w(\sigma_u^j, v)$. Hence, we can define $f(\sigma, u) := \{(\sigma_u^j, v) | u^* \textit{ is matched to } v \textit{ in } \sigma_u^j, j \geq \sigma(u)\} \subseteq Q$, where $|f(\sigma, u)| = m - \sigma(u) + 1$, and $w(\sigma, u) \leq w(\sigma', v)$ for all $(\sigma', v) \in f(\sigma, u)$.

We next check the injectivity of f. Suppose $(\rho, v) \in f(\sigma, u)$. Then, u^* is the current partner of v in ρ, and this uniquely determines u, which is unmatched in ρ. Hence, $\sigma = \rho_u^j$, where j is uniquely determined as the marginal position of (ρ, u). Therefore, the injectivity is 1.

Hence, our weighting principle (3.1) gives the following:

$\sum_{i=1}^{m} \sum_{(\sigma,u) \in S_i} (m-i+1)\psi(i) w_u = \sum_{a \in S} |f(a)| \cdot w(a) \leq \sum_{b \in Q} w(b) = \sum_{i=1}^{m} \sum_{(\rho,v) \in Q_i} \psi(i) w_v$.

Dividing both sides by $w(\mathsf{OPT}) \cdot m^n$ gives $\frac{1}{m} \sum_{i=1}^{m} (m-i+1)\psi(i)\alpha_i \leq \frac{1}{m} \sum_{i=1}^{m} \psi(i) x_i$.

Since we do not wish α_i's to appear in our constraints, we derive a lower bound for the LHS in terms of x_i's by applying Lemma 2.1 with $b_i := (m-i+1)\psi(i)$, where $\psi(m+1)$ can be chosen to be any value. Rearranging gives the required inequality. ■

3.2 Promoting Marginally Bad Instances

Lemma 3.2 *We have:* $\frac{2}{m} \sum_{i=1}^{m} \psi(i) \cdot x_m + \frac{1}{m} \sum_{i=1}^{m} [5\psi(i) - i(\psi(i+1) - \psi(i))] \cdot x_i \geq \frac{3}{m} \sum_{i=1}^{m} \psi(i)$.

To use the weighting principle, we shall define relations from marginally bad instances S to the following subsets of special good instances.

Definition 3.1 (*If* v *is matched, would* v^* *still be matched?*) *For* $i \in [m]$, *let the graceful instances be* $Y_i := \{(\sigma, u) \in Q_i | u$ *is matched in* σ *to some* v *s.t.* v^* *does not exist or is also matched in* $\sigma\}$. *Let* $y_i := \frac{\sum_{(\sigma,u) \in Y_i} w_u}{w(\mathsf{OPT}) \cdot m^{n-1}}$ *and* $Y := \cup_{i \in [m]} Y_i$.

Definition 3.2 (*You will be matched even at the bottom.*) *For $i \in [m]$, let the perpetual instances be $Z_i = \{(\sigma, u) \in Q_i | (\sigma_u^m, u) \in Q_m\}$. Let $z_i = \frac{\sum_{(\sigma,u)\in Z_i} w_u}{w(OPT)\cdot m^{n-1}}$ and $Z := \cup_{i\in[m]} Z_i$.*

By definition, we know that $Y_i \subseteq Q_i$ and hence $x_i \geq y_i \geq 0$. Moreover, observing that there exists a bijection between Z_i and Q_m that maps each $(\sigma, u) \in Z_i$ to $(\sigma_u^m, u) \in Q_m$, we have $z_i = x_m$.

Suppose (σ, u) is a good instance that has marginal position j. We wish to compare the matchings produced by σ and σ_u^j. Sometimes it is more convenient to consider an unmatched node as being ignored. Specifically, given a configuration σ and a node u, running Ranking with σ_u means that we still use σ to generate the probing order, but any edge involving u is ignored. Observe that if (σ, u) has a marginal position j, then σ_u and σ_u^j will produce the same matching.

Lemma 3.3 (Ignoring One Node.) *Suppose u is covered by the matching $M(\sigma)$ produced by σ, and $M(\sigma_u)$ is the matching produced by using the same probing list, but any edge involving u is ignored. The symmetric difference $M(\sigma) \oplus M(\sigma_u)$ is an alternating path $P = (u = u_1, u_2, \ldots, u_p)$ such that for all $i \in [p-2]$, $(\sigma, u_i) > (\sigma, u_{i+2})$.*

Proof: We can view probing G with σ_u as using the same list L of unordered node pairs to probe another graph G_u, which is the same as G except that the node u is labelled *unavailable* and will not be matched in any case. After each round of probing, we compare what happens to the partially constructed matchings $M(\sigma)$ in G and $M(\sigma_u)$ in G_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 $M(\sigma) \oplus M(\sigma_u)$ is an alternating path P connecting u to the current crucial node; initially, both $M(\sigma)$ and $M(\sigma_u)$ are empty, and path P is degenerate and contains only u. (iii) If the path $P = (u = u_1, u_2, \ldots, u_l)$ contains $l \geq 3$ nodes, then for all $i \in [l-2]$, then $(\sigma, u_i) > (\sigma, u_{i+2})$.

Consider the inductive step. Suppose currently the alternating path $M(\sigma) \oplus M(\sigma_u)$ contains l nodes, where u_l is crucial. Observe that the crucial node and $M(\sigma) \oplus M(\sigma_u)$ 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 u_l with another currently unmatched node u_{l+1} in G, which is also unmatched in G_u (because the induction hypothesis states that all nodes but u_l have the same availability status in G and G_u).

Since u_l has different availability in G and G_u, but u_{l+1} is unmatched in both G and G_u, it follows that the edge $e := \{u_l, u_{l+1}\}$ is added to exactly one of $M(\sigma)$ and $M(\sigma_u)$. Hence, the edge e is added to extend the alternating path $M(\sigma) \oplus M(\sigma_u)$, and the node u_{l+1} becomes crucial.

Next, it remains to show that if $l \geq 2$, then $(\sigma, u_{l-1}) > (\sigma, u_{l+1})$. Suppose we go back in time, and consider at the beginning of the round when the edge $\{u_{l-1}, u_l\}$ is about to be probed, and u_{l-1} is crucial. By the induction hypothesis, both u_l and u_{l+1} are unmatched in both G and G_u. It follows that $(\sigma, u_{l-1}) > (\sigma, u_{l+1})$, because otherwise the edge $\{u_{l-1}, u_l\}$ would have lower probing priority than $\{u_{l+1}, u_l\}$. This completes the inductive step. ■

Lemma 3.4 (Two Unmatched Neighbors Implies Perpetual) *Suppose in configuration σ, node u is matched and has two unmatched neighbors in G. Then, $(\sigma, u) \in Z$ is perpetual.*

Proof: If we assume the opposite, then u will be unmatched in σ_u^m. Suppose x and y are two neighbors of u that are unmatched in σ. Then, by Lemma 3.3, the symmetric difference

$M(\sigma) \oplus M(\sigma_u^m)$ is an alternating path starting from u, and hence at most one of x and y will remain unmatched in σ_u^m.

This implies that in σ_u^m, the unmatched node u will have at least one unmatched neighbor; this contradicts the fact that that Ranking will always produce a maximal matching. ■

Next we derive inequalities involving the graceful instances. Combining the inequalities, we can obtain the crucial constraint involving only x_i's for achieving a ratio that is strictly larger than 0.5.

Lemma 3.5 (***You are unmatched because someone is not graceful.***) *We have the following inequality:* $\frac{1}{m}\sum_{i=1}^m \psi(i) y_i \leq \frac{1}{m}\sum_{i=1}^m \psi(i)(2x_i - 1)$.

Proof: We define a relation from the set R of bad instances to the set $Q \setminus Y$ of good instances that are not graceful.

Given any bad instance $(\sigma, u) \in R$, we know that u^* exists and is matched to some node v such that $w(\sigma, v) \geq w(\sigma, u)$, by Fact 2.1. Moreover, since v is matched to u^* such that u is unmatched, we know that $(\sigma, v) \in Q \setminus Y$ is good but not graceful. Hence, we define $f(\sigma, u) := \{(\sigma, v)\}$, where v is the current partner of u^*. Observe that each $(\sigma, v) \in Q \setminus Y$ can be related to a unique $(\sigma, u) \in R$, where u is the optimal partner of v's current partner in σ. Hence, the injectivity of f is 1.

Hence, the weighting principle (3.1) gives: $\sum_{(\sigma,u)\in R} w(\sigma, u) \leq \sum_{(\sigma,v)\in Q\setminus Y} w(\sigma, v)$. Dividing both sides by $w(\mathsf{OPT}) \cdot m^n$ gives: $\frac{1}{m}\sum_{i=1}^m \psi(i)\overline{x}_i \leq \frac{1}{m}\sum_{i=1}^m \psi(i)(x_i - y_i)$.

Finally, using $\overline{x}_i \geq 1 - x_i$ from (2.1) and rearranging gives the required inequality. ■

Lemma 3.6 (***If you are marginal, someone else is either graceful or perpetual.***) *We have the inequality:* $\frac{1}{m}\sum_{i=1}^m (i-1)\psi(i)\alpha_i \leq \frac{1}{m}\sum_{i=1}^m \psi(i)(3y_i + 2z_i)$.

Proof: As mentioned earlier, we shall define two relations f and g from marginally bad S to graceful Y and perpetual Z, respectively, such that the following properties hold.

1. For each $a \in S$, for each $b \in f(a) \cup g(a)$, $w(a) \leq w(b)$.
2. For each $a \in S$, $|f(a)| + |g(a)| = \sigma(u) - 1$.
3. The injectivity of f is at most 3 and the injectivity of g is at most 2.

Suppose we have f and g with these properties. Then, our weighting principle (3.1) gives:

$$\sum_{(\sigma,u)\in S}(\sigma(u) - 1)w(\sigma, u) \leq \sum_{(\rho,v)\in Y} 3w(\rho, v) + \sum_{(\rho,v)\in Z} 2w(\rho, v),$$

which by definition is equivalent to

$$\sum_{i=1}^m (i-1)\psi(i)\sum_{(\sigma,u)\in S_i} w_u \leq \sum_{i=1}^m \psi(i)(3\sum_{(\rho,v)\in Y_i} w_u + 2\sum_{(\rho,v)\in Z_i} w_u).$$

Dividing both sides by $w(\mathsf{OPT}) \cdot m^n$ gives the required inequality.

Next we show how f and g are constructed such that all required properties hold.

Given marginally bad $(\sigma, u) \in S$, we consider good instance $(\sigma', u) \in Q$, where $\sigma' = \sigma_u^j, j < \sigma(u)$ is obtained by "promoting" u's rank in σ. Note that by Fact 2.1, u^* must be matched in σ to some node v_0 such that $(\sigma, v_0) > (\sigma, u)$. Let the partner of u in (σ', u) be p. The next claim is crucial for the construction of f and g.

Claim 3.1 *If $w(\sigma', p) < w(\sigma, u)$, then u^* is matched in σ' to some node v such that $w(\sigma', v) \geq w(\sigma, v_0) \geq w(\sigma, u)$.*

Proof: By Lemma 3.3, we know that the symmetric difference $M(\sigma') \oplus M(\sigma)$ is an alternating path $(u = u_1, p = u_2, u_3, u_4 \ldots)$ that starts with u. Moreover, we have $w(\sigma', u) \geq w(\sigma', u_3) \geq w(\sigma', u_5) \geq \ldots$ and $w(\sigma', p) \geq w(\sigma', u_4) \geq w(\sigma', u_6) \geq \ldots$. If u^* is not contained in the alternating path, then directly we have $v = v_0$ and hence the claim holds.

Assume that u^* is contained in the alternating path. Then, v_0 must also appear in the alternating path. Let $v_0 = u_i$. Since $w(\sigma', v_0) = w(\sigma, v_0) \geq w(\sigma, u) > w(\sigma', p)$, we conclude that i must be odd. By Lemma 3.3, we know that u^* must be u_{i-1} since u_i is matched to u_{i-1} in σ. Moreover, we know that $u^* = u_{i-1}$ is matched to u_{i-2} in σ' such that $w(\sigma', u_{i-2}) \geq w(\sigma', u_i) = w(\sigma, v_0)$. ■

Next we include instances in Y into $f(\sigma, u)$ and instances in Z into $g(\sigma, u)$ on a case by case basis. Recall that for each $1 \leq j < \sigma(u)$, we consider $\sigma' = \sigma_u^j$; moreover, after promoting u to rank j, u is matched in σ' to p.

Case-1(a). u^* is matched in σ' and $w(\sigma', p) = w(\sigma, p) \geq w(\sigma, u)$.

In this case, (σ', p) is graceful, because p is matched in σ' to u, whose optimal partner u^* is also matched. Hence, we include $(\sigma', p) \in Y$ in $f(\sigma, u)$.

Case-1(b). u^* is matched in σ' and $w(\sigma', p) = w(\sigma, p) < w(\sigma, u)$.

By Claim 3.1, u^* is matched in σ' to some node v such that $w(\sigma', v) \geq w(\sigma, u)$. Observe that (σ', v) is graceful, and we include $(\sigma', v) \in Y$ in $f(\sigma, u)$.

Case-2(a). u^* is unmatched in σ', and p^* (if it exists) is also matched in σ'.

Note that after promoting u, we have $w(\sigma', u) \geq w(\sigma, u)$. Moreover, (σ', u) is graceful, because the optimal partner p^* either does not exist or is matched in σ'. We include $(\sigma', u) \in Y$ in $f(\sigma, u)$.

Case-2(b). u^* is unmatched in σ', p^* exists and is the only unmatched neighbor of p in σ'.

By Claim 3.1, since u^* is unmatched in σ', we have $w(\sigma, p) = w(\sigma', p) \geq w(\sigma, u)$; also, since p is matched in σ', $p \neq u^*$. Moreover, by Lemma 3.3, the symmetric difference $M(\sigma) \oplus M(\sigma')$ is an alternating path, and only two nodes (u and u^*) can have different matching status in σ and σ'.

Hence, in σ, p must remain matched and p^* must remain unmatched; this means that p has exactly two unmatched neighbors, namely u and p^*, in σ. By Lemma 3.4, we conclude that (σ, p) is perpetual, and include $(\sigma, p) \in Z$ in $g(\sigma, u)$.

Case-2(c). u^* is unmatched in σ', p^* exists and is not the only unmatched neighbor of p in σ'.

Similar to Case-2(b), in this case, $w(\sigma', p) = w(\sigma, p) \geq w(\sigma, u)$ and p has two different unmatched neighbors in σ', so (σ', p) is perpetual by Lemma 3.4. We include $(\sigma', p) \in Z$ in $g(\sigma, u)$.

By construction, property 1 holds. Moreover, for each $1 \leq j < \sigma(u)$ and $\sigma' = \sigma_u^j$, exactly one of the above 5 cases happens. Hence, we also have property 2: $|f(\sigma, u)| + |g(\sigma, u)| = \sigma(u) - 1$. Next, we prove the injectivity.

Injectivity Analysis. Observe that in our construction, if $(\rho, v) \in f(\sigma, u) \cup g(\sigma, u)$, then $\sigma = \rho_u^t$, where t is the marginal position of (ρ, u). Hence, in the injectivity analysis, once (ρ, v) and u are identified, σ can be uniquely determined.

For relation f, suppose $(\rho, v) \in Y$ is included in some $f(\sigma, u)$ in the following cases.

Case-1(a). Node u is uniquely identified as the current partner of v in ρ.

Case-1(b). Node u is uniquely identified as the optimal partner of v's current partner.

Case-2(a). Node u is the same as v.

Hence, each $(\rho, v) \in Y$ is related to at most 3 instances in S, which means that f has injectivity at most 3.

For relation g, suppose $(\rho, v) \in Z$ is included in some $g(\sigma, u)$ in the following cases.

Case-2(b). By construction $\rho = \sigma$, and v has exactly two neighbors that are unmatched in ρ, one of which is v^*. Node u is uniquely identified as the other unmatched neighbor.

Case-2(c). Node u is uniquely identified as the current partner of v in ρ.

Hence, each $(\rho, v) \in Z$ is related to at most 2 instances in S, which means that g has injectivity at most 2. This completes the proof of Lemma 3.6. ■

We can now derive the main constraint of this subsection.

Proof of Lemma 3.2: We start from the inequality in Lemma 3.5. Observing that $z_i = x_m$, and using the upper bound for $\frac{1}{m}\sum_{i=1}^{m} \psi(i) y_i$ in Lemma 3.6, we have $\frac{1}{m}\sum_{i=1}^{m}(i-1)\psi(i)\alpha_i \leq \frac{1}{m}\sum_{i=1}^{m}\psi(i)(6x_i + 2x_m - 3)$.

We next use Lemma 2.1 by setting $b_i := \psi(i)$ and $c_i := i - 1$; observe that $c_1 = 0$, and we set $\psi(m+1) := 0$, which is consistent with $\psi(m) \geq 0 = \psi(m+1)$. Hence, we have the following lower bound for the LHS: $\frac{1}{m}\sum_{i=1}^{m}(i-1)\psi(i)\alpha_i \geq \frac{1}{m}\sum_{i=1}^{m}(\psi(i) + i(\psi(i+1) - \psi(i))) \cdot x_i$.

Rearranging gives the required inequality. ■

3.3 Using LP to Bound Performance Ratio

Putting all achieved constraints on x_i's together, we obtain the following linear program LP_m^{ψ}, which is a lower bound on the performance ratio when weighted Ranking is run with weight adjustment function ψ and sample space $\Omega_m = [m]^V$.

$$
\begin{array}{llll}
\mathsf{LP}_m^{\psi} & \min & \frac{1}{m}\sum_{i=1}^{m} x_i & \\
 & \text{s.t.} & x_i - x_{i+1} \geq 0, & i \in [m-1] \\
\multicolumn{3}{r}{\frac{2}{m}\sum_{i=1}^{m}\psi(i)\cdot x_m + \frac{1}{m}\sum_{i=1}^{m}[5\psi(i) - i(\psi(i+1) - \psi(i))]\cdot x_i \geq \frac{3}{m}\sum_{i=1}^{m}\psi(i)} & (3.2) \\
\multicolumn{3}{r}{\frac{1}{m}\sum_{i=1}^{m}[2\psi(i) + (m-i)(\psi(i) - \psi(i+1))]x_i \geq \psi(1)} & (3.3) \\
 & & x_i \geq 0, & i \in [m].
\end{array}
$$

Achieving ratio strictly larger than 0.5. Observe that LP_m^{ψ} is independent of the size of G. Hence, to obtain a lower bound on the ratio, we can use an LP solver to solve LP_m^{ψ} for some large enough m and some appropriate non-negative non-increasing sequence $\{\psi(i)\}_{i=1}^{m}$. In particular, there exists a weighted Ranking algorithm with ratio strictly above 0.5.

Theorem 3.1 *Using* $m = 10000$ *and* $\psi(i) := 1 - \frac{e^{\frac{17i}{m}} - 1}{e^{17} - 1}$, *the weighted* Ranking *algorithm has performance ratio at least the value given by* LP_m^{ψ}*:* 0.501505.

Although the function $\varphi(t) := 1 - e^{t-1}$ (that is used in [1, 4]) cannot give a ratio better 0.5 from our LP, it is still possible that the function could have good performance ratio. More experimental results and our source code can be downloaded at:

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

Limiting case when m tends to infinity. Experiments show that LP_m^ψ is increasing in m. This suggests that a (slightly) better analysis may be achieved if Ranking samples σ from the continuous space $\Omega_\infty = [0,1]^V$, and uses adjusted weight $w(\sigma, u) := \varphi(\sigma(u)) \cdot w_u$ for each node u.

The variables x_i's are replaced by the function $z(t) := \frac{\sum_{u \in V} \Pr_\sigma[(\sigma,u) \text{ is good}|\sigma(u)=t] \cdot w_u}{w(\mathsf{OPT})}$. Our combinatorial counting argument can be replaced by measure analysis. For instance, $\Omega_\infty = [0,1]^V$ is equipped with the uniform n-dimensional measure, while $z(t)$ has measure of dimension $n-1$. Since we assume that $\psi(m+1) = 0$ in the finite analysis, this corresponds to $\varphi(1) = 0$ in continuous case.

Observe that it is possible to describe a continuous version of the weighting principle using measure theory to derive all the corresponding constraints involving z. However, the formal rigorous proof is out of the scope of this paper, and one can intuitively see that each constraint involving the x_i's translates naturally to a constraint involving z in the limiting case. Hence, the following continuous $\mathsf{LP}_\infty^\varphi$ gives a lower bound on the ratio when Ranking samples continuously, and we analyze it in Section 5.2 as a case study.

$$\begin{array}{rll}
\mathsf{LP}_\infty^\varphi \qquad \min & \int_0^1 z(t)dt & \\
\text{s.t.} \qquad\qquad & z'(t) \leq 0 & \forall t \in [0,1] \\
2\Phi \cdot z(1) + \int_0^1 [5\varphi(t) - t\varphi'(t)]\, z(t)dt & \geq 3\Phi & \\
\int_0^1 [2\varphi(t) - (1-t)\varphi'(t)]\, z(t)dt & \geq \varphi(0) & \\
z(t) & \geq 0 & \forall t \in [0,1] \\
\Phi & = \int_0^1 \varphi(t)dt. &
\end{array}$$

4 Improved LP for Unweighted Case

We show in this section that the technique of keeping track of when both a node and its optimal partner are both matched [14, 7] can be applied to unweighted Oblivious Matching Problem on general graphs to improve the analysis of the previously best ratio of 0.523 in [3].

In the unweighted case, the notation is simpler as in [3]. The sample space Ω is the set of all permutations on V, and Ranking simply samples bijection $\sigma : V \to [n]$ uniformly at random from Ω to obtain a permutation on nodes to derive the lexicographical order on node pairs. For a permutation σ, let σ_u^j be the permutation obtained by moving u to position j while keeping the relative order of other nodes unchanged.

As in [14, Corollary 2], without loss of generality, we can assume that the optimal matching in the graph $G = (V, E)$ matches all nodes in V. For each $i \in [n]$, events are defined similarly.

Definition 4.1 **(Q_i, R_i, Y_i, Z_i, S_i)** *Rank-i good event:* $Q_i = \{(\sigma, u) | \sigma(u) = i$ *and* u *is matched in* $\sigma\}$; *rank-i bad event:* $R_i = \{(\sigma, u) | \sigma(u) = i, u$ *is not matched in* $\sigma\}$; *rank-i extra good event:* $Y_i = \{(\sigma, u) \in Q_i | (\sigma, u^*) \in Q\}$; *rank-i perpetual event* $Z_i = \{(\sigma, u) \in Q_i | (\sigma_u^n, u) \in Q_n\}$; *rank-i marginally bad event:* $S_i = \{(\sigma, u) \in R_i | (\sigma_u^{i-1}, u) \in Q_{i-1}\}$. *Let* $Q = \cup_{i=1}^n Q_i$, $R = \cup_{i=1}^n R_i$, $Y = \cup_{i=1}^n Y_i$, $Z = \cup_{i=1}^n Z_i$ *and* $S = \cup_{i=1}^n S_i$.

For each $i \in [n]$, the variables are defined: $x_i = \frac{|Q_i|}{n!}$, $y_i = \frac{|Y_i|}{n!}$, $z_i = \frac{|Z_i|}{n!}$ and $\alpha_i = \frac{|S_i|}{n!}$. Moreover, under the perfect matching assumption, it is not hard to derive the following equalities (let $x_0 = 1$): $x_1 = 1$, $1 - x_i = \frac{|R_i|}{n!}$, $z_i = x_n$ and $\alpha_i = x_{i-1} - x_i$ for all $i \in [n]$.

Note that at least one of u and u^* must be matched in any permutation σ. Hence, the number of nodes matched in each permutation σ is at least $\frac{n}{2}$. Hence, instances in Y is the "extra gain" above the trivial performance ratio 0.5. A simple counting analysis yields the following lemma.

Lemma 4.1 (Extra Gain) *The performance ratio is* $\frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{2} + \frac{1}{2n}\sum_{i=1}^{n} y_i$.

Lemma 4.2 (Evolving Constraints [3]) *For all* $i \in [n]$, *we have* $\left(1 - \frac{i-1}{n}\right) x_i + \frac{2}{n}\sum_{j=1}^{i-1} x_j \geq 1$.

Next we show a lemma very similar to Lemma 3.6 that provides a lower bound for the extra gain.

Lemma 4.3 (Mixed Constraints) *For all* $i \in [n]$, *we have* $\frac{i}{n} \cdot x_n + \frac{i}{n} \cdot x_i + \frac{1}{n}\sum_{j=1}^{i}(-x_j + 2y_j) \geq 0$.

Proof: The inequality is trivial when $i = 1$. For $i \geq 2$, similar to the proof of Lemma 3.6, it suffices to construct a relation f from $\cup_{j=1}^{i} S_j$ to $\cup_{j=1}^{i} Y_j$, and a relation g from $\cup_{j=1}^{i} S_j$ to $\cup_{j=1}^{i} Z_j$ such that the following properties hold.

1. For each $(\sigma, u) \in S_j$, $|f(\sigma, u)| + |g(\sigma, u)| = j - 1$.
2. The injectivity of f is at most 2 and the injectivity of g is at most 1.

Suppose we have those two relations, then immediately we have $\sum_{j=1}^{i}(j-1)|S_j| \leq \sum_{j=1}^{i}(2|Y_j| + |Z_j|)$. Observing that $z_i = x_n$, dividing the both sides by $n!$ gives the following:

$\sum_{j=1}^{i}(j-1)(x_{j-1} - x_j) \leq \sum_{j=1}^{i}(2y_j + x_n)$, from which the required inequality can be obtained.

Next we show how f and g are constructed. Given marginally bad $(\sigma, u) \in S_j$, where $j \leq i$, we consider the good instance $(\sigma', u) \in Q$, where $\sigma' = \sigma_u^k, k < j$.

From Fact 2.1, note that u^* must be matched in σ to some node v such that $\sigma(v) < \sigma(u)$ (otherwise u will be considered first). We include instances in $\cup_{j=1}^{i} Y_j$ into $f(\sigma, u)$ and instances in $\cup_{j=1}^{i} Z_j$ into $g(\sigma, u)$ as follows.

Case-1. u^* is matched in σ'. Observe that both u and u^* are matched in σ', and $\sigma'(u) \leq i$. Hence, we include extra good $(\sigma', u) \in \cup_{j=1}^{i} Y_j$ in $f(\sigma, u)$

Case-2. u^* is unmatched in σ' and v^* is matched in σ'. Since u^* is unmatched, we know that v must be matched in this case. Since v^* is also matched and $\sigma'(v) \leq \sigma(v) + 1 < \sigma(u) + 1 = j + 1 \leq i + 1$, we include extra good $(\sigma', v) \in \cup_{j=1}^{i} Y_j$ in $f(\sigma, u)$.

Case-3. u^* is unmatched in σ' and v^* is also unmatched in σ'. Since v has two different unmatched neighbors in σ' in this case, (σ', v) is perpetual by Lemma 3.4. We include $(\sigma', v) \in \cup_{j=1}^{i} Z_j$ in $g(\sigma, u)$.

Observe that for each $k < j$, we have $\sigma' = \sigma_u^k$, and exactly 1 of the 3 cases happens. Hence, for $(\sigma, u) \in S_j$, $|f(\sigma, u)| + |g(\sigma, u)| = j - 1$, and so the first property holds.

Injectivity Analysis. We shall verify that each good instance (σ', v) can be included by a unique marginally bad instance (σ, u) for each case. Observe that if u can be identified, then σ can be recovered from σ' by moving u to its marginal position in (σ', u).

Case-1. If $(\sigma', v) \in \cup_{j=1}^{i} Y_j$ is included in Case-1 by marginal bad (σ, u), then u is uniquely identified as v.

Case-2 and 3. Note that if a good instance (σ', v) is included by Case-2 or 3 because of some marginally bad (σ, u), then (σ, u) can be recovered using [3, Lemma 3.3 $\mathsf{R}(6)$]. For completeness, we give an alternative analysis here.

By Lemma 3.3, since u^* is unmatched in σ', the symmetric difference $M(\sigma') \oplus M(\sigma)$ is an alternating path P that starts at u and the last three nodes on the path are w, v and u^*, where node w is the partner of v in σ'. Recall that in the proof of Lemma 3.3, running Ranking with σ' on G_u with node u marked as unavailable is equivalent to running with σ. When we compare running σ' on G and G_u, at any moment, exactly one node on path P is crucial, i.e, it has different availability in G and G_u. Consider the round in which node w is crucial, and the pair $\{w, v\}$ is about to be probed. Node w is matched in G_u, while unmatched in G. At this moment, if we also make w unavailable in G, then the edges included after this point will be the same in both G and G_u; in particular, v will be matched to u^* if we mark w as unavailable in G. Therefore, if we mark the current partner w of v in σ' as unavailable, and still use the same probing order as given by σ', v will be matched to u^*. Hence, we can recover u^* and uniquely identify u.

Therefore, as in the proof of Lemma 3.6, we conclude that the injectivity of f is at most 2 and the injectivity of g is at most 1. This completes the proof of Lemma 4.3. ■

Putting all achieved constraints on x_i's together, we achieve the following linear program LP_n^U, whose optimal value is a lower bound for the performance ratio for Ranking when the input graph has n nodes. To express the LP in a convenient form, we have relaxed the equality in Lemma 4.1 to an inequality.

$$
\begin{array}{lrll}
\mathsf{LP}_n^U & \min & \frac{1}{n}\sum_{i=1}^{n} x_i & \\
 & \text{s.t.} & x_i - x_{i+1} \geq 0 & i \in [n-1] \\
 & & (1 - \frac{i-1}{n})x_i + \frac{2}{n}\sum_{j=1}^{i-1} x_j \geq 1 & i \in [n] \\
 & & \frac{i}{n} \cdot x_n + \frac{i}{n} \cdot x_i + \frac{1}{n}\sum_{j=1}^{i}(-x_j + 2y_j) \geq 0 & i \in [n] \\
 & & \frac{1}{n}\sum_{i=1}^{n}(2x_i - y_i) \geq 1 & \\
 & & x_i, y_i \geq 0 & i \in [n].
\end{array}
$$

As in [3], the value of LP_n^U decreases as n increases. Hence, to give a lower bound on the performance ratio of Ranking, we use continuous LP to analyze the limiting behavior in Section 5.3.

5 Analyzing Finite LP via a General Class of Continuous LP

In this section, we analyze the finite LPs constructed in Sections 3 and 4 via continuous LP to give lower bounds on the performance ratios of (weighted and unweighted) Ranking.

We formulate a general class of continuous LP, and develop new duality and complementary slackness characterization that can capture both the weighted and the unweighted cases. Unlike the previous framework in [3], the functions in the new framework do not have to be continuous everywhere. As inspired from the optimal solutions from our finite LPs, the continuity requirement is weakened such that in the primal LP, there is a component of the function that is monotone and can be allowed to have jump discontinuities. However, as we shall see later, other components of the function may take unconventional form such as the Dirac delta function.

For both the weighted and unweighted cases, we construct dual feasible solutions, the objective values of which give corresponding lower bounds on the performance ratios.

5.1 Primal-Dual for a New Class of Continuous LP

Tyndall [16] and Levinson [11] formulated a class of continuous LP that can handle the evolving constraint. In previous work [3], a class of continuous LP was developed to handle the monotonicity and the boundary constraint. In this paper, we study a new class of continuous linear program CP with a unified constraint that incorporates both evolving and boundary constraints; in particular, it includes the continuous LPs for Ranking as special cases.

Primal. Suppose m, k and n are positive integers. By default a *vector* is considered as a column vector. Let $P \in \mathbb{R}^{k \times n}$ be a matrix. Let $z : [0,1] \to \mathbb{R}^n$ be a measurable primal function variable such that

(i) Pz is continuous except at a finite number of jump discontinuities in $[0,1]$;
(ii) Pz is non-decreasing in $[0,1]$; and
(iii) Pz is differentiable almost everywhere in $[0,1]$.

Let $B, E, F : [0,1] \to \mathbb{R}^{m \times n}$ be measurable functions. Let $A \in \mathbb{R}^n$, $K \in \mathbb{R}^k$, $D \in \mathbb{R}^{m \times n}$ and $C \in \mathbb{R}^m$ be constants. In the rest of this paper, we use "$\forall t$" to denote "for almost all t", which means for all but a measure zero set. The primal LP is

$$\begin{array}{lll} \mathsf{CP} & \min & p(z) = \int_0^1 A^T z(t)dt \\ & \text{s.t.} & Pz'(t) \geq 0 \qquad \forall t \in [0,1] \qquad (5.1) \\ & & Pz(0) = K \qquad (5.2) \\ \end{array}$$

$$E(t)z(1) + B(t)z(t) + \int_0^1 F(s)z(s)ds + \int_0^t Dz(s)ds \geq C \qquad \forall t \in [0,1] \qquad (5.3)$$

$$z(t) \geq 0 \qquad \forall t \in [0,1].$$

Remark. The continuous LPs that we encounter in this paper do not have a constraint of the form (5.2). However, we include it here to be compatible with the continuous LP framework developed in [3]. We shall describe how the existence of constraint (5.2) will affect the form of the dual.

Dual. Let $\zeta : [0,1] \to \mathbb{R}^k$ and $w : [0,1] \to \mathbb{R}^m$ be measurable dual function variables such that ζ is continuous in $[0,1]$ and differentiable almost everywhere in $[0,1]$. In order to satisfy weak duality with the primal, if there is no constraint (5.2) in CP, then we require constraint (5.4) to appear in the following dual LP (which means the term $K^T\zeta(0)$ in the objective function vanishes).

$$\begin{array}{lll} \mathsf{CD} & \max & d(\zeta, w) = K^T\zeta(0) + \int_0^1 C^T w(t)dt \\ & \text{s.t.} & \zeta(0) = 0 \qquad (5.4) \\ & & P^T\zeta(1) + \int_0^1 E^T(t)w(t)dt \leq 0 \qquad (5.5) \end{array}$$

$$-P^T\zeta'(t) + B^T(t)w(t) + F^T(t)\int_0^1 w(s)ds + \int_t^1 D^T w(s)ds \leq A \qquad \forall t \in [0,1] \qquad (5.6)$$

$$\zeta(t), w(t) \geq 0 \qquad \forall t \in [0,1].$$

For vectors $u = (u_1, \ldots, u_n)^T$ and $v = (v_1, \ldots, v_n)^T$, denote the point-wise product of u and v by $u \circ v := (u_1v_1, \ldots, u_nv_n)^T$. We have the following result for CP and CD.

Lemma 5.1 (Weak Duality and Complementary Slackness) *Suppose z and (ζ, w) are feasible primal and dual solutions, respectively. Then, $d(\zeta, w) \leq p(z)$. Moreover, suppose z and (ζ, w)*

satisfy the following complementary slackness conditions $\forall t \in [0,1]$*:*

$$[Pz'(t)] \circ \zeta(t) = 0 \tag{5.7}$$

$$[E(t)z(1) + B(t)z(t) + \textstyle\int_0^1 F(s)z(s)ds + \int_0^t Dz(s)ds - C] \circ w(t) = 0 \tag{5.8}$$

$$[-P^T\zeta'(t) + B^T(t)w(t) + F^T(t)\textstyle\int_0^1 w(s)ds + \int_t^1 D^T w(s)ds - A] \circ z(t) = 0 \tag{5.9}$$

$$[P^T\zeta(1) + \textstyle\int_0^1 E^T(t)w(t)dt] \circ z(1) = 0 \tag{5.10}$$

and if in addition z *has a discontinuity at* $\mu \in [0,1]$*,*

$$\zeta(\mu) = 0. \tag{5.11}$$

Then, z *and* (ζ, w) *are optimal to* CP *and* CD*, respectively, and achieve the same optimal value; conversely, if* $d(\zeta, w) = p(z)$*, then the complementary slackness conditions hold.*

Proof: To prove $d(\zeta, w) \le p(z)$, by (5.3) we have

$$\begin{aligned}
&d(\zeta, w) = K^T\zeta(0) + \textstyle\int_0^1 C^T w(t)dt \\
&\le K^T\zeta(0) + \textstyle\int_0^1 [E(t)z(1) + B(t)z(t) + \int_0^1 F(s)z(s)ds + \int_0^t Dz(s)ds]^T w(t)dt \\
&= K^T\zeta(0) + \textstyle\int_0^1 [B^T(t)w(t) + F^T(t)\int_0^1 w(s)ds + \int_t^1 D^T w(s)ds]^T z(t)dt + [\int_0^1 E^T(t)w(t)dt]^T z(1),
\end{aligned}$$

where in the last step we change the order of integration by using Tonelli's Theorem on measurable function g: $\int_0^1\int_0^t g(t,s)dsdt = \int_0^1\int_t^1 g(s,t)dsdt$. Using (5.6) we obtain

$$\begin{aligned}
d(\zeta, w) &\le K^T\zeta(0) + \textstyle\int_0^1 [A + P^T\zeta'(t)]^T z(t)dt + [\int_0^1 E^T(t)w(t)dt]^T z(1) \\
&= K^T\zeta(0) + \textstyle\int_0^1 A^T z(t)dt + \int_0^1 (P^T\zeta'(t))^T z(t)dt + [\int_0^1 E^T(t)w(t)dt]^T z(1).
\end{aligned}$$

Recall that ζ is continuous in $[0,1]$, while Pz is continuous except at a finite number of jump discontinuities in $[0,1]$. Let $\mu_1, \mu_2, \ldots, \mu_K$ be the jump discontinuities of Pz. Since Pz is non-decreasing by definition, we have $Pz(\mu_k^-) \le Pz(\mu_k^+)$ for $1 \le k \le K$. Let $\mu_0 := 0$ and $\mu_{K+1} := 1$. Using integration by parts and the Fundamental Theorem of Calculus on the intervals separated by the jump discontinuities, we obtain

$$\begin{aligned}
&\textstyle\int_0^1 (P^T\zeta'(t))^T z(t)dt \\
&= \textstyle\sum_{k=0}^K \int_{\mu_k}^{\mu_{k+1}} d\left[(Pz(t))^T\zeta(t)\right] - \int_0^1 (Pz'(t))^T\zeta(t)dt \\
&= \textstyle\sum_{k=0}^K \left[(Pz(\mu_{k+1}^-))^T\zeta(\mu_{k+1}) - (Pz(\mu_k^+))^T\zeta(\mu_k)\right] - \int_0^1 (Pz'(t))^T\zeta(t)dt \\
&= (Pz(1))^T\zeta(1) + \textstyle\sum_{k=1}^K \left[Pz(\mu_k^-) - Pz(\mu_k^+)\right]\zeta(\mu_k) - (Pz(0))^T\zeta(0) - \int_0^1 (Pz'(t))^T\zeta(t)dt \\
&\le (P^T\zeta(1))^T z(1) - (P^T\zeta(0))^T z(0) - \textstyle\int_0^1 (Pz'(t))^T\zeta(t)dt.
\end{aligned}$$

Substituting $\int_0^1 (P^T\zeta'(t))^T z(t)dt$ with the above expression, we have

$$\begin{aligned}
d(\zeta, w) &\le K^T\zeta(0) + \textstyle\int_0^1 A^T z(t)dt + (P^T\zeta(1))^T z(1) - (P^T\zeta(0))^T z(0) \\
&\qquad - \textstyle\int_0^1 (Pz'(t))^T\zeta(t)dt + [\int_0^1 E^T(t)w(t)dt]^T z(1) \\
&\le [K - Pz(0)]^T\zeta(0) + \textstyle\int_0^1 A^T z(t)dt + [P^T\zeta(1) + \int_0^1 E^T(t)w(t)dt]^T z(1) \\
&= \textstyle\int_0^1 A^T z(t)dt + [P^T\zeta(1) + \int_0^1 E^T(t)w(t)dt]^T z(1) \\
&\le \textstyle\int_0^1 A^T z(t)dt \\
&= p(z),
\end{aligned}$$

where the second inequality follows from (5.1), the first equality from (5.4) (or (5.2) if it exists), and the last inequality from (5.5). In conclusion, we have $d(\zeta, w) \leq p(z)$. Moreover, if z and (ζ, w) satisfy conditions (5.7) – (5.11), then all the inequalities above hold with equality. Hence, $d(\zeta, w) = p(z)$; so z and (ζ, w) are optimal for CP and CD, respectively.

Conversely, if $d(\zeta, w) = p(z)$, then all the inequalities above must hold with equality, which implies that the complementary conditions are all satisfied. ■

5.2 Performance Ratio of Ranking for the Weighted Case

Recall that in Section 3 we obtained $\mathsf{LP}^{\varphi}_{\infty}$ as a continuous counterpart for the discrete LP^{ψ}_{m}. We now derive $\mathsf{LP}^{\varphi}_{\infty}$ and $\mathsf{LD}^{\varphi}_{\infty}$ from CP (without constraint (5.2)) and CD, respectively, where the optimal value of $\mathsf{LP}^{\varphi}_{\infty}$ gives a lower bound for the performance ratio of weighted Ranking (that uses the continuous sample space $\Omega = [0,1]^V$). Let $m = 2$, $n = 1$ and $k = 1$. Set the coefficients as follows:

$$A = 1, P = -1, B(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, D(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, E(t) = \begin{bmatrix} 2\Phi \\ 0 \end{bmatrix}, F(t) = \begin{bmatrix} 5\varphi(t) - t\varphi'(t) \\ 2\varphi(t) - (1-t)\varphi'(t) \end{bmatrix}, C = \begin{bmatrix} 3\Phi \\ \varphi(0) \end{bmatrix}.$$

Let z be the primal function variable, and we recover the primal LP:

$$\begin{array}{lll}
\mathsf{LP}^{\varphi}_{\infty} & \min & \int_0^1 z(t)dt \\
& \text{s.t.} & z'(t) \leq 0 \qquad \forall t \in [0,1] \\
& & 2\Phi \cdot z(1) + \int_0^1 [5\varphi(t) - t\varphi'(t)]\, z(t)dt \geq 3\Phi \qquad (5.12) \\
& & \int_0^1 [2\varphi(t) - (1-t)\varphi'(t)]\, z(t)dt \geq \varphi(0) \qquad (5.13) \\
& & z(t) \geq 0 \qquad \forall t \in [0,1] \\
& & \Phi = \int_0^1 \varphi(t)dt.
\end{array}$$

Let ζ and $w = (w_1, w_2)$ be the dual variables, where ζ is continuous in $[0,1]$ and differentiable almost everywhere in $[0,1]$. Note that w_1 and w_2 appear only in the form $\int_0^1 w_1(t)dt$ and $\int_0^1 w_2(t)dt$. Therefore, we can replace $\int_0^1 w_1(t)dt$ and $\int_0^1 w_2(t)dt$ by real numbers w_1 and w_2, respectively. The dual LP is as follows:

$$\begin{array}{lll}
\mathsf{LD}^{\varphi}_{\infty} & \max & 3\Phi \cdot w_1 + \varphi(0) \cdot w_2 \\
& \text{s.t.} & \zeta(0) = 0 \\
& & -\zeta(1) + 2\Phi \cdot w_1 \leq 0 \qquad (5.14) \\
& & \zeta'(t) + [5\varphi(t) - t\varphi'(t)]\, w_1 + [2\varphi(t) - (1-t)\varphi'(t)]\, w_2 \leq 1 \qquad \forall t \in [0,1] \qquad (5.15) \\
& & \zeta(t), w_1, w_2 \geq 0 \qquad \forall t \in [0,1] \\
& & \Phi = \int_0^1 \varphi(t)dt.
\end{array}$$

We discuss a procedure for constructing a pair of primal feasible solution z and dual feasible solution (ζ, w_1, w_2) that are "nearly" optimal, where z has a jump discontinuity $\mu \in [0,1]$. The complementary slackness condition $\zeta(\mu) = 0$ can only be checked experimentally, as the closed forms for these solutions could not be found. However, since the primal and the dual objective values are close, we can conclude that both are nearly optimal by Lemma 5.1.

Constructing primal feasible solution z. Experiments suggest that the optimal primal has the following form. Let $a > b \geq 0$ be real numbers. Moreover, z has a jump discontinuity μ. Define

$$z(t) = \begin{cases} a, & 0 \leq t \leq \mu \\ b, & \mu < t \leq 1. \end{cases}$$

Then, we have $\int_0^1 z(t)dt = a\mu + b(1-\mu)$. Let $\Phi_\mu = \int_0^\mu \varphi(t)dt$. Observe the following:

$$\begin{aligned} \int_0^1 \varphi(t)z(t)dt &= a\Phi_\mu + b(\Phi - \Phi_\mu) \\ \int_0^1 \varphi'(t)z(t)dt &= a(\varphi(\mu) - 1) + b(\varphi(1) - \varphi(0) - \varphi(\mu) + 1) \\ \int_0^1 t\varphi'(t)z(t)dt &= a(\mu\varphi(\mu) - \Phi_\mu) + b(\varphi(1) - \Phi - \mu\varphi(\mu) + \Phi_\mu). \end{aligned}$$

For z to be optimal in $\mathsf{LP}_\infty^\varphi$, the constraints (5.12) and (5.13) should hold with equality. Hence $\mathsf{LP}_\infty^\varphi$ can be rewritten as

$$\begin{aligned} \min \quad & a\mu + b(1-\mu) \\ \text{s.t.} \quad & (6\Phi_\mu - \mu\varphi(\mu))a + (8\Phi - 6\Phi_\mu + \mu\varphi(\mu) - \varphi(1))b = 3\Phi \\ & (\Phi_\mu + \mu\varphi(\mu) - \varphi(\mu) + 1)a + (\Phi - \Phi_\mu - \mu\varphi(\mu) + \varphi(\mu) + \varphi(0) - 1)b = \varphi(0). \end{aligned}$$

By solving the equality constraints as a linear system with respect to a and b and applying substitution to $a\mu + b(1-\mu)$, we can obtain a minimization problem on a function of μ. Setting $\varphi(t) := 1 - \frac{e^{17t}-1}{e^{17}-1}$ for $t \in [0,1]$, and running experiments on this problem (with precision 1×10^{-6}), we conclude that the function achieves minimum value at $\mu \approx 0.895033$, where $a \approx 0.547528$, $b \approx 0.109144$ and $a\mu + b(1-\mu) \approx 0.501512$. Note that the function z defined by a, b and the jump discontinuity μ is feasible in $\mathsf{LP}_\infty^\varphi$. Hence, z achieves a value of 0.501512 for $\mathsf{LP}_\infty^\varphi$.

Constructing the dual feasible solution (ζ, w_1, w_2). We first observe that for (ζ, w_1, w_2) to be optimal in $\mathsf{LD}_\infty^\varphi$, the constraints (5.14) and (5.15) should hold with equality. Hence, we obtain

$$\zeta(1) = 2\Phi \cdot w_1 \tag{5.16}$$

$$\zeta'(t) = 1 - \big[5\varphi(t) - t\varphi'(t)\big] w_1 - \big[2\varphi(t) - (1-t)\varphi'(t)\big] w_2. \tag{5.17}$$

Taking integral on both sides of equation (5.17) and using $\zeta(0) = 0$, we have

$$\begin{aligned} \zeta(t) &= t - \left[6\int_0^t \varphi(s)ds - t\varphi(t)\right] w_1 - \left[\int_0^t \varphi(s)ds - (1-t)\varphi(t) + \varphi(0)\right] w_2 \\ &= t + (w_1 t + w_2(1-t))\varphi(t) - (6w_1 + w_2)\int_0^t \varphi(s)ds - w_2\varphi(0). \end{aligned} \tag{5.18}$$

Note that given φ, the function ζ is defined by the above equation in terms of t, w_1 and w_2. Setting $t = 1$ we have

$$\zeta(1) = 1 - (6\Phi - \varphi(1))w_1 - (\Phi + \varphi(0))w_2. \tag{5.19}$$

Combining (5.16) and (5.19) we get

$$(8\Phi - \varphi(1))w_1 + (\Phi + \varphi(0))w_2 = 1. \tag{5.20}$$

We observe that for any $w_1, w_2 \geq 0$ satisfying (5.20), the function ζ determined by (5.18) satisfies (5.19), and hence (5.16) and (5.17) also hold.

Therefore, to show that (ζ, w_1, w_2) is feasible in $\mathsf{LD}^{\varphi}_{\infty}$, it remains to check that $\zeta \geq 0$. Recall that $\mathsf{LD}^{\varphi}_{\infty}$ has an objective function in terms of w_1 and w_2, which by (5.20) can be rewritten as

$$3\Phi \cdot w_1 + \varphi(0) \cdot w_2 = \tfrac{\varphi(0)}{\Phi+\varphi(0)} - \left(\tfrac{8\Phi-\varphi(1)}{\Phi+\varphi(0)} - 3\Phi\right) w_1.$$

Set $\varphi(t) := 1 - \frac{e^{17t}-1}{e^{17}-1}$ for $t \in [0,1]$. Then it can be checked that $\frac{8\Phi-\varphi(1)}{\Phi+\varphi(0)} - 3\Phi \geq 0$. Hence the objective function is decreasing with respect to w_1. On the other hand, the function ζ defined by (5.18) may be negative when w_1 is too small. Let w_1 be the minimum value such that (substituting w_2 using (5.20))

$$\zeta(t) = t + \left[w_1 t + \tfrac{1-(8\Phi-\varphi(1))w_1}{\Phi+\varphi(0)}(1-t)\right]\varphi(t) - \left[6w_1 + \tfrac{1-(8\Phi-\varphi(1))w_1}{\Phi+\varphi(0)}\right]\int_0^t \varphi(s)ds - \tfrac{1-(8\Phi-\varphi(1))w_1}{\Phi+\varphi(0)}\varphi(0)$$
$$\geq 0$$

for all $t \in [0,1]$. Running experiments on this minimization problem (with precision 1×10^{-6}) we obtain $w_1 \approx 0.0129253$ and hence $w_2 \approx 0.465017$ by (5.20). Moreover, we check that ζ is non-negative at the local minimum $t_0 \approx 0.895033$. The objective function $3\Phi \cdot w_1 + \varphi(0) \cdot w_2 \approx 0.501512$. Therefore, we achieve a dual feasible solution (ζ, w_1, w_2) with objective value at least 0.501512. Since the primal and the dual solutions have very close values, we conclude that both are nearly optimal.

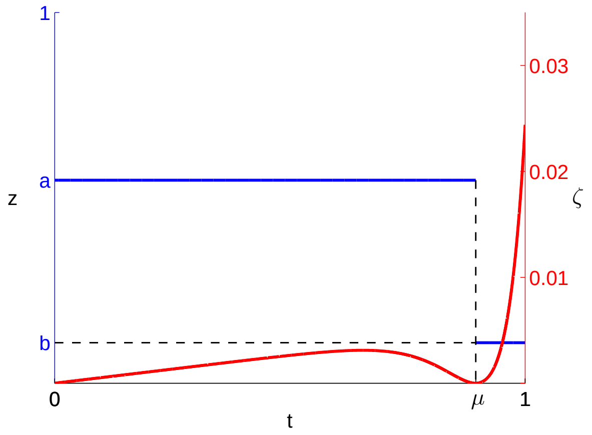

Figure 5.1: Primal z and Dual ζ

Remark on complementary slackness. It can be easily checked that the primal and dual solutions we construct above satisfy all the complementary slackness conditions, where the last one $\zeta(\mu) = 0$ we can verify empirically. From experiments the dual ζ achieves value zero at $t_0 \approx 0.895033$, which is the same (with precision 1×10^{-6}) as the jump discontinuity μ of the primal z (Figure 5.1).

Proof of Theorem 1.2: By using the procedure described above we can construct a dual feasible solution (ζ, w_1, w_2) to $\mathsf{LD}^{\varphi}_{\infty}$ that achieves objective value 0.501512 when $\varphi(t) = 1 - \frac{e^{17t}-1}{e^{17}-1}$ for $t \in [0,1]$. By Lemma 5.1 the performance ratio of weighted Ranking is at least 0.501512. ■

5.3 Performance Ratio of Ranking for the Unweighted Case

We first derive LP^{U}_{∞} and LD^{U}_{∞} from CP (without constraint (5.2)) and CD, respectively, where LP^{U}_{∞} serves as a continuous counterpart for the discrete LP^{U}_{n} for unweighted Ranking.

Let $m = 3$, $n = 2$ and $k = 1$. Set the coefficients as follows:

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, P = \begin{bmatrix} -1 & 0 \end{bmatrix},$$

$$E(t) = \begin{bmatrix} 0 & 0 \\ t & 0 \\ 0 & 0 \end{bmatrix}, B(t) = \begin{bmatrix} (1-t) & 0 \\ t & 0 \\ 0 & 0 \end{bmatrix}, F(t) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 2 & -1 \end{bmatrix}, D = \begin{bmatrix} 2 & 0 \\ -1 & 2 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 \\ 0 \\ 1. \end{bmatrix}$$

Let $z = (\alpha, \beta)$ be the primal variable. Observe the monotone and continuity assumptions apply to only α; moreover, since β appears only within an integral in the following primal LP, it may take unconventional form such as the Dirac delta function.

$$\begin{array}{llrll}
\mathsf{LP}^U_\infty & \min & \int_0^1 \alpha(t)dt & & \\
& \text{s.t.} & \alpha'(t) \leq 0 & \forall t \in [0,1] & (5.21) \\
& & (1-t)\alpha(t) + 2\int_0^t \alpha(s)ds \geq 1 & \forall t \in [0,1] & (5.22) \\
& & t \cdot \alpha(1) + t \cdot \alpha(t) + \int_0^t [-\alpha(s) + 2\beta(s)]ds \geq 0 & \forall t \in [0,1] & (5.23) \\
& & \int_0^1 [2\alpha(s) - \beta(s)]ds \geq 1 & & (5.24) \\
& & \alpha(t), \beta(t) \geq 0 & \forall t \in [0,1]. &
\end{array}$$

Let ζ and $w = (\xi, \eta, \gamma)$ be the dual variables where ζ is continuous in $[0, 1]$ and differentiable almost everywhere in $[0, 1]$. Note that γ appears only in the form $\int_0^1 \gamma(t)dt$. Therefore, we can replace $\int_0^1 \gamma(t)dt$ by a real number γ. The dual LP is as follows:

$$\begin{array}{llrll}
\mathsf{LD}^U_\infty & \max & -\zeta(0) + \int_0^1 \xi(t)dt + \gamma & & \\
& \text{s.t.} & \zeta(0) = 0 & & (5.25) \\
& & -\zeta(1) + \int_0^1 t\eta(t)dt \leq 0 & & (5.26) \\
& & \zeta'(t) + [(1-t)\xi(t) + t\eta(t)] + 2\gamma + \int_t^1 [2\xi(s) - \eta(s)]ds \leq 1 & \forall t \in [0,1] & (5.27) \\
& & -\gamma + 2\int_t^1 \eta(s)ds \leq 0 & \forall t \in [0,1] & (5.28) \\
& & \zeta(t), \xi(t), \eta(t), \gamma \geq 0 & \forall t \in [0,1]. &
\end{array}$$

Lower bounding the ratio of Ranking via LP^U_∞. Before we discuss how to compute the optimal value p_∞ of LP^U_∞, we would like to show that LP^U_∞ indeed gives a lower bound on the performance ratio of unweighted Ranking. Since we already know that the performance ratio on a graph with n nodes is at least the optimal value p_n of LP^U_n, it suffices to show that $p_n \geq p_\infty$ for all n. A simple idea is to consider a feasible solution for LP^U_n, and convert it into the corresponding step function that is feasible in LP^U_∞. However, we find that it is not straightforward to convert the step function into a feasible solution in LP^U_∞, and we instead show a weaker result ($p_\infty \leq p_n + \frac{1}{n}$), which is sufficient for our purpose.

Lemma 5.2 (Relating p_∞ and p_n) *For all $n \in \mathbb{Z}^+$, we have $p_\infty \leq p_n + \frac{1}{n}$.*

Proof: Let (x, y) be an optimal solution to LP^U_n. Our goal is using (x, y) to construct a feasible solution (α, β) to LP^U_∞ that has an objective value at most $p_n + \frac{1}{n}$. Recall that to form a feasible solution to LP_∞^U the function α should be continuous except at a finite number of jump discontinuities in $[0,1]$. Define the step function α with jump discontinuities $\{\frac{i-1}{n} : 2 \le i \le n, i \in \mathbb{Z}\}$ in $[0,1]$ as follows: $\alpha(0) := x_1 + \frac{1}{n}$; and $\alpha(t) := x_i + \frac{1}{n}$ for $t \in \left(\frac{i-1}{n}, \frac{i}{n}\right]$ and $1 \le i \le n$. On the other hand, we note that β appears only within integrals in LP_∞^U; hence, we only need β to be non-negative and integrable. As we shall see later, we would require $\int_0^t \beta(s)ds$ to be sufficiently large even for small $t > 0$. Therefore, it will be convenient for β to take the form of a Dirac delta function δ_u, which can be viewed as a distribution with mass concentrated at a single point $u \in \mathbb{R}$:

$$\delta_u(t) = \begin{cases} +\infty, & t = u \\ 0, & t \neq u \end{cases} \qquad \text{and} \qquad \int_{-\infty}^{+\infty} \delta_u(t)dt = 1.$$

For our purpose, we can assign $\beta(t) := (\frac{1}{n}\sum_{i=1}^n y_i + \frac{1}{n}) \cdot \delta_0(t)$, where all the mass is concentrated at $t = 0$.

It can be checked that the weak duality shown in Lemma 5.1 still holds for this generalized function variable β with other appropriate variables. Then, the objective value of (α, β) is

$$\int_0^1 \alpha(t)dt = \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} \alpha(t)dt = \frac{1}{n}\sum_{i=1}^n \left(x_i + \frac{1}{n}\right) = p_n + \frac{1}{n}.$$

It remains to prove that (α, β) is feasible to LP_∞^U, i.e., it satisfies the constraints (5.21) – (5.24) of LP_∞^U. Clearly $\alpha'(t) \le 0$ for $t \in [0,1] \setminus \{\frac{i}{n} : 0 \le i \le n, i \in \mathbb{Z}\}$. Hence (5.21) is satisfied.

To check (5.22) – (5.24), we fix $t \in (0,1]$, and let i be the integer such that $t \in \left(\frac{i-1}{n}, \frac{i}{n}\right]$. Then $\alpha(t) = x_i + \frac{1}{n}$. Also, observe that $\int_0^t \alpha(s)ds = \frac{1}{n}\sum_{j=1}^{i-1} x_j + \left(t - \frac{i-1}{n}\right) x_i + \frac{t}{n}$ and $\int_0^t \beta(s)ds = \frac{1}{n}\sum_{i=1}^n y_i + \frac{1}{n}$. The feasibility of (α, β) to LP_∞^U is verified via the feasibility of (x, y) to LP_n^U as follows.

- Constraint (5.22). We have

$$\begin{aligned} &(1-t)\alpha(t) + 2\int_0^t \alpha(s)ds = (1-t)\left(x_i + \tfrac{1}{n}\right) + \tfrac{2}{n}\sum_{j=1}^{i-1} x_j + 2\left(t - \tfrac{i-1}{n}\right) x_i + \tfrac{2t}{n} \\ \ge\ &(1-t)x_i + \tfrac{2}{n}\sum_{j=1}^{i-1} x_j + \left(t - \tfrac{i-1}{n}\right) x_i = \left(1 - \tfrac{i-1}{n}\right) x_i + \tfrac{2}{n}\sum_{j=1}^{i-1} x_j \ge 1. \end{aligned}$$

- Constraint (5.23). Note that $x_n \le 1$ (otherwise, this implies that LP_n^U has optimal value at least 1). Then we have

$$\begin{aligned} &t \cdot \alpha(1) + t \cdot \alpha(t) + \int_0^t [-\alpha(s) + 2\beta(s)]ds \\ &= t\left(x_n + \tfrac{1}{n}\right) + t\left(x_i + \tfrac{1}{n}\right) - \tfrac{1}{n}\sum_{j=1}^{i-1} x_j - \left(t - \tfrac{i-1}{n}\right) x_i - \tfrac{t}{n} + \tfrac{2}{n}\sum_{j=1}^n y_j + \tfrac{2}{n} \\ &= t \cdot x_n + \tfrac{t+2}{n} + \tfrac{i}{n} \cdot x_i - \tfrac{1}{n}\sum_{j=1}^{i} x_j + \tfrac{2}{n}\sum_{j=1}^n y_j \\ &\ge \tfrac{i}{n} \cdot x_n + \tfrac{i}{n} \cdot x_i + \tfrac{1}{n}\sum_{j=1}^{i} (-x_j + 2y_j) \ge 0, \end{aligned}$$

  where the first inequality follows from $t \ge \frac{i-1}{n}$ and hence $t \cdot x_n + \frac{t+2}{n} \ge \frac{(i-1)x_n + 1}{n} \ge \frac{i}{n} \cdot x_n$.
- Constraint (5.24). We have

$$\int_0^1 [2\alpha(s) - \beta(s)]ds = \frac{2}{n}\sum_{i=1}^n x_i + \frac{2}{n} - \frac{1}{n}\sum_{i=1}^n y_i - \frac{1}{n} \ge \frac{1}{n}\sum_{i=1}^n (2x_i - y_i) \ge 1.$$

■

Lemma 5.3 (Lower Bounding the Ratio) *The performance ratio of unweighted* Ranking *is at least the optimal value of* LP_∞^U.

Proof: For $n \in \mathbb{Z}^+$, let ρ_n be the worst-case performance ratio of unweighted $\mathsf{Ranking}$ on graphs with n nodes. It suffices to show that $p_\infty \le \rho_n$ for $n \in \mathbb{Z}^+$, where p_∞ is the optimal value of LP^U_∞.

We first claim that given $n, m \in \mathbb{Z}^+$, there exists $N \ge m$ such that $\rho_n \ge \rho_N$. Let k be a positive integer such that $m \le kn =: N$. Suppose $\mathsf{Ranking}$ has performance ratio ρ_n on G_n with n nodes. We make k copies of G_n to form a graph G_N with N nodes. Then, $\mathsf{Ranking}$ also has performance ratio ρ_n on G_N. Since ρ_N is the worst-case performance ratio for graphs with N nodes, it follows that $\rho_n \ge \rho_N$.

Now suppose there exists $n \in \mathbb{Z}^+$ such that $p_\infty > \rho_n$. Then, there exists $\epsilon_n > 0$ such that $p_\infty > \rho_n + \epsilon_n$. Setting $m := \lceil \frac{1}{\epsilon_n} \rceil$, there exists $N \ge \lceil \frac{1}{\epsilon_n} \rceil$ such that $\rho_n \ge \rho_N$. Hence $p_\infty > \rho_N + \epsilon_n \ge p_N + \epsilon_n$. On the other hand, by Lemma 5.2, we have $p_\infty \le p_N + \frac{1}{N} \le p_N + \epsilon_n$, which is a contradiction. ■

Next we discuss a procedure for constructing a pair of primal feasible solution (α, β) and dual feasible solution $(\zeta, \xi, \eta, \gamma)$ that are "nearly" optimal. The complementary slackness conditions are not rigorously proved, since closed forms for some of these solutions could not be found. However, we use these conditions as a guidance to construct a feasible dual solution that is nearly optimal, the objective value of which is a lower bound for the performance ratio of unweighted $\mathsf{Ranking}$.

Constructing a primal feasible solution (α, β). By running experiments on the discrete LP^U_n we have the following observation. The optimal solution (of the discrete LP^U_n) involves two transition points, dividing $[0, 1]$ into three intervals. The constraint corresponding to (5.22) is tight in the first interval, the one corresponding to (5.23) is tight in the second interval. The constraint corresponding to (5.24) is also tight. The variables corresponding to α are always positive, which remain constant in the last interval, while those corresponding to β are positive only in the first interval. This gives us a clue for finding an optimal solution of LP^U_∞ in a similar form as follows.

For (α, β), we consider two transition points $0 \le \lambda < \theta \le 1$. Set $\int_0^1 (2\alpha(s) - \beta(s))ds = 1$ and $\beta(t) = 0$ for $\lambda < t \le 1$. Moreover, we seek for a continuous function α satisfying the following equations:

$$\begin{cases} (1-t)\alpha(t) + 2\int_0^t \alpha(s)ds - 1 = 0, & t \in [0, \lambda] \qquad (5.29) \\ t \cdot \alpha(1) + t \cdot \alpha(t) + \int_0^t [-\alpha(s) + 2\beta(s)]ds = 0, & t \in [\lambda, \theta] \qquad (5.30) \\ \alpha(t) = \alpha(1), & t \in [\theta, 1]. \qquad (5.31) \end{cases}$$

From equation (5.29) we get $\alpha(t) = 1-t$ for $t \in [0, \lambda]$. In particular, $\alpha(\lambda) = 1-\lambda$. In equation (5.30), $\int_0^t \beta(s)ds = \int_0^1 \beta(s)ds$ is a constant, assuming $\beta(t) = 0$ for $\lambda < t \le 1$. Solving this differential equation we get $\alpha(t) = c_0 - \alpha(1)\ln t$ for $t \in [\lambda, \mu]$, where c_0 is some constant. The continuity of α at λ with $\alpha(\lambda) = 1 - \lambda$ implies $1 - \lambda = c_0 - \alpha(1)\ln\lambda$. Then, $c_0 = 1 - \lambda + \alpha(1)\ln\lambda$ and hence $\alpha(t) = 1 - \lambda - \alpha(1)\ln(t/\lambda)$ for $t \in [\lambda, \mu]$. Moreover, $\alpha(\theta) = \alpha(1)$ from (5.31). Then, the continuity of α at θ gives $1 - \lambda - \alpha(1)\ln(\theta/\lambda) = \alpha(1)$, and hence $\alpha(1) = \frac{1-\lambda}{1+\ln(\theta/\lambda)}$. It follows that

$$\alpha(t) = \begin{cases} 1-t, & t \in [0, \lambda] \\ (1-\lambda)\left(1 - \frac{\ln(t/\lambda)}{1+\ln(\theta/\lambda)}\right), & t \in [\lambda, \mu] \\ \frac{1-\lambda}{1+\ln(\theta/\lambda)}, & t \in [\mu, 1]. \end{cases} \qquad (5.32)$$

By definition of α and (5.30) we have $\int_0^1 \beta(s)ds = \int_0^\lambda \beta(s)ds = \frac{\lambda^2}{4} - \frac{(1-\lambda)\lambda}{2(1+\ln(\theta/\lambda))}$. Recall that $\beta(t) = 0$ for $\lambda < t \le 1$. Also note that β is not required to be continuous in $[0, 1]$. Hence, we can simply set $\beta(t) = \frac{\lambda}{4} - \frac{(1-\lambda)}{2(1+\ln(\theta/\lambda))}$ for $0 \le t \le \lambda$. Then, by using $\int_0^1 (2\alpha(s) - \beta(s))ds = 1$, we have

$$\int_0^1 \alpha(s)ds = \frac{\lambda^2}{8} - \frac{(1-\lambda)\lambda}{4(1+\ln(\theta/\lambda))} + \frac{1}{2}. \qquad (5.33)$$

From (5.32) and (5.33), we obtain the following relation between λ and θ:

$$\tfrac{\lambda^2}{2} + (1-\lambda)\theta + \tfrac{1-\lambda}{1+\ln(\theta/\lambda)} \cdot [1-\lambda-\theta\ln(\theta/\lambda)] = \tfrac{\lambda^2}{8} - \tfrac{(1-\lambda)\lambda}{4(1+\ln(\theta/\lambda))} + \tfrac{1}{2}. \tag{5.34}$$

It can be easily checked that the solution (α, β) constructed above is feasible to LP^U_∞. Hence it remains to find optimal λ and θ. This reduces to the problem of minimizing (5.33) subject to (5.34) with $\lambda, \theta \in [0,1]$. Experimental results (with precision 1×10^{-6}) show that the optimal transition points are $\lambda \approx 0.739924$ and $\theta \approx 0.864958$, and the corresponding objective value is 0.526824.

Constructing a dual feasible solution $(\zeta, \xi, \eta, \gamma)$. Using the above form of primal solution and the complementary slackness conditions as a guidance, we now construct $(\zeta, \xi, \eta, \gamma)$ with transition points λ and θ that satisfies the following. The function ζ is positive only in the third interval, ξ positive only in the first interval, η positive only in the second interval, and γ is a positive real number. The constraint (5.27) is always tight in $[0,1]$, while (5.28) is tight in the first interval. With these restrictions we can construct a dual solution in the following manner.

Let $H := \int_0^1 \eta(s)ds$. Setting constraint (5.28) to be equal for $t = 0$, we get $\lambda = 2H$. Setting constraint (5.27) to be equal, we have for $t \in [0,1]$

$$\zeta'(t) + [(1-t)\xi(t) + t\eta(t)] + 2\gamma + \textstyle\int_t^1 [2\xi(s) - \eta(s)]ds = 1. \tag{5.35}$$

Next we derive ξ, η and ζ in their corresponding "non-zero" intervals, respectively.

- For $0 \le t \le \lambda$, we have $\eta(t) = \zeta(t) = 0$. Equation (5.35) becomes $(1-t)\xi(t) + \int_t^\lambda 2\xi(s)ds + 3H = 1$. Solving this equation we get $\xi(t) = \frac{(1-3H)(1-\lambda)^2}{(1-t)^3}$.
- For $\lambda < t \le \theta$, we have $\xi(t) = \zeta(t) = 0$. Equation (5.35) becomes $t\eta(t) - \int_t^\theta \eta(s)ds + 4H = 1$. Solving this equation we get $\eta(t) = \frac{(1-4H)\theta}{t^2}$.
- For $\theta < t \le 1$, we have $\xi(t) = \eta(t) = 0$. Equation (5.35) becomes $\zeta'(t) + 4H = 1$. Solving this equation with (extra) condition $\zeta(\theta) = 0$ we get $\zeta(t) = (1-4H)(t-\theta)$.

From the definition of η we have $H = \int_0^1 \eta(s)ds = \int_\lambda^\theta \frac{(1-4H)\theta}{s^2}ds = (1-4H)(\frac{\theta}{\lambda} - 1)$. Hence $H = \frac{\theta-\lambda}{4\theta-3\lambda}$. Substituting H in the expressions for ζ, ξ, η and γ, we obtain:

$$\zeta(t) = \begin{cases} 0 \\ 0 \\ \frac{\lambda(t-\theta)}{4\theta-3\lambda} \end{cases} \qquad \xi(t) = \begin{cases} \frac{\theta(1-\lambda)^2}{(4\theta-3\lambda)(1-t)^3} \\ 0 \\ 0 \end{cases} \qquad \eta(t) = \begin{cases} 0, & 0 \le t \le \lambda \\ \frac{\lambda\theta}{(4\theta-3\lambda)t^2}, & \lambda < t \le \theta \\ 0, & \theta < t \le 1 \end{cases}$$

$$\gamma = \tfrac{2(\theta-\lambda)}{4\theta-3\lambda}.$$

Moreover, the objective function of LD^U_∞ is

$$-\zeta(0) + \textstyle\int_0^1 \xi(t)dt + \gamma = \frac{\theta(\lambda-\lambda^2/2)+2(\theta-\lambda)}{4\theta-3\lambda}. \tag{5.36}$$

It can be easily checked that the solution $(\zeta, \xi, \eta, \gamma)$ constructed above satisfies (5.25), (5.27) and (5.28). On the other hand, the satisfiability of (5.26) depends on the specific values of λ and θ, which can be expressed as follows

$$-\zeta(1) + \textstyle\int_0^1 t\eta(t)dt = \frac{\lambda}{4\theta-3\lambda} \cdot [\theta(1+\ln(\theta/\lambda)) - 1] \le 0. \tag{5.37}$$

Hence, solving LD^U_∞ reduces to the problem of maximizing (5.36) subject to (5.37) with $\lambda, \theta \in [0,1]$. By running experiments we find that there exist transition points $\lambda \approx 0.739924$ and $\theta \approx 0.864954$ such that (5.37) is satisfied and the corresponding value of (5.36) is 0.526823.

Proof of Theorem 1.3: By using the procedure described above we can construct a feasible solution $(\zeta, \xi, \eta, \gamma)$ to LD_{∞}^{U} with objective value at least 0.526823. By Lemmas 5.1 and 5.3 the performance ratio of unweighted Ranking is at least 0.526823. ■

References

[1] Gagan Aggarwal, Gagan Goel, Chinmay Karande, and Aranyak Mehta. Online vertex-weighted bipartite matching and single-bid budgeted allocations. In *SODA*, pages 1253–1264, 2011.

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

[3] T.-H. Hubert Chan, Fei Chen, Xiaowei Wu, and Zhichao Zhao. Ranking on arbitrary graphs: Rematch via continuous lp with monotone and boundary condition constraints. In *SODA*, pages 1112–1122, 2014.

[4] Nikhil R. Devanur, Kamal Jain, and Robert D. Kleinberg. Randomized primal-dual analysis of ranking for online bipartite matching. In *SODA*, pages 101–107, 2013.

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

[6] Gagan Goel and Aranyak Mehta. Online budgeted matching in random input models with applications to adwords. In *SODA*, pages 982–991, 2008.

[7] Gagan Goel and Pushkar Tripathi. Matching with our eyes closed. In *FOCS*, pages 718–727, 2012.

[8] Gagan Goel and Pushkar Tripathi. Matching with our eyes closed. *CoRR*, abs/1306.2988, 2013.

[9] Chinmay Karande, Aranyak Mehta, and Pushkar Tripathi. Online bipartite matching with unknown distributions. In *STOC*, pages 587–596, 2011.

[10] Richard M. Karp, Umesh V. Vazirani, and Vijay V. Vazirani. An optimal algorithm for on-line bipartite matching. In *STOC*, pages 352–358, 1990.

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

[12] Mohammad Mahdian and Qiqi Yan. Online bipartite matching with random arrivals: an approach based on strongly factor-revealing LPs. In *STOC*, pages 597–606, 2011.

[13] Silvio Micali and Vijay V. Vazirani. An $O(\sqrt{V}E)$ algorithm for finding maximum matching in general graphs. In *FOCS'80*, pages 17–27. IEEE Computer Society, 1980.

[14] Matthias Poloczek and Mario Szegedy. Randomized greedy algorithms for the maximum matching problem with new analysis. In *FOCS*, pages 708–717, 2012.

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

[16] 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.