Online Matching with Stochastic Rewards: Optimal Competitive Ratio via Path Based Formulation

Vineet Goyal
vgoyal@ieor.columbia.edu
Rajan Udwani
rhu2104@columbia.edu

Abstract

The problem of online matching with stochastic rewards is a generalization of the online bipartite matching problem where each edge has a probability of success. When a match is made it succeeds with the probability of the corresponding edge. Introducing this model, Mehta and Panigrahi (FOCS 2012) focused on the special case of identical and vanishingly small edge probabilities and gave an online algorithm which is 0.567 competitive against a deterministic offline LP. For the case of vanishingly small but heterogeneous probabilities Mehta et al. (SODA 2015), gave a 0.534 competitive algorithm against the same LP benchmark.

We study a generalization of the problem to vertex-weighted graphs. To the best of our knowledge, no results beating $1/2$ were previously known for this setting, even for identical probabilities. We improve on this in two ways. First, we show that a natural generalization of the Perturbed-Greedy algorithm achieves the best possible competitive ratio of $(1 - 1/e)$, when probabilities decompose as a product of two factors, one corresponding to each vertex of the edge. Second, we give a deterministic 0.596 competitive algorithm for the case of fully heterogeneous but vanishingly small edge probabilities. A key contribution of our approach is the use of novel path-based formulations. These allow us to compare against the natural benchmark of clairvoyant (offline) algorithms that know the sequence of arrivals and the edge probabilities in advance, but not the outcomes of potential matches. The idea of a path-based program to compare against clairvoyant algorithms may be of independent interest in other online settings.

1 Introduction

Online bipartite matching problems and its variants and generalizations have been intensely studied in the last two decades [KVV90; MSVV05; DJK13; MP12; MWZ15; AGKM11; GM08; M13; DH09; CHN14; MGST12; KMT11; FMMM09; BM08; BJN07]. Internet advertising is a major application domain for these problems, along with Crowdsourcing [HV12; KOS14], resource allocation [CF09; MGL13; GNR14] and more recently personalized recommendations [GNR14]. Despite a rich body of work on these problems, there are several basic questions that remain open. In this work, we study one such problem.

In the classical online bipartite matching problem introduced by Karp et al. [KVV90], we have a graph $G(I, T, E)$ where the vertices in $I$, which might correspond to resources in allocation problems, advertisers in Internet advertising, and tasks in Crowdsourcing, are known in advance and and vertices in $T$, also referred to as arrivals (corresponding to customers, ad slots, workers), are sequentially revealed one at a time. When a vertex $t \in T$ arrives, the set of edges $(i, t) \in E$ is revealed. After each arrival the (online) algorithm must make an irrevocable decision to offer at most one (available/unmatched) vertex $i$ with an edge to $t$, with the goal of maximizing the total number of matches. The performance of the algorithm is compared against an optimal offline algorithm which knows all edges in advance. In particular, let $T$ denote a sequence of arrivals, $OPT(T)$ the optimal achievable by an offline algorithm and $A(T)$
the expected value achieved by a (possibly randomized) online algorithm \( A \). The goal is to design an algorithm that maximizes the competitive ratio
\[
\min_T \frac{A(T)}{OPT(T)}.
\]
Karp et al. [KVV90] proposed and analyzed the Ranking algorithm, which attains the best possible competitive ratio of \((1 - 1/e)\) for online bipartite matching (see [BM08] GM08 DJK13 for a corrected and simplified analysis). This was generalized to vertex-weighted matchings in Aggarwal et al. [AGKM11]. While this gives a more or less complete picture for the classical setting we described above, in all the applications we mentioned there is in fact a non-zero probability that any given match might fail. For instance, in Internet advertising users might not click on the ad and in the popular pay-per-click model an advertiser pays only if the user actually clicks i.e. the match succeeds. Similarly in personalized recommendations, for generating revenue the customer actually has to buy an offered product. Motivated by these observations, Mehta and Panigrahi [MP12] introduced a generalization of the problem where the objective is to maximize the number of “successful” matches. They associate a probability with every edge, \( p_{it} \) for edge \((i, t)\), which is revealed online along with the edge. When an algorithm makes a match, the arrival accepts (successful match or reward) with probability given by the corresponding edge and the success of a given match is independent of past ones. The outcome of the stochastic reward is revealed only after the match is made and the objective is to maximize the expected number of successful matches. A natural generalization of this problem allows each resource (vertex in \( I \)) to have a possibly different reward \( r_i \) (vertex weighted rewards). This is the setting we study.

The introduction of stochastic rewards to the setting of online matching raises an interesting question regarding the nature of the offline algorithm one should compare against. In addition to knowing all edges and their probabilities in advance, should the offline algorithm know the outcomes of the random rewards too? Unsurprisingly, such a benchmark turns out to be too strong for any meaningful bound. A perhaps natural alternative is to compare against adaptive offline algorithms that know the arrivals, edges and edge probabilities in advance but must adapt to the outcome of reward from matches in real time. Such clairvoyant algorithms are an oft-used benchmark [GGI10, GNR14, MSL18, CF09]. While it can be difficult to understand the performance of such algorithms, letting \( OPT(C) \) denote the expected reward of the best clairvoyant algorithm, observe that an upper bound on \( OPT(C) \) is easily given by the optimum of the following LP (see [MSL18, GNR14] for a formal proof).

\[
OPT(LP) = \max \sum_{(i, t) \in E} p_{it} r_i x_{it}
\]

\[
s.t. \quad \sum_{t \in T} p_{it} x_{it} \leq 1 \quad \forall i \in I
\]

\[
\sum_{i \in I} x_{it} \leq 1 \quad \forall t \in T
\]

\[
0 \leq x_{it} \leq 1 \quad \forall (i, t) \in E
\]

As a result, even when evaluating competitive ratios against clairvoyant algorithms one often chooses to use the upper bound \( OPT(LP) \). To the best of our knowledge there are no better upper bounds known on \( OPT(C) \) in related literature. Note that in absence of the stochastic component (all probabilities one), the two benchmarks converge and we have \( OPT(C) = OPT(LP) \).

1Note that we compare against a non-adaptive adversary and thus take the minimum over all fixed arrival sequences. This is a standard assumption in the literature on online matching.

2This algorithm orders all offline vertices in a random order and matches every online vertex to the highest ranked neighbour available.

3If rewards are allowed to depend on both the arrival and resource the problem becomes much harder. It was shown in [AGKM11] that in this case no algorithm with competitive ratio better than \( O(1/n) \) is possible.
The benchmark given by this LP is at least as strong as the clairvoyant benchmark and sufficiently weaker than offline algorithms that know the outcomes of the rewards in advance. In fact, it is possible to get algorithms that are \(1/2\) competitive against such an LP in even more general settings using a deterministic greedy or even non-adaptive policies (see [GNR14, MWZ15]). However, beating \(1/2\) has proven to be a major challenge. Progress has been made in special cases with remarkable new insights. In particular, Mehta and Panigrahi [MP12] found that for identical rewards and identical probabilities i.e., \(r_i = 1, \forall i \in I\) and \(p_{it} = p \forall (i, t) \in E\): Assigning arrivals to available neighbors with the least number of failed matches in the past (called StochasticBalance) is 0.567 competitive as \(p \to 0\). Further, they showed that no algorithm has a competitive ratio better than 0.621 < \((1 - 1/e)\) against the LP. For the same special case, they also showed that Ranking is 0.534 approximate for vanishing \(p\). When the probabilities are heterogeneous but vanishingly small, Mehta et al. [MWZ15] gave a 0.534 competitive algorithm (called SemiAdaptive). No results beating \(1/2\) are known to us for other cases, or when the rewards \(r_i\) are not identical, even for identical probabilities. We show positive results to this end by giving a new, tighter bound on clairvoyant algorithms via path based formulations. Our formulations, and the associated weak duality, allows us to leverage and build on insights from primal-dual analysis for the classical online matching setting [DJK13, BJN07], leading also to a simpler and more unified analysis.

Roadmap: In the next section, we summarize our results and techniques used followed by further discussion of related work. In Section 1.3, we discuss an obstacle to using the randomized primal-dual framework of Devanur et al. [DJK13], for the stochastic rewards problem. This motivates the consideration of our path based analysis in Section 2, where we formally state and prove our main results. Finally, in Section 3 we conclude with a review of some relevant open problems.

1.1 Our Results

Consider the following natural generalization of the Perturbed-Greedy algorithm.

**Algorithm 1: Generalized Perturbed-Greedy**

\[
S = I, \ g(t) = e^{t-1};
\]

For every \(i \in I\) generate i.i.d. r.v. \(y_i \in U[0,1]\);

for every new arrival \(t\) do

\[
i^* = \arg \max_{i \in (i, t) \in E, i \in S} p_{it} r_i (1 - g(y_i));
\]

offer \(i^*\) to \(t\);

If \(t\) accepts \(i^*\) update \(S = S \setminus \{i^*\}\);

end

Remarks: When \(p_{it} = 1 \forall (i, t) \in E\) this is exactly the Perturbed-Greedy algorithm of [AGKM11], shown to be \((1 - 1/e)\) competitive for the deterministic case. Additionally, the algorithm is adaptive as it takes into account the outcome (success/failure) of every match (refer [MP12, MWZ15] for detailed discussions on adaptivity/non-adaptivity).

**Theorem 1.** For decomposable probabilities \(p_{it} = p_i p_t \forall (i, t) \in E\), Algorithm 1 achieves the best possible competitive ratio of \((1 - 1/e)\) w.r.t. clairvoyant algorithms.

Note that the case of decomposable probabilities (with otherwise arbitrary magnitudes) includes the case of identical probabilities \(p_{it} = p \forall (i, t) \in E\), for which [MP12] showed that there is no 0.621 < \((1 - 1/e)\) competitive algorithm when comparing with \(OPT(LP)\). In contrast, Theorem 1 shows that one can achieve a guarantee of \((1 - 1/e)\) when comparing against \(OPT(C)\).

To show this result we need to address the hurdle with existing techniques (outlined in Section 1.3) and get a tighter bound on \(OPT(C)\). We do so by introducing a new path based program that has a constraint for every possible sample path instance of the stochastic reward
The choice of arbitrary probabilities $p_{it}$, the analysis primarily breaks down due to the lack of a unique preference order across arrivals. This difficulty is not entirely unexpected. Indeed, based on the insights in [MP12], the case of heterogeneous probabilities bears resemblance to the (hard) open problem of AdWords without the small bid assumption [MSVV05, M+13]. Previously, [MP12] and [MWZ15] focused on the special, yet practically important, case of vanishingly small probabilities. The relevance of this case stems from the observation that in applications such as online-advertisement, the probabilities or click-through-rates are often very small [MSVV05, M+13]. As observed by [MP12], this assumption is closely tied to the common assumption of small bids in the AdWords setting. Nonetheless, while there is a simple and elegant primal dual analysis for the (small bids) AdWords problem (due to [DJK13, BJN07]), the case of stochastic rewards (with small probabilities) seems to be quite different and not as amenable to the primal-dual framework. In fact, the analysis of algorithms proposed for this case in [MWZ15] is significantly more complicated. While Theorem 1 bridges this gap for the case of decomposable probabilities, the case of fully heterogeneous probabilities calls for a different algorithm and approach. Expanding on insights from [MP12, MWZ15], we propose the following deterministic algorithm.

ALGORITHM 2: Generalized Fully-Adaptive

Input: Scaling function $g(\cdot) : [0, \infty] \to [0, 1]$, offline vertex set $I$; $S = I$;

For every $i \in I$, $l_i = 0$;

for every new arrival $t$ do

\[ i^* = \arg \max_{i \in (I \setminus S) \times E} p_i r_i g(l_i); \]

offer $i^*$ to $t$;

If $t$ accepts $i^*$ update $S = S \setminus \{i^*\}$;

Else $l_{it} = l_{it} + p_{it}$;

end

The choice of scaling function $g(\cdot)$ naturally influences the competitive ratio. We consider several choices. In particular, we give a choice that leads to a 0.596 competitive algorithm. Note that when $g(t) = e^{-t}$, we recover the Fully-Adaptive algorithm proposed, but not analyzed, in [MP12, MWZ15]. Let $p_{max} = \max_{(i, t) \in E} p_{it}$, then we will show the following.

---

4The Adwords problem is a generalization of online bipartite matching where vertices $i \in I$ have budgets $B_i$, and edges $(i, t)$ have bids $b_{it}$. An arriving vertex can be matched to any neighbour that has not yet spent all its budget. If $i$ and $t$ are matched, $b_{it}$ units of $i$’s budget is consumed. If after some match $i$ has depleted its entire budget, it becomes unavailable. The goal is to maximize the total budget used. The small bid assumption enforces that bids $b_{it} = o(B_i)$. 

---
Competitive Ratio (Against Clairvoyant)

|                     | Identical rewards | Vertex-weighted rewards |
|---------------------|-------------------|-------------------------|
| \( p_{it} = p = o(1) \) | 0.567 → (1−1/e)   | 0.5 → (1−1/e)           |
| \( p_{it} = p_{it} \) | 0.500 → (1−1/e)   | (1−1/e) \footnote{MSVV05} |
| \( p_{it} = o(1) \)  | 0.534 → 0.596     | 0.5 → 0.596             |
| General \( p_{it} \) | 1/2 \footnote{MWZ15} | 1/2 \footnote{MSVV05}   |

Table 1: Summary of best known lower bounds for all relevant settings. \( c_i \) represents the number of copies of each vertex \( i \in I \) available, with \( c_i = 1 \) being the most general case (see footnote 5). Entries in bold represent our contributions.

**Theorem 2.** For \( g(t) = e^t (e^t + 1)^{-1} \), and \( p_{\text{max}} = o(1) \), the competitive ratio of Algorithm 2 is at least \( g(0)(>0.596) \).

We also analyze the inverse scaling function \( g(t) = \frac{1}{\beta^2 t + 1} \), which we show is at least 0.588 competitive. For the function \( g(t) = e^{-\beta t} \), a natural generalization of the function \( e^{-t} \) proposed in [MWZ15], we establish 0.581 competitiveness.

Setting \( g(t) = 1 - e^{-\beta t} \), we recover the algorithm of Mehta et al. [MSVV05], for the AdWords problem with budgets \( r_i \) and bids \( p_{it} \) (\( l_i \) translates to the fraction of the budget of advertiser \( i \) consumed so far). This underpins the connection between our problem and the AdWords problem, and puts our results in a clear correspondence with equivalent results for special cases of AdWords.

1. The case of decomposable probabilities resembles the classical online vertex weighted bipartite matching problem (where Perturbed-Greedy is known to give the best guarantee).
2. The small probabilities case resembles the AdWords setting with small bids (where the algorithm of [MSVV05] gives the best guarantee).

In case of vanishing probabilities however, a very important distinction from AdWords with small bids is that we do not know the budgets in advance. This results in a very different analysis where we do not employ the standard primal-dual framework. In particular, the standard dual fitting seeks to satisfy a dual constraint for every edge, and it is not clear if this can be accomplished with unknown (random) budgets. We instead show that it is sufficient to satisfy a single constraint for every offline vertex. This is an important change, in that we are able to focus on a global constraint that shows that our online algorithm performs competitively in terms of the profit earned from every offline resource, but over the course of all arrivals as opposed to every arrival.

Finally, in Section 2.3 we show that when we have a large number of copies \( c_i \) for every vertex \( i \in I \) (corresponds to budget in AdWords, and initial inventory in the resource allocation setting), our path based program from the decomposable case is equivalent to the classical LP. This is to be expected since the case of stochastic rewards with large inventory admits a \((1−1/e)\) competitive algorithm against the standard LP via a simple extension of the main result in [MSVV05] (Section 6, extension number 5 in [MSVV05], also in [GNR14]).

**1.2 Related Work**

Karp et al. [KVV90] introduced the online bipartite matching model and proposed the optimal \((1−1/e)\) competitive \textbf{Ranking} algorithm. Birnbaum and Mathieu [BM08], Goel and Mehta [GM08] considerably simplified the original analysis. Subsequently, Aggarwal et al. [AGKMT11], proposed the \textbf{Perturbed-Greedy} algorithm and showed that it was \((1−1/e)\) competitive more
generally for the case of online vertex-weighted bipartite matching. Devanur et al. [DJK13] gave an elegant and intuitive randomized primal-dual interpretation of these results. Their framework applies to and simplifies several other settings, such as the related AdWords with small bids problem of Mehta et al. [MSVV05] (and the earlier special case by Kalyanasundaram and Pruhs [KP00]). There have also been a series of results on stochastic arrival models, where there is distributional information in the arrivals that can be exploited for better results [BSSX16, DH09, MGS12, KMT11, FMMD09], as well as simultaneous guarantees in adversarial and stochastic arrival settings [MGZ12]. For a detailed survey we refer the reader to the monograph by Mehta [M+13].

Apart from [MP12, MWZ15] (discussed in Sections 1 and 1.1), other settings closest to ours have been considered in Golrezaei et al. [GNR14] and Ma and Simchi-Levi [MSL18]. Golrezaei et al. [GNR14] consider a broad generalization of our setting where one offers an assortment of products to each arrival, and the arrival then chooses based on a choice model that is revealed online at the time of arrival. With the objective of maximizing total expected reward, they show that when the number of copies (inventory) of each resource approaches infinity, an inventory balancing algorithm is asymptotically \((1 - 1/e)\) competitive. They seek to compare against clairvoyant algorithms however, for the purpose of analysis they use the upper bound \(OPT(LP)\) for a suitably generalized LP. Nonetheless, they achieve the best possible competitive ratio asymptotically. However, for the case of unit inventory,\(^5\) their guarantee converges to 0.5. Recall that beating 0.5 is an open problem even in our setting, which is a special case of theirs. Note that their asymptotic result also implies a deterministic \((1 - 1/e)\) competitive algorithm for our setting when we have infinitely many copies of each vertex \(i \in I\). As previously mentioned, this also follows directly from earlier work by Mehta et al. [MSVV05]. More recently, Ma and Simchi-Levi [MSL18] also studied a generalization of our setting in the resource allocation framework, where each resource \(i \in I\) can be sold at multiple rewards rates with possibly different probabilities of successful reward for each rate. Similar to Golrezaei et al. [GNR14], they focus on the asymptotic regime where the inventory of each resource approaches infinity and give the optimal competitive ratio against clairvoyant algorithms, using \(OPT(LP)\) as an upper bound on \(OPT(C)\) for analysis. Concurrent to our work, Brubach et al. [BGS19] look at more general version where a finite number of rematches are allowed in case of unsuccessful matches. They observe that in this case the problem for a single arrival is optimally solved by a DP. Using this insight show that the DP based greedy algorithm is 0.5 competitive. Further, they highlight several limitations of the LP benchmark and advocate the use of clairvoyant algorithms as the more natural benchmark for future work.

1.3 Preliminaries

Devanur et al. [DJK13] introduced a unifying primal-dual framework for understanding and analyzing algorithms for online matching and related problems. Given the LP for the stochastic rewards setting, a natural approach would be to explore a similar primal-dual algorithm and analysis. At the outset it might even seem that Algorithm 1 offers a natural extension, so the analysis might generalize directly. Yet there are fundamental obstacles on this path and previous work explores novel approaches instead [MP12, MWZ15]. Let us understand one such

\(^5\) Guarantees for the case of unit inventory leads to stronger results that generalize to the case of arbitrary inventories. Consider a unit inventory setting where in place of each \(i \in I\) in the original setting, we have \(c_i\) resources each with inventory of 1 and arrivals that have edges to all \(c_i\) vertices for every edge \((i, t)\) in the original instance. Now the offline/clairvoyant algorithm knows all arrivals in advance and thus knows these copies represent the same resource. Therefore, \(OPT(C)\) remains unchanged and an algorithm for the unit inventory case can be used for arbitrary inventory levels without loss in guarantee.
hurdle from the context of the framework in [DJK13]. We start with the dual of the LP,

$$\begin{align*}
\min & \quad \sum_t \lambda_t + \sum_i \theta_i \\
\text{s.t.} & \quad \lambda_t + p_d \theta_i \geq p_d r_i \quad \forall (i, t) \in E \\
& \quad \lambda_t, \theta_i \geq 0 \quad \forall t \in T, i \in I
\end{align*}$$

In line with [DJK13], when the algorithm offers vertex $i$ to arrival $t$: if the match succeeds let us set dual variables to $\lambda_t = r_i(1 - g(y_i))$ and $\theta_i = r_i g(y_i)$ and we let the variables be zero otherwise. The sum $\sum_t \lambda_t + \sum_i \theta_i$, clearly captures the reward of the algorithm. Further, if the setting also ensured that the dual constraints were satisfied to within a constant factor, we would have a corresponding competitive ratio guarantee for Algorithm 1 via a standard dual fitting argument. While this assignment of dual values will not lead to dual feasibility for all values $\{y_i\}$, for the case of unit probabilities [DJK13], observed that it suffices to satisfy dual constraints (within a constant factor) in expectation over the values $\{y_i\}$ i.e.,

$$E_{y_i}[\lambda_t + \theta_i] \geq (1 - 1/e)r_i.$$ 

As a direct generalization, taking expectations over the success of the match $(i, t)$, in the stochastic case it suffices to show that for the above setting of dual values,

$$E_{(i, t)}[E_{y_i}[\lambda_t + \theta_i]] \geq (1 - 1/e)p_d r_i.$$ 

However, in case of stochastic rewards this need not hold to within any constant factor. For instance, consider the dual constraint corresponding to match $(i, t)$ in our example,

$$\lambda_t + p_d \theta_i = r_i(1 - g(y_i) + p_d g(y_i)) \mathbb{1}(i, t).$$

Where $\mathbb{1}(i, t)$ is an indicator variable that is 1 if edge $(i, t)$ succeeds. Taking expectation over the outcome of the match $(i, t)$ with $y_i$ fixed we get,

$$E_{(i, t)}[\lambda_t + \theta_i] = p_d r_i (1 - g(y_i)) + p_d^2 r_i g(y_i).$$

For small $p_d$, this approaches $p_d r_i (1 - g(y_i))$. So if the overall instance is such that $i$ is matched to $t$ only for $y_i \sim 1$, and $t$ is otherwise unmatched we have, $E_{y_i}[E_{(i, t)}[\lambda_t + \theta_i]] \to 0$. Part of the problem seems to be that the formulation only insists that each resource $i$ is used at most once in expectation, giving rise to a $p_d^2$ term as we saw above. Alternatively, similar to [BJN07, GNR14] we could define the duals to guarantee feasibility. For instance, we can let $\lambda_t$ be the same as above but set $\theta_i = r_i g(y_i)$ (regardless of success/failure of the match). Now, $\lambda_t + p_d \theta_i = p_d r_i$. However, the sum $\lambda_t + \theta_i \geq r_i g(y_i)$ can be much larger than the expected reward $p_d r_i$, of the algorithm for match $(i, t)$. The latter can even approach 0 for $p_d \to 0$. It is not obvious (to us) if one can overcome these hurdles while still considering an expectation based LP. Our path dependent formulation circumvents this problem by imposing constraints on every sample path as opposed to in expectation over all paths.

### 1.4 Notation

We now review and introduce new notation before proceeding with a formal presentation of the results in the next section. First, recall the problem definition. We have a bipartite graph $G$ with a set of $n$ offline vertices $i \in I$ and an arbitrary number of vertices $t \in T$ arriving online. We use the index of arrivals $t$ to also denote their order in time. So assume vertex $t \in T$ arrives at time $t$. Now, all edges $(i, t)$ incident on vertex $t$ are revealed when $t$ arrives, along with a corresponding probability of success $p_d$. On each new arrival, we must make an irrevocable decision to match the arrival to any one of the available offline neighbours. Once a match,
say \((i, t)\), is made it succeeds with probability \(p_{it}\), making \(i\) unavailable to future arrivals and leading to a reward \(r_i\). The objective is to maximize the expected reward summed over all arrivals. Recall that \(OPT(C)\) denotes the expected reward of the best clairvoyant algorithm.

For edge \((i, t)\), let \(1(i, t)\) be an indicator random variable that takes value one w.p. \(p_{it}\). Let \(\omega\) denote a sample path given by an instance of stochastic rewards over all arrivals, and let \(\omega^t\) represent the partial sample path described by \(\omega\) up to and including arrival \(t\). In other words, \(\omega^t\) determines values of random variables \(1(i, t')\), for all edges \((i, t')\) incident on arrivals \(t' \leq t\).

Also, we define \(\omega_t\) to be consistent with \(\omega_{t-1}\) for every \(t \in T\), meaning both represent the same path up to arrival \(t\). Let \(\omega\) and \(\omega_{t-1}\) represent the corresponding random variables. Let \(1^\omega(i, t)\) denote the value of random variable \(1(i, t)\) on sample path \(\omega\). Finally, let \(\Omega\) denote the universe of all sample paths (full and partial).

### 2 Main Results

#### 2.1 Decomposable Probabilities \((p_{it} = p_i \cdot p_t)\)

In this section, we analyze Algorithm 1, which is equivalent to the vertex weighted version of the Perturbed-Greedy algorithm with rewards/revenues \(r_i\) replaced by \(p_{it}r_i\) for arrival \(t\), and show a competitive ratio of \((1 - 1/e)\) in expectation against a non-adaptive adversary.

As discussed in Section 1.3, considering an upper bound on the value \(OPT(C)\) of the clairvoyant via an LP that only imposes constraints in expectation raises several issues. Let \(x_{it}\) denote the decision variable corresponding to whether clairvoyant matches \(i\) to \(t\). This is a random variable that depends on the sample path up to time \(t\). So consider sample paths \(\omega\), and for every edge \((i, t)\), let \(x_{it}^\omega \in [0, 1]\) represent the decision of clairvoyant on whether \(t\) is matched to \(i\) on this sample path. Clearly, the following must be satisfied on every path \(\omega\),

\[
\sum_{t' \in E(i,t)} x_{it}^\omega \leq 1 \quad \forall i \in I
\]

\[
\sum_{t' \in E(i,t)} x_{it}^\omega \leq 1 \quad \forall t \in T
\]

Constraints [1] capture the fact that any resource \(i\) is used at most once on every sample path. This is in contrast to the LP earlier, where this condition was imposed only in expectation over all sample paths. Similarly, constraints [2] capture that \(t\) is matched to at most one vertex on every sample path. Recall that we assume clairvoyant knows all edges and edge probabilities in advance but not actual outcomes of future matches. In fact, when deciding the match for arrival \(t\) we can let clairvoyant have access to values \(1(i, t')\), for all edges \((i, t')\) incident on arrivals \(t' < t\). Knowing if other matches in the past would have been (un)successful does not give any additional useful information to the clairvoyant, due to the irrevocability of decisions and independence of rewards across arrivals. However, \(x_{it}\) must be independent of the edge random variables \(1(j, \tau)\) for all edges \((j, \tau)\) revealed in the future i.e., \(\tau > t\). In other words, we must have,

\[
x_{it}^\omega = x_{it}^{\omega_0} \quad \forall \omega, \omega_0 \in \Omega \text{ such that } \omega^{t-1} = \omega_0^{t-1}.
\]

Using [1], [2] and [3], we now formulate a linear program with these constraints and the
objective of maximizing the total expected reward.

$$\text{PBP} : \max E_{\omega \sim \omega} \left[ \sum_{(i,t) \in E} r_i x^\omega_{it} 1^\omega(i,t) \right]$$

s.t. \( \sum_{t \mid (i,t) \in E} x^\omega_{it} 1^\omega(i,t) \leq 1 \quad \forall i \in I, \omega \in \Omega \)

\( \sum_{i \mid (i,t) \in E} x^\omega_{it} \leq 1 \quad \forall t \in T, \omega \in \Omega \)

\( x^\omega_{it} = x^\omega_{it} \quad \forall (i,t) \in E, \{ \omega, \omega_0 \in \Omega \mid \omega^{t-1} = \omega_0^{t-1} \} \)

\( 0 \leq x^\omega_{it} \leq 1 \quad \forall (i,t) \in E, \omega \in \Omega. \)

Since a clairvoyant algorithm must satisfy all constraints in the program, the values \( x^\omega_{it} \) generated by executing any clairvoyant algorithm over sample paths \( \omega \), yield a feasible solution for PBP. Let us simplify the objective of PBP, refer to the optimal objective value of PBP as \( OPT(\text{PBP}) \) and consider the following,

\[
OPT(\text{PBP}) = E_{\omega} \left[ \sum_{(i,t)} r_i x^\omega_{it} 1^\omega(i,t) \right] = \sum_{(i,t)} E_{\omega} \left[ E_{\omega} \left[ r_i x^\omega_{it} 1^\omega(i,t) \mid \omega^{t-1} = \omega^{t-1} \right] \right] = \sum_{(i,t)} E_{\omega} \left[ r_i x^\omega_{it} E_{\omega} \left[ 1^\omega(i,t) \mid \omega^{t-1} = \omega^{t-1} \right] \right] = \sum_{(i,t)} E_{\omega} \left[ \sum_{(i,t)} p_{it} r_i E_{\omega} [x^\omega_{it}] \right].
\]

Where the first equality follows form the tower property of expectation, the second from condition \( (3) \), the third equality follows from the independence of each reward \( 1^\omega(i,t) \) from past rewards, and the final equality also from condition \( (3) \).

Remarks: Note that, given a random seed \( s \) for a randomized clairvoyant algorithm, variables \( x^\omega_{it, s} \) corresponding to the output of the clairvoyant on sample path \( \omega \), are binary. However, \( x^\omega_{it} = E_{\omega} [x^\omega_{it, s}] \) can be fractional. Further note, in the deterministic case where all edge probabilities are one, PBP is equivalent to the classical LP.

In light of the primal-dual framework in \([DJK13]\), the natural next step would be try to construct a dual fitting argument. However, the dual of PBP has a constraint for every sample path and there may not exist any setting of dual variables that results in an appropriate fitting. The lemma that follows establishes a weak duality result instead, which lets us upper bound \( OPT(\text{PBP}) \) using suitable dual fitting.

**Lemma 1.** Consider non-negative variables \( \lambda^\omega_i, \theta^\omega_i \) satisfying,

\[
E_{\omega} [\lambda^\omega_i + \theta^\omega_i 1^\omega(i,t) \mid \omega^{t-1} = \omega^{t-1}] \geq p_{it} r_i,
\]

for every edge \((i,t)\). Then for every feasible solution \( \{x^\omega_{it}\} \) for PBP,

\[
\sum_{(i,t)} p_{it} r_i E_{\omega} [x^\omega_{it}] \leq E_{\omega} \left[ \sum_{i} \lambda^\omega_i + \sum_{i} \theta^\omega_i \right].
\]

Therefore, \( OPT(\text{PBP}) \leq E_{\omega} \left[ \sum_{i} \lambda^\omega_i + \sum_{i} \theta^\omega_i \right]. \)

**Proof.** Fix an arbitrary feasible solution \( \{x^\omega_{it}\} \) for PBP, then for every edge \((i,t)\) and sample path \( \omega \) we have,

\[
\lambda^\omega_i \sum_{t} x^\omega_{it} + \theta^\omega_i \sum_{t} x^\omega_{it} 1^\omega(i,t) \leq \lambda^\omega_i + \theta^\omega_i.
\]
Where we multiplied inequality (2) for $i, \omega$ by $\lambda_i^\omega$ and inequality (1) for $t, \omega$ by $\theta_t^\omega$, and used non-negativity. Following the standard procedure for LP duality, let us sum over all $(i, t)$ for a fixed $\omega$,

$$\sum_{(i,t)} x_{it}^\omega \lambda_i^\omega + \sum_{(i,t)} x_{it}^\omega \theta_t^\omega \mathbb{1}^\omega(i, t) \leq \sum_t \lambda_t^\omega + \sum_i \theta_i^\omega$$

Next, taking expectation (convex combination of linear constraints) over sample paths $\omega$ and using $E_{\omega}[\lambda_t^\omega + \theta_t^\omega \mathbb{1}^\omega(i, t) \mid \omega^{t-1} = \omega^{t-1}] \geq p_t r_i$ we have,

$$E_{\omega}\left[ \sum_t \lambda_t^\omega + \sum_i \theta_i^\omega \right] \geq E_{\omega}\left[ \sum_{(i,t)} x_{it}^\omega (\lambda_t^\omega + \theta_t^\omega \mathbb{1}^\omega(i, t)) \right]$$

$$= \sum_{(i,t)} E_{\omega^{t-1}} \left[ E_{\omega}\left[ x_{it}^\omega (\lambda_t^\omega + \theta_t^\omega \mathbb{1}^\omega(i, t)) \mid \omega^{t-1} = \omega^{t-1} \right] \right]$$

$$\geq p_t r_i \sum_{(i,t)} E_{\omega}\left[ x_{it}^\omega \right]$$

Where we used the tower property of expectation and condition (3) to get the final inequality. \hfill \Box

Remarks: The above lemma implies that we need to set dual variables such that the ‘dual constraint’,

$$\lambda_t^\omega + \theta_t^\omega \mathbb{1}^\omega(i, t) \geq \mathbb{1}^\omega(i, t) r_i,$$

is satisfied for every edge $(i, t)$ in expectation, conditioned on knowing the success failure of all edges before $t$. Demanding dual feasibility in expectation may seem similar to the randomized primal-dual framework of [DJK13]. However, there is a subtle but important distinction here. The dual fitting in [DJK13] actually leads to a feasible solution for the dual program, which is not the case here. Recall that in [DJK13], if the online algorithm matches $i$ to $t$, we set $\lambda_t = r_i(1 - g(\omega_i))$ and $\theta_t = r_i g(\omega_i)$. Let $M(\mathbf{Y}, t)$ denote the match of arrival $t$ when random values are given by $\mathbf{Y}$ (if $t$ is unmatched let $r_{M(\mathbf{Y}, t)} = 0$). Also, let $\mathbb{1}(M(\mathbf{Y}, i))$ be the indicator that denotes whether $i$ is matched for random vector $\mathbf{Y}$. Now, consider instead the following dual fitting,

$$\lambda_t' = \frac{1}{(1 - 1/e)} E_{\mathbf{Y}}[\lambda_t] = \frac{1}{(1 - 1/e)} E_{\mathbf{Y}}[\mathbb{1}(M(\mathbf{Y}, t))(1 - g(\mathbb{1}(M(\mathbf{Y}, t))))] \quad \forall t \in T$$

$$\theta_t' = \frac{1}{(1 - 1/e)} E_{\mathbf{Y}}[\theta_t] = \frac{1}{(1 - 1/e)} r_i E_{\mathbf{Y}}[\mathbb{1}(M(\mathbf{Y}, i)) g(\omega_i)] \quad \forall i \in I.$$

Due to [DJK13], this setting of dual variables gives a feasible dual solution for the classical case i.e.,

$$\lambda_t' + \theta_t' \geq r_i,$$

for every edge $(i, t)$. In contrast, our fitting will be dual infeasible and only satisfy the weaker constraint given in Lemma 1 which is a specific convex combination of the dual constraints. The interested reader may refer to Appendix A for further discussion on the issue with proceeding directly via standard duality.

Armed with Lemma 1 we now focus on finding dual variables $\lambda_t^\omega \geq 0$ and $\theta_t^\omega \geq 0$ such that,

$$E_{\omega}\left[ \sum_t \lambda_t^\omega + \sum_i \theta_i^\omega \right] = E_{\omega, \mathbf{Y}}[\text{Reward for Algorithm 1}],$$

and for every $(i, t) \in E$,

$$E_{\omega}[\lambda_t^\omega + \theta_t^\omega \mathbb{1}^\omega(i, t) \mid \omega^{t-1} = \omega^{t-1}] \geq \alpha p_t r_i.$$
Equation \((4)\) requires that the expected sum of dual variables matches the expected reward of Algorithm \(\Pi\) where expectation is over the stochastic rewards and random variables \(Y = (y_i)\) (samples denoted using \(Y = (y_i)\)) generated by the algorithm. Inequalities \((5)\) require that the dual feasibility condition required by Lemma \(\Pi\) is met to within a constant factor. If we can find such variables, then \(\lambda^i_t/\alpha\) and \(\theta^i_t/\alpha\) would satisfy the condition in Lemma \(\Pi\) resulting in a proof of \(\alpha\) competitiveness for Algorithm \(\Pi\). In what follows, we will find such a setting of duals with \(\alpha = (1 - 1/\epsilon)\).

Consider the following process for setting duals. Initialize variables \(\lambda^i_t, Y\) and \(\theta^i_t, Y\) to 0. Consider the online arrival process and execute Algorithm \(\Pi\) for fixed \(Y = (y_i)\). Let \(\omega\) be the sample path of the stochastic rewards experienced by the algorithm. Now, whenever the algorithm offers \(i\) to \(t\) set,

\[
\lambda^i_t = r_i(1 - g(y_i)) \mathbb{1}^\omega(i, t) \quad \text{and} \quad \theta^i_t = r_i g(y_i) \mathbb{1}^\omega(i, t).
\]

Clearly, \(\lambda^i_t\) is set uniquely since the algorithm offers at most one \(i\) to arrival \(t\), and \(\theta^i_t\) takes a non-zero value only if it is also accepted by some \(t\), and if this occurs \(\theta^i_t\) is never re-set. Taking expectation over \(Y = (y_i)\), we define our dual variables \(\lambda^i_t, \theta^i_t\) as,

\[
\lambda^i_t = E_Y [\lambda^i_t] \quad \text{and} \quad \theta^i_t = E_Y [\theta^i_t].
\]

First, note that if the algorithm matches \(t\) to \(i\) we have, \(\lambda^i_t + \theta^i_t = r_i \mathbb{1}^\omega(i, t)\). Therefore, \((4)\) is satisfied and we have,

\[
E_{\omega, Y} [\text{Reward for Algorithm } \Pi] = E_{\omega, Y} [\sum_t \lambda^i_t + \sum_i \theta^i_t] = E_\omega [\sum_t \lambda^t + \sum_i \theta^i].
\]

Moreover, for all edges \((i, t)\) that are output by the algorithm for a given \(Y\) we have, \(E_\omega [\lambda^i_t + \theta^i_t \mathbb{1}^\omega(i, t)] = p_i r_i\). Thus, we do not face the issue highlighted in Section \(\Pi\) where setting dual variables similar to above for the expectation based LP led to a \(p^2_\omega\) term in the dual constraints.

In the rest of this section, we focus on showing that inequalities \((5)\) are satisfied for every edge \((i, t)\), with \(\alpha = (1 - 1/\epsilon)\), as long as the probabilities are decomposable. We are interested in conditional expectations where the values of variables \(\mathbb{1}^\omega(i, t')\) are fixed for all \(t' \leq t - 1\). So we proceed by fixing an arbitrary sample path up to time \(t - 1\) and perform the rest of the analysis conditioned on this path. So all expectations for the rest of this section are conditioned on \(\omega^{t-1} = \omega^{t-1}\), and we suppress this in the notation. Also for notational convenience, we write \(\lambda^i_t, \lambda^i_t, Y\) simply as \(\lambda^i_t, \lambda^i_t\), and similarly, \(\theta^i_t, \theta^i_t, Y\) as \(\theta^i_t, \theta^i_t\). So, expectation \(E_\omega [\cdot | \omega^{t-1} = \omega^{t-1}]\) is written simply as \(E[\cdot]\), but we continue to use \(E_Y [\cdot]\) to distinguish the two expectations. This is summarized in the following.

**Proposition 1.** Arbitrarily fix an edge \((i, t) \in E\) and a sample path up to (and including) arrival \(t - 1\) i.e., \(\omega^{t-1} \in \Omega\). Also fix (arbitrary) values \(y_j\) for \(j \neq i\), denoted \(Y_{-i}\). Then for duals set according to \((5)\), it suffices to show,

\[
E_Y [E_\omega [\lambda^i_t + \theta^i_t \mathbb{1}^\omega(i, t)]] \geq (1 - 1/\epsilon)p_i r_i.
\]

Where the inner expectation is conditioned on \(\omega^{t-1} = \omega^{t-1}\), and we ignored \(\omega\) from sub/superscript.

To ease comparing the performance of Algorithm \(\Pi\) to clairvoyant, we use the following coupling for the analysis.

**Proposition 2.** A sample path \(\omega \in \Omega\) for the stochastic rewards is randomly sampled. Instead of using different sample paths for clairvoyant and Algorithm \(\Pi\), we use the same path for both. So whenever a match is made by either algorithm, we use the corresponding variable from \(\omega\) to see if it is successful. Hence w.l.o.g., both algorithms are subject to the same values \(\mathbb{1}^\omega(i, t)\).
To show the inequality in Proposition\(^1\), we separately lower bound \(E_{\lambda_i}(E[\theta_i^Y])\) and \(E_{\lambda_i}(E[\theta_i^Y 1(i, t)])\). The lower bound on \(E_{\lambda_i}(E[\theta_i^Y])\) will be similar to Lemmas 1 and 2 in [DJK13]. In contrast, the bound on \(E_{\lambda_i}(E[\theta_i^Y 1(i, t)])\) has an interesting subtlety absent in the classical case, owing to the fact that matches can be unsuccessful. This is also where we need the decomposability assumption for edge probabilities. We discuss this further in the proof of Lemma\(^3\).

To proceed, consider the ‘matching’ \(M_{-i}\) given by Algorithm\(^1\) when it is executed with the reduced set of vertices \(I \setminus i\) (recall, the values \(y_i\) are fixed for all \(j \in I \setminus i\)). Unlike the deterministic case, here \(M_{-i}\) may have one offline vertex matched multiple times (though only one match could have actually succeeded). Let \(I'_{-i}\) denote the set of available neighbours of \(i\), in this execution of Algorithm\(^1\) without vertex \(i\). Define \(y_i^c\) such that \(p_{it}r_i(1 - g(y_i^c)) = \max_{j \in I'_{-i}} p_{jt}(1 - g(y_j))\).

Set \(y_i^c = 1\) if the set \(I'_{-i}\) is empty, and \(y_i^c = 0\) if no such value exists. Due to the monotonicity of \(g(t) = e^{t-1}\), we have a unique value of \(y_i^c\).

**Lemma 2.** With \(Y_{-i}\) fixed, \(E[\lambda_i^Y] \geq p_{it}r_i(1 - g(y_i^c))\) for every \(y_i \in [0, 1]\). Thus, \(E_{\lambda_i}(E[\lambda_i^Y]) \geq p_{it}r_i(1 - g(y_i^c))\).

**Proof.** Let \(I'_{y_i}\) be the set of available neighbours at \(t\) when Algorithm\(^1\) is executed with the full vertex set and value \(y_i\) for \(i\). It suffices to show that for every \(y_i\), \(\max_{j \in I'_{y_i}} p_{jt}(1 - g(y_j)) \geq p_{it}(1 - g(y_i^c))\), which follows directly from \(I'_{y_i} \subseteq I'_{-i}\). So fix \(y_i\) and suppose \(t'\) is the first arrival \(i\) is offered to by the algorithm. If \(t' \geq t\), the output of the algorithm prior to arrival \(t\) coincides with \(M_{-i}\), and we have \(I'_{y_i} = I'_{-i} \cup \{i\}\). When \(t' < t\), the set of available matches for arrival \(t' + 1\) is a superset (not necessarily strict) of the set \(I'_{y_i}\). Inductively, this is true for every arrival after \(t'\), including \(t\), giving us the desired.

**Lemma 3.** If the probabilities decompose such that \(p_{it} = p_{it} p_t\) for every \((i, t) \in E\), then for fixed \(Y_{-i}\) we have, \(E_{\lambda_i}(E[\theta_i^Y 1(i, t)]) \geq p_{it} r_i \int_0^{\theta_i^Y} g(x)dx\).

**Proof.** We focus on the interval \(y_i \in [0, y_i^c]\). There are two possibilities, either \(i\) is successfully matched before \(t\) and unavailable for \(t\), or \(i\) is available when \(t\) arrives. We show that in the latter scenario the algorithm matches \(i\) to \(t\). Let us first see how this proves the claim in the lemma. When \(i\) is unavailable for \(t\), \(\theta_i^Y = r_i g(y_i)\) since \(i\) was successfully matched to an arrival preceding \(t\). In case \(i\) is available and thus matched to \(t\) (by assumption), \(\theta_i^Y = r_i g(y_i)^1(i, t)\). Therefore, in both cases \(E[\theta_i^Y 1(i, t)] = p_{it} r_i g(y_i)\). Since this holds for all values of \(y_i \leq y_i^c\) we have, \(E_{\lambda_i}(E[\theta_i^Y 1(i, t)]) \geq p_{it} r_i \int_0^{\theta_i^Y} g(x)dx\).

To finish the proof we need to show that for \(y_i \leq y_i^c\), \(i\) is matched to \(t\) if available. In the classical/deterministic case this follows directly from \(r_i(1 - g(y_i)) \geq r_i(1 - g(y_i^c))\), and the fact that \(i\) could not have been offered to any arrival prior to \(t\) (since in that case \(i\) would be unavailable at \(t\)). In our case, we still have \(p_{it} r_i(1 - g(y_i)) \geq p_{it} r_i(1 - g(y_i^c))\). However, for some small enough value of \(y_i\), \(i\) may be unsuccessfully matched to some arrival \(t'\) preceding \(t\), freeing up some \(j\neq i\) that was successfully matched to \(t'\) for larger \(y_i\). If \(i\) is not matched offered to any arrival preceding \(t\) then the claim follows as in the deterministic case. So the interesting case is when \(i\) is matched (unsuccessfully) prior to \(t\) and consequently \(t\) is matched to some \(j\neq i\) such that, \(E[\lambda_i^Y] = p_{it} r_j(1 - g(y_j)) > p_{it} r_i(1 - g(y_i^c))\). In this case, consider the graph given by the difference between the current ‘matching’ with value \(y_i\) (including the unsuccessful edges) and the ‘matching’ \(M_{-i}\) (including unsuccessful edges), where \(i\) is removed from consideration during the execution. On this difference graph, there exists a unique alternating path that includes both \(i\) and \(t\). Using the decomposition of probabilities, for every edge \((i', t')\) on the alternating path we have, \(p_{it}(1 - g(y_i)) \geq p_{it} r_i(1 - g(y_i^c))\) (note we may have \(p_{it} r_i(1 - g(y_i)) < p_{it} r_i(1 - g(y_i^c))\) ). In particular, \(p_{it} r_i(1 - g(y_i)) \geq p_{it} r_i(1 - g(y_i^c))\) and thus, \(p_{it} r_i(1 - g(y_i)) \geq p_{it} r_i(1 - g(y_i^c))\). Therefore, if \(i\) is available on arrival of \(t\) the algorithm matches \(i\) to \(t\) (since ties occur w.p. 0).

While this completes the proof, observe that if there are no successful matches prior to \(t\) in the matching \(M_{-i}\), then \(t\) is matched to \(i\) (if available) for all \(y_i \leq y_i^c\) for arbitrary (not necessarily decomposable) probabilities, and the claim in the lemma still holds.

\[\square\]
2.2 Vanishing Probabilities ($p$)

As we saw in the example above, ensuring dual feasibility for every edge ($a_1$) does not hold and the previous expectation can be lower bounded. We will focus on the dual feasibility of constraint corresponding to edge ($i, t$) so we fix $y_j, y_k$ with $y_j > y_k$ and consider the following sample path before $t$ arrives, $1(k, t'') = 0, 1(i, t'') = 1$ and $1(i, t') = 0, 1(j, t') = 1$. Observe that $y_j^r = 1$ and consider $E[y_j^r]$ as $y_i$ varies. For $y_i \geq y_j, k$ is offered to $t''$, $j$ is offered to and accepted by $t'$ and $i$ is offered to $t$ therefore, $E[y_j^r] = p_{it} r_i g(y_i)$ and $E[y_j^r] = p_{it} r_i (1 - g(y_i))$. For $y_k < y_i < y_j, i$ is offered to but not accepted by $t'$ and suppose $p_{it}$ is sufficiently smaller than $p_{jt}$ so $j$ is offered to $t$. Thus, $E[y_j^r] = p_{it} r_i (1 - g(y_j))$ but $\theta_j^i = 0$ for $y_k < y_i < y_j$. For $y_i \leq y_k, i$ is offered to and accepted by $t''$, $j$ is offered to and accepted by $t'$ and $t$ is unmatched with $\lambda_j^y = 0 = p_{it} r_i (1 - g(y_j))$ and $E[y_j^r] = p_{it} r_i g(y_i)$. Combining all pieces we have

$$E_y [E[y_j^r + \theta_j^i 1(i, t)]] = p_{it} r_i (1 - y_j) + p_{it} r_i (1 - g(y_j)) (y_j - y_k) + p_{it} \int_0^{y_k} g(x) dx$$

Clearly, Lemma 3 does not hold and the previous expectation can be lower bounded. More concretely, let $y_j = 1 - \epsilon$ and $p_{it}, p_{it}$ be such that $p_{it} (1 - g(y_j)) = p_{it} (1 - g(y_k))$. Then for $g(x) = e^{x-1}$,

$$E_y [E[y_j^r + \theta_j^i 1(i, t)]] = O(\epsilon) + p_{it} r_i (1 - g(y_k)) (1 - y_k) + p_{it} \int_0^{y_k} g(x) dx$$

$$= O(\epsilon) + p_{it} r_i \min_y \left(1 - g(y) (1 - y) + \int_0^y g(x) dx\right)$$

$$< O(\epsilon) + 0.44 p_{it} r_i (for \ y = 0.5571)$$

Note, we chose the worst case $y$ above but in reality this would involve an expectation over random variables $y_j, y_k$ that we fixed. In general, an analysis that considers this more complicated expectation could help tackle such issues. Also note, using our analysis for $g(x) = 1/2$ we get a 1/2 competitive (greedy) algorithm for arbitrary probabilities.

### 2.2 Vanishing Probabilities ($p_{it} = o(1)$)

As we saw in the example above, ensuring dual feasibility for every edge ($i, t$) (while satisfying $a_1$) appears to be much more involved (and may not always be possible) when the probabilities do not decompose. When examining the dual constraint for an edge ($i, t$), a key difference between the proof of Lemma 3 and the equivalent statement for the classical deterministic case, arises from the fact that in the stochastic case $i$ could be unsuccessfully matched to vertices arriving before $t$. In the deterministic case where matches always succeed, $i$ being matched to a previous arrival ensures a lower bound on the value of $E_y[\theta_j^i]$. From an intuitive standpoint, while a match may not always succeed in the stochastic case, every time $i$ is matched its probability of being matched successfully when the algorithm terminates certainly increases. This motivates us to consider an approach where we are not focused on showing feasibility for every edge ($i, t$), but rather on showing that an online algorithm matches $i$ with a suitably large probability in comparison to the clairvoyant. This allows us to move away from the local issues that seem to arise only over the feasibility of a single edge. To this end, let us recall a keen and remarkable insight first made in [MP12]. They considered the case of identical edge probabilities $p_i$ and observed that instead of letting the success of each match be determined by an i.i.d. random variable, one could consider the following process. For each vertex $i \in I$, independently generate a threshold value $B_i$ according to the following distribution,

$$\mathbb{P}(B_i = \tau p) = p(1 - p)^{\tau}.$$
A match made to $i$ succeeds only if the threshold is met i.e., if the threshold is $\tau p$, the $\tau + 1$-th match made to $i$ would succeed, and none before. These thresholds are generated in advance, but the algorithm (offline or online) is unaware of the values. This process is probabilistically equivalent to the original stochastic rewards setting, but with an important contrast for our analysis. Whereas in the previous analysis, for every edge $(i, t)$ we fixed a sample path up to arrival $t$ and considered an expectation over the edges incident on $t$. Leveraging this new view of the reward process we can now fix thresholds $B_{-i}$ for all vertices in $I \setminus \{i\}$, and take expectation over $B_i$ while considering all relevant matches made to $i$, instead of just a single edge $(i, t)$. We shall make this more formal later, but to effectively use this insight we add the assumption that the probabilities are all $o(1)$. Observe that under this assumption the distribution for the thresholds is well approximated by an exponential r.v. i.e.,

$$[1 - e^{-\tau p}] \leq \mathbb{P}(B_t \leq \tau p) \leq [1 - e^{-\tau p}](1 + O(p)). \quad (7)$$

So for $p = o(1)$, using the exponential distribution as an approximation only incurs a multiplicative $(1 + o(1))$ error. More generally, when the probabilities are heterogeneous but $o(1)$, the above approximation still holds i.e., we can randomly set a threshold $B_i$ using the exponential distribution above, and a match made to $i$ only succeeds if the threshold is crossed.

More formally, consider an arbitrary algorithm $A$ (could be online of offline) that does not know the outcomes of future rewards. Let $A(S, t, s)$ denote the matching decision made by the algorithm for arrival $t$, set $S$ of available offline vertices, and random seed (for a randomized algorithm) $s$. Then the stochastic reward process is equivalently simulated by the following.

**ALGORITHM:** Thresholding Process for Heterogeneous Probabilities

**Input:** Algorithm $A$, offline vertex set $I$;
$S = I$;
For every $i \in I$, $l_i = 0$;
Generate i.i.d. r.v. $B_i \in \text{Exponential}(1)$;

**for** every new arrival $t$ **do**

Let $(i^*, t) = A(S, t, s)$;
$l_{i^*} = l_{i^*} + p_{i^*} t$;
If $l_{i^*} > B_{i^*}$ **update**;
$S = S \setminus \{i^*\}$ and generate reward $r_{i^*}$;

end

Given that $i$ is available at arrival $t$, the probability of success of match $(i, t)$ is given by $e^{-\tau p} p_i \approx p_i$, under the vanishing probabilities assumption. So the independence from past failures is preserved through the memoryless property of exponentials. Notice that the error term $(1 + O(p))$ in (7) now becomes $(1 + O(p_{\text{max}}))$, where we let $p_{\text{max}}$ denote the maximum probability $p_i$ over all edges. In what follows, we leverage this perspective to derive a new ‘global dual’ constraint.

In the following, we refer to a sample path by referring to the vector of thresholds $B$. Fix an algorithm $A$ as above (possibly offline, but does not know future outcomes). Focus on an offline vertex $i$ and fix arbitrary values $B_{-i}$, and for a randomized algorithm also fix the seed $s$. Suppose $B_i = \infty$, i.e., none of the matches that the algorithm makes to $i$ succeed. Then we use $\tau(B_{-i}, s)$ to denote the sum of the probabilities over all matches to $i$ made by the algorithm for this setting of thresholds and seed $s$. In other words $\tau(B_{-i}, s) = l_i(T)$, where $T$ is the last arrival and $l_i(T)$ denotes the value of variable $l_i$ given by the thresholding process above, at termination of $A$. Since the algorithm does not know future outcomes, we have the following key observation.

For every value $B_i$ that is less than $\tau(B_{-i}, s)$, $A$ successfully matches $i$ to some arrival. We therefore call $\tau(B_{-i}, s)$, the effort threshold of $A$ for vertex $i$. Using this insight, we can
now write the expected reward of $A$ as follows,

$$E_B[\text{Reward of } A] = \sum_i r_i E_{B_{-i}, s}[\mathbb{P}(B_i \leq \tau(B_{-i}, s))].$$

$$\leq (1 + O(p_{\max})) \sum_i r_i E_{B_{-i}, s}[e^{-\tau(B_{-i}, s)}]. \tag{8}$$

Where recall, the $(1 + O(p_{\max}))$ term captures the approximation error. Now, let OPT$(B, s)$ denote the matching generated by clairvoyant for threshold values $B$ and random seed $s$. Abusing notation, henceforth let $\tau(B_{-i}, s)$ denote the effort threshold of clairvoyant for vertex $i$.

Observe that the expected revenue for any algorithm $A$ is also given by,

$$\sum_{(i,t)} p_{it} r_i E_{B, s}[x_{it}^{B,s}].$$

Where, $x_{it}^{B,s}$ is an indicator r.v. that is 1 iff $A$ matches $i$ to $t$ for thresholds $B$ and seed $s$. Therefore, we can associate with match $(i, t)$ a deterministic reward $p_{it} r_i$, regardless of the outcome of the match. Given this observation, we are now ready to define our setting of dual values. Fix a set of values $B$ and consider a run of Algorithm 2 (ALG), with output given by ALG$(B)$. Then, for every match $(i, t)$ in ALG$(B)$ we set,

$$\lambda_t^B = p_{it} r_i g(t^{ALG}(t)) \quad \text{and} \quad \theta_t^B = p_{it} r_i (1 - g(t^{ALG}(t))).$$

Where we denote the sum $\sum_{(i,t') \in \text{ALG}(B)} \mu_{it'}$ as $t^{\text{ALG}}_i(t)$. Note that for every $(i, t) \in \text{ALG}(B)$, the sum $\lambda_t^B + \theta_t^B$ equals $p_{it} r_i$. Letting,

$$\theta_t^B = \sum_{(i,t) \in \text{ALG}(B)} \theta_{it}^B,$$

we can write the total reward of the algorithm as the sum $\sum_i \lambda_t^B + \sum_i \theta_t^B$. We can further rewrite this sum as the follows,

$$E_B[\text{Reward of Algorithm 2}] = E_{B, s}[\sum_i \lambda_t^B + \sum_{t|(i,t) \in \text{OPT}(B,s)} \theta_t^B]. \tag{9}$$

Where we used the fact that for a fixed seed $s$, clairvoyant matches at most one vertex to every arrival. So if we can show that for every $i, B_{-i}, s$, the following ‘dual’ constraint is satisfied,

$$E_{B_i}[\theta_t^B + \sum_{t|(i,t) \in \text{OPT}(B,s)} \lambda_t^B | B_{-i}] \geq (1 - O(p_{\max}))0.596 \cdot r_i (1 - e^{-\tau(B_{-i}, s)}).$$

Then, combining this with (8), we have the following,

$$\sum_i E_{B, s}[\theta_t^B + \sum_{t|(i,t) \in \text{OPT}(B,s)} \lambda_t^B] = \sum_i E_{B_{-i}, s}[E_{B_i}[\theta_t^B + \sum_{t|(i,t) \in \text{OPT}(B,s)} \lambda_t^B | B_{-i}]]$$

$$\geq \sum_i (1 - O(p_{\max}))0.596 \cdot E_{B_{-i}, s}[r_i (1 - e^{-\tau(B_{-i}, s)})]$$

$$= (1 - O(p_{\max}))0.596 \cdot E_{B, s}[\text{Reward of Clairvoyant}].$$

Using (9) then establishes 0.596-competitiveness for vanishing probabilities. It remains to give a proof of (10) for our setting of duals.
Lemma 4. Let $g(.)$ be a strictly decreasing function and let,

$$f(x) = 1 - e^{-x}[1 - g(x)(x + 1)] - \int_0^x g(z)e^{-z}dz.$$ 

Suppose that,

$$\min_{x \geq 0} f(x) = g(0).$$

Then for every $i$, $B_{-i}$, $s$, we have,

$$E_{B_i} [\theta^B_i + \sum_{t(i,t) \in \text{OPT}(B,s)} \lambda^B_t \mid B_{-i}] \geq (1 - O(p_{max}))g(0) \cdot r_i (1 - e^{-\tau(B_{-i},s)}).$$

Proof. Fix $i, B_{-i}$ and $s$. We shall take an expectation over $B_i$. Recall that $\tau(B_{-i},s)$ denotes the effort threshold for clairvoyant (OPT). Similarly, let $\tau'(B_{-i})$ be the effort threshold for Algorithm 2 (ALG). In the following, we denote $\tau(B_{-i},s)$ and $\tau'(B_{-i})$ more simply as $\tau$ and $\tau'$ respectively. Now clearly, ALG successfully matches $i$ iff $B_i < \tau'$. Using this, we can write for any given $B_i$,

$$\theta^B_i \geq r_i \left\int_0^{\min\{\tau', B_i\}} (1 - g(y))dy - \epsilon_1(\min\{\tau', B_i\}) \right\.$$ 

$$= r_i \left[\min\{\tau', B_i\} - \int_0^{\min\{\tau', B_i\}} g(y)dy \right] - r_i \epsilon_1(\min\{\tau', B_i\}).$$

Where we used the vanishing probabilities assumption to write the sum $\sum_{t(i,t) \in \text{ALG}(B)} r_i (1 - g(t^{\text{ALG}}(t)))p_{it}$, as an integral. Since $(1 - g(\cdot))$ is an increasing function, the error incurred due to this approximation is given by,

$$\epsilon_1(\min\{\tau', B_i\}) = O(p_{max})[g(0) - g(\min\{\tau', B_i\})].$$

Taking expectation over $B_i$, we have,

$$E_{B_i} [\theta^B_i \mid B_{-i}] \geq r_i \left\int_0^{\tau'} \left( z - \int_0^z g(y)dy - \epsilon_1(\tau) \right)e^{-z}dz \right\.$$ 

$$+ \int_{\tau'}^\infty \left( \tau' - \int_0^{\tau'} g(y)dy - \epsilon_1(\tau') \right)e^{-z}dz \right\.$$ 

Where,

$$\epsilon_2(\tau') = O(p_{max})[\tau' - \int_0^{\tau'} g(y)dy]e^{-\tau'},$$

denotes the error due to the first integration (ignoring $O(p_{max}^2)$ terms due to $\epsilon_1(\cdot) \cdot \epsilon_2(\cdot)$). The second integration can be written as $c \int_0^{\infty} e^{-z}dz$, for constant $c$. This incurs no error since $e^{-z}$ is decreasing and the integral approximation is in fact, a lower bound on the actual sum.

Now, for $\lambda^B_t$ we have that for $B_i \geq \tau'$,

$$\sum_{t(i,t) \in \text{OPT}(B,s)} \lambda^B_t \geq r_i g(\tau') \int_0^{\min\{\tau, B_i\}} dy = r_i g(\tau') \min\{\tau, B_i\}.$$

Note that in this case there is no error term as we integrate a constant. For $B_i < \tau'$ the contribution from $\sum_{t(i,t) \in \text{OPT}(B,s)} \lambda^B_t$ can be as small as zero, since ALG may successfully
match $i$ before it is ever matched in $\text{OPT}$. Letting $\int g(z)dz = G(z) + c$, consider the expectation over $B_i$,

$$E_{B_i} \left[ \sum_{t \in \text{OPT}(B_s)} \lambda^B_t \mid B_{-i} \right] \geq r_i \left[ \int_{\tau'}^{\infty} g(\tau') \min \{ \tau, z \} e^{-z} dz - \epsilon_3(\tau, \tau') \right]$$

$$= r_i g(\tau') \left[ \int_{\tau'}^{\max \{ \tau', \tau' \}} ze^{-z} dz + \tau \int_{\min \{ \tau, \tau' \}}^{\infty} e^{-z} dz \right] - r_i \epsilon_3(\tau, \tau')$$

$$= r_i g(\tau') \left[ \tau' e^{-\tau'} - \max \{ \tau, \tau' \} e^{-\max \{ \tau, \tau' \}} + e^{-\tau'} - e^{-\max \{ \tau, \tau' \}} \right]$$

$$+ \tau e^{-\max \{ \tau, \tau' \}} - r_i \epsilon_3(\tau, \tau')$$

$$\geq -r_i \epsilon_3(\tau, \tau') + r_i g(\tau') \begin{cases} \tau e^{-\tau'} & \tau \leq \tau' \\ \tau' e^{-\tau'} + e^{-\tau'} - e^{-\tau} & \tau > \tau' \end{cases}$$

Where note that the integral $\int_{\max \{ \tau, \tau' \}}^{\infty} e^{-z} dz$ is a lower bound on the actual sum it replaces. Therefore, the error $\epsilon_3(\tau, \tau')$ must only account for the approximation due to $\int_{\max \{ \tau, \tau' \}}^{\infty} ze^{-z} dz$. The function $ze^{-z}$ is decreasing for $z \geq 1$ therefore, the integral only incurs an error for the portion intersecting with $z \in [0, 1]$. Hence,

$$\epsilon_3(\tau, \tau') = O(p_{\max}) \mathbb{1}(\tau' \leq \min \{ 1, \tau \}) \cdot g(\tau') [\min \{ 1, \tau \} e^{-\min \{ 1, \tau \}} - \tau' e^{-\tau'}]. \quad (10)$$

Where $\mathbb{1}(\tau' \leq \min \{ 1, \tau \})$ is one if $\tau' \leq \min \{ 1, \tau \}$, and zero otherwise.

We now split the analysis in two parts. In the first part, we ignore all the approximation errors and show that the resulting sum satisfies the dual condition. In the second part, we finish the proof by showing that the loss due to errors leads to at most a multiplicative factor of $(1 - O(p_{\max}))$. So ignoring the error terms, consider the following,

$$\hat{\theta}^{B_{-i}} = r_i \left[ \int_0^{\tau'} \left( z - \int_0^z g(y) dy \right) e^{-z} dz \right.$$

$$+ \int_{\tau'}^{\infty} \left( \tau' - \int_0^{\tau'} g(y) dy \right) e^{-z} dz \right] \quad (11)$$

$$= r_i \left[ -\tau' e^{-\tau'} + 1 - e^{-\tau'} - \int_0^{\tau'} (G(z) - G(0)) e^{-z} dz \right.$$  

$$+ \tau' e^{-\tau'} - (G(\tau') - G(0)) e^{-\tau'} \right] \quad (12)$$

$$= r_i \left[ 1 - e^{-\tau'} - \int_0^{\tau'} g(z) e^{-z} dz \right]$$

Where $\int g(z)dz = G(z) + c$. Similarly let,

$$\hat{\lambda}^{B_{-i}} = r_ig(\tau') \begin{cases} \tau e^{-\tau'} & \tau \leq \tau' \\ \tau' e^{-\tau'} + e^{-\tau'} - e^{-\tau} & \tau > \tau' \end{cases}$$

Note that for $\tau \to \infty$, we have $\hat{\theta}^{B_{-i}} + \hat{\lambda}^{B_{-i}} = f(\tau')$. Now, we wish to show,

$$H(\tau, \tau') := \frac{\hat{\theta}^{B_{-i}} + \hat{\lambda}^{B_{-i}}}{r_i(1 - e^{-\tau})} \geq g(0) \quad \forall \tau, \tau' \geq 0.$$  

Observe that $H(\tau, 0)$ equals $g(0)$, for all $\tau \geq 0$. Therefore, we focus on the more interesting case of $\tau' > 0$. Here since $g(\cdot)$ is strictly decreasing we have $g(\tau') < g(0)$. Suppose that there
exists some $\tau = \tau_1$ and $\tau' = \tau_2 > 0$, such that $H(\tau_1, \tau_2)$ is smaller than $g(0)$. We will derive a contradiction. To this end, consider the following derivative w.r.t. $\tau$,

$$\frac{d[\hat{\theta}^{B-i} + \hat{B}^{B-i}]}{d\tau} = r_1 g(\tau') \begin{cases} e^{-\tau'} & \tau \leq \tau' \\ e^{-\tau} & \tau > \tau' \end{cases}$$

Observe that the above piecewise function is strictly smaller than $r_1 g(0) e^{-\tau}$ for all $\tau \geq 0$. Therefore, $H(\tau, \tau_2) < g(0)$ for every $\tau \geq \tau_1$ (as $\tau$ increases, the numerator increases at a rate strictly smaller than $g(0)$ times the rate of increase in denominator). This implies that $H(\infty, \tau_2) = f(\tau_2) < g(0)$, which in turn implies that $0$ is not a minimizer of $f(\cdot)$, contradiction. Therefore, $H(\tau, \tau') \geq g(0)$ for all values $\tau, \tau' \geq 0$.

To finish the proof, we need to address the error terms. In particular, we will show that,

$$\int_0^{\tau'} r_1 e_1(z) e^{-z} dz + r_1 e_2(\tau') + r_1 e_1(\tau') e^{-\tau'} \leq O(p_{\max}) \hat{\theta}^{B-i} \quad \text{and} \quad r_1 e_3(\tau, \tau') \leq O(p_{\max}) \hat{B}^{B-i}.$$

Combined, this will give us the desired. Note that $r_1 e_2(\tau') = O(p_{\max}) r_1 [\tau' - \int_0^{\tau'} g(z) dz] e^{-\tau'}$ is $O(p_{\max}) \hat{\theta}^{B-i}$ by definition (see (11)). For the terms corresponding to $e_1(\cdot)$ we have,

$$\int_0^{\tau'} r_1 e_1(z) e^{-z} dz + r_1 e_1(\tau') e^{-\tau'} = O(p_{\max}) r_1 \left[ \int_0^{\tau'} (g(0) - g(z)) e^{-z} dz + (g(0) - g(\tau')) e^{-\tau'} \right]$$

$$= O(p_{\max}) r_1 \left[ g(0) - g(\tau') e^{-\tau'} - \int_0^{\tau'} g(z) e^{-z} dz \right] \leq O(p_{\max}) \hat{B}^{B-i}.$$

Where the last inequality follows from (11). Finally, consider the term $e_3(\tau, \tau')$ given by (11). This is non-zero only for $\tau' < \min\{1, \tau\}$. Therefore, it suffices to show that for all $\tau' < \min\{1, \tau\}$,

$$\min\{1, \tau\} e^{-\min\{1, \tau\}} - \tau' e^{-\tau'} \leq \tau' e^{-\tau'} + e^{-\tau'} - e^{-\tau}$$

(13)

Where the RHS is by definition of $\hat{B}^{B-i}$ for $\tau > \tau'$. Observe that $e^{-z}(z + 1)$ is maximized at $z = 0$, whereas $e^{-z}(2z + 1) \geq 1$ for all $z \in [0, 1]$. Therefore, $e^{-\tau}(\tau + 1) \leq e^{-\tau}(2\tau' + 1)$, proving (14) for $\tau \leq 1$. For $\tau > 1$, we have $e^{-1} + e^{-\tau} < 2e^{-1} < 1$, implying (14) also for $\tau > 1$.

With this we are now ready to finish the proof of Theorem 2.

**Proof of Theorem 2.** The function $g(t) = e^t (e^\infty e^{-t} dy)$, is monotonically decreasing and we have $0.5963 \leq g(0) \leq 0.5964$ (bounds numerically evaluated). It remains to show that $\min_{x \geq 0} f(x) = g(0)$, then a direct application of Lemma 4 gives the desired. Observe that by definition of $g(\cdot)$, we have, $g(t) - g(t') = \frac{t'}{t' + 1}$. Using this one can verify that, $\frac{df(x)}{dx} = 0$. So $f(x)$ is constant for all $x \geq 0$, and we have the desired. \hfill \Box

**Theorem 3.** For $g(t) = \beta_1 \frac{e^{-\beta_2 t}}{\beta_2 e^t + 1}$, with $\beta_1 = 0.588$, $\beta_2 = 0.575$ and $p_{\max} = o(1)$, the competitive ratio of Algorithm 4 is at least 0.588.

**Proof.** Notice that $g(0) = \beta_1 = 0.588$. One can verify analytically/numerically that 0 is a minimizer of $f(\cdot)$ given this choice of $g(\cdot)$. \hfill \Box

Similarly, we have the below result for a modified version of the scaling function $g(\cdot)$ proposed in [MP12, MWZ13].

**Theorem 4.** For $g(t) = \beta_1 e^{-\beta_2 t}$, with $\beta_1 = 0.581$, $\beta_2 = 0.535$ and $p_{\max} = o(1)$, the competitive ratio of Algorithm 4 is at least 0.581.
Remarks: In case of vanishing probabilities, the classical upper bound of \((1 - 1/e)\) does not necessarily hold when comparing against clairvoyant. Unlike the decomposable case, which includes the classical case of unit probabilities, the vanishing probabilities case does not immediately include any previous setting with the \((1 - 1/e)\) barrier. Further, recall that from the thresholding process perspective, the marginal increase in objective value obtained by matching \(i\) to \(t\) is given by \(e^{-l_i(t)}p_{it}\). This value decreases over time as \(i\) is unsuccessfully matched and \(l_i(t)\) increases. As a consequence of these diminishing marginal gains, the hard instances that lead to the \((1 - 1/e)\) bound in deterministic settings are not as bad here, and it may even be possible to break the \((1 - 1/e)\) barrier. In particular, for Algorithm 2 it might be possible to improve on the competitive ratio bound by using some ideas that bring out a (solvable) factor revealing LP. Our current analysis handles each item separately but there is a clear connection between items which for instance, may be captured via a factor revealing LP.

2.3 General Probabilities with Large Inventory

So far we focused on the case where there is exactly one unit of every offline vertex \(i \in I\). The case where we have multiple units \(c_i \geq 1\) for each \(i \in I\), is a special case of the single unit setting. Indeed, recall that under the assumption that \(c_i \to \infty\) for every \(i\), the algorithm of [MSVV05] is \((1 - 1/e)\) competitive against the expectation based LP for arbitrary probabilities. This raises a natural question: Is the path based program PBP equivalent to the expectation based LP for large inventory? In this section, we formally show this to be the case.

Clearly, \(OPT(PBP) \leq OPT(LP)\), and every feasible solution for PBP, can be converted into a feasible solution for the LP, regardless of \(c_i\). For the reverse, we shall use standard Chernoff bounds relying on the assumption of large \(c_i\).

Theorem 5. Let \(c_{\text{min}} = \min_{i \in I} c_i\). Then we have,

\[
OPT(LP) \left(1 - O\left(\sqrt{\frac{\log c_{\text{min}}}{c_{\text{min}}}}\right)\right) \leq OPT(PBP) \leq OPT(LP).
\]

Hence, for \(c_{\text{min}} \to \infty\), \(OPT(PBP) \to OPT(LP)\).

Proof. We focus on the lower bound and more strongly show that every feasible solution of the standard LP can be turned into a feasible solution for PBP, with nearly the same objective value. Let \(\{x_{it}\}_{i,t} \in \mathbb{E}\) be a feasible solution for the standard LP. Then we have for all \(i \in I\),

\[
\sum_t p_{it}x_{it} \leq c_i.
\]

Consider indicator random variables \(\mathbb{1}(i, t)\) that take value 1 w.p. \(p_{it}\). Using Chernoff bounds for binary random variables we have that the value of the sum \(\sum_t \mathbb{1}(i, t)x_{it}\) strongly concentrates around the mean \(\sum_t p_{it}x_{it}\), which is at most \(c_i\). More formally,

\[
P\left(\sum_t \mathbb{1}(i, t)x_{it} \geq c_i(1 + \delta)\right) \leq e^{-\delta^2/c_i}.
\]

For \(\delta = \sqrt{\log c_i/c_i}\), the RHS is \(O(1/c_i)\). Therefore, for most sample paths \(\omega\) over the indicator r.v.s., we have that,

\[
\sum_t \mathbb{1}(i, t)x_{it} \leq c_i\left(1 + \sqrt{\frac{\log c_i}{c_i}}\right).
\]

Using this we construct the following feasible solution for PBP. On all sample paths where the above inequality holds, set \(x_{it}^\omega = \frac{x_{it}}{1 + \sqrt{\log c_i/c_i}}\). If on some path \(\omega\), we have \(\sum_t \mathbb{1}(i, t)x_{it} > \)
then consider the last arrival $t_0$ such that the partial sum, $\sum_{t \leq t_0} 1^\omega(i,t)x_{it}$, is at most $c_i \left(1 + \sqrt{\log c_i} \right)$. Set $x_{it}^\omega = 0$ for all arrivals $t > t_0$. It is easy to see that this is a feasible solution for PBP. Further, with $\delta = \sqrt{\log c_i} / c_i$, the objective value of this solution is at least,

$$\sum_{(i,t)} p_{it} r_i E_\omega[ x_{it}^\omega ] \geq \sum_{(i,t)} p_{it} r_i \left[ \mathbb{P} \left( \sum_{t} 1^\omega(i,t)x_{it} \leq c_i (1 + \delta) \right) \cdot \frac{x_{it}}{1 + \delta} \right]$$

$$\geq \sum_{(i,t)} p_{it} r_i \left(1 - \frac{1}{\sqrt{c_i}}\right) \cdot \frac{x_{it}}{1 + \sqrt{\log c_i} / c_i}$$

$$\geq OPT(LP) \left(1 - O \left( \sqrt{\frac{\log c_{\min}}{c_{\min}}} \right) \right).$$

### 3 Conclusion and Open Problems

We considered a vertex-weighted version of the problem of online matching with stochastic rewards with the goal of designing online algorithms to compete against clairvoyant solutions that know the entire sequence of arrivals in advance but learn the stochastic outcomes online. No prior algorithm beating $1/2$ was known for the problem, even for the special case of identical reward probabilities. We considered the case where the probabilities $p_{it}$ for every edge $(i,t)$ decompose as $p_i \times p_t$ and showed that a natural generalization of the Perturbed-Greedy algorithm gives the best possible competitive guarantee of $(1 - 1/e)$. When probabilities are $o(1)$ but otherwise fully heterogeneous, we showed that a deterministic algorithm that generalizes the Fully-Adaptive algorithm suggested in [MP12], is $0.596$ competitive.

Our analysis in the first case involved a new path based program that better approximates the optimal value of clairvoyant algorithms than previously considered LPs, and an accompanying notion of weak duality to relate the value of an online algorithm to that of the path based program. This analysis runs into a fundamental obstacle for more general probabilities. So for the second case we adopted a new dual fitting approach that develops on an alternative perspective of the stochastic reward process introduced in [MP12]. Using this we showed that it suffices to satisfy ‘global dual’ conditions, and proposed a setting of duals to meet these conditions. To the best of our knowledge, our results are the first to give a provably tighter bound on clairvoyant beyond the standard LP. Finally, we showed that for arbitrary probabilities but large inventory (many copies of each offline vertex), the path based program and the LP are asymptotically equivalent.

Beating $1/2$ without any assumptions on probabilities still remains open for this problem. Similarly, in the more general setting of online assortments, achieving a competitive guarantee better than $1/2$ remains open for unit inventory. A more immediate line of investigation would be to determine the best one could do in the vanishing probabilities case. Here it may in fact, be possible to beat $(1 - 1/e)$.

### Acknowledgements

The authors would like to thank Will Ma for a helpful discussion on the results in [MP12]. This research was partially supported by NSF Grant CMMI 1636046 and NSF CAREER Award CMMI 1351838.
References

[AGKM11] 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, pages 1253–1264. SIAM, 2011.

[BGS19] Brian Brubach, Nathaniel Grammel, and Aravind Srinivasan. Vertex-weighted online stochastic matching with patience constraints. arXiv preprint arXiv:1907.03963, 2019.

[BJN07] Niv Buchbinder, Kamal Jain, and Joseph Seffi Naor. Online primal-dual algorithms for maximizing ad-auctions revenue. In European Symposium on Algorithms, pages 253–264. Springer, 2007.

[BM08] Benjamin Birnbaum and Claire Mathieu. On-line bipartite matching made simple. ACM SIGACT News, 39(1):80–87, 2008.

[BSSX16] Brian Brubach, Karthik Abinav Sankararaman, Aravind Srinivasan, and Pan Xu. New algorithms, better bounds, and a novel model for online stochastic matching. In 24th Annual European Symposium on Algorithms (ESA 2016). Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2016.

[CF09] Carri W Chan and Vivek F Farias. Stochastic depletion problems: Effective myopic policies for a class of dynamic optimization problems. Mathematics of Operations Research, 34(2):333–350, 2009.

[CHN14] Moses Charikar, Monika Henzinger, and Huy L Nguyen. Online bipartite matching with decomposable weights. In European Symposium on Algorithms, pages 260–271. Springer, 2014.

[DH09] Nikhil R Devanur and Thomas P Hayes. The adwords problem: online keyword matching with budgeted bidders under random permutations. In Proceedings of the 10th ACM conference on Electronic commerce, pages 71–78. ACM, 2009.

[DJK13] Nikhil R Devanur, Kamal Jain, and Robert D Kleinberg. Randomized primal-dual analysis of ranking for online bipartite matching. In Proceedings of the twenty-fourth annual ACM-SIAM symposium on Discrete algorithms, pages 101–107. Society for Industrial and Applied Mathematics, 2013.

[FMMM09] Jon Feldman, Aranyak Mehta, Vahab Mirrokni, and S Muthukrishnan. Online stochastic matching: Beating 1-1/e. In Foundations of Computer Science, 2009. FOCS’09. 50th Annual IEEE Symposium on, pages 117–126. IEEE, 2009.

[GGI+19] Xiao-Yue Gong, Vineet Goyal, Garud Iyengar, David Simchi-Levi, Rajan Udwani, and Shuangyu Wang. Online assortment optimization with reusable resources. Available at SSRN 3334789, 2019.

[GM08] 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, pages 982–991. Society for Industrial and Applied Mathematics, 2008.

[GNR14] Negin Golrezaei, Hamid Nazerzadeh, and Paat Rusmevichientong. Real-time optimization of personalized assortments. Management Science, 60(6):1532–1551, 2014.
A Issue with Standard Dual Fitting for PBP

To make the dual simpler and more intuitive, let us simplify the formulation by introducing a change of variables, from $x_{it}^\omega$ to $x_{it}^{\omega_{t-1}}$. This change subsumes equalities (3), since the new variables incorporate their independence from future part of the sample path by definition. To understand how the other conditions can be modified, let us denote a collection of consistent partial sample paths up to time $t$ using $\omega^\tau$ for $\tau \leq t$. Now, consider the following constraints for every partial sample path $\omega^\tau$,

$$\sum_{\tau:j|\tau \leq t; (i,\tau) \in E} x_{it}^{\omega_{t-1}} 1^{\omega^\tau}(i,\tau) \leq 1 \quad \forall i \in I.$$
These constraints are directly implied by (1) and can in fact, be further strengthened. Fix an edge \((i, t)\) and notice that the space of all partial sample paths \(\omega^t\) that are consistent with \(\omega^{t-1}\) can be partitioned into two sets, ones which have \(1(i, t) = 1\) and ones with \(1(i, t) = 0\). Considering a sample path from the set of paths with \(1(i, t) = 1\), we have from the previous inequality that variables \(x_{it}^{\omega^{t-1}}\) satisfy,
\[
\sum_{\tau | \tau \leq t - 1; (i, \tau) \in E} x_{it}^{\omega^{t-1}} \cdot \mathbb{1}^{\omega^{t-1}}(i, \tau) + x_{it}^{\omega^{t-1}} \leq 1 \quad \forall \omega^{t-1} \in \Omega.
\]

We will replace constraints (1) by the above in our new LP. Before we can state the LP formally, it remains to simplify the objective. To that end, we define \(p(\omega, t - 1)\) as the probability of observing sample path \(\omega^{t-1}\) at arrival \(t\) i.e.,
\[
p(\omega, t - 1) = \mathbb{P}[\cap_{\tau \leq t - 1} \{1(i, \tau) = \mathbb{1}^{\omega}(i, \tau)\}].
\]

The objective value for clairvoyant can now be stated as follows,
\[
\sum_{(i, t) \in E} r_i E_\omega[x_{it}^{\omega^{t-1}} \cdot \mathbb{1}^{\omega}(i, t)] = \sum_{(i, t) \in E} r_i E_{\omega^{t-1}}[E_\omega[x_{it}^{\omega^{t-1}} \cdot \mathbb{1}^{\omega}(i, t) | \omega^{t-1} = \omega^{t-1}]]
\]
\[
= \sum_{(i, t) \in E} r_i \sum_{\omega^{t-1} \in \Omega} p(\omega, t - 1) \cdot x_{it}^{\omega^{t-1}} E_\omega[\mathbb{1}^{\omega}(i, t) | \omega^{t-1} = \omega^{t-1}]
\]
\[
= \sum_{\omega^{t-1} \in \Omega} \sum_{(i, t) \in E} r_i p(\omega, t - 1) \cdot p_{it} \cdot x_{it}^{\omega^{t-1}}.
\]

Where the first equality follows from the tower property of expectation, the second by definition of \(x_{it}^{\omega^{t-1}}\), and the third equality from the independence of each reward \(\mathbb{1}^{\omega}(i, t)\) from past rewards and an interchange of summation. Now, consider the new path based program (NewPBP) and its dual,

**NewPBP:**
\[
\max \sum_{\omega^{t-1} \in \Omega} \sum_{(i, t) \in E} p(\omega, t - 1) \cdot p_{it} r_i \cdot x_{it}^{\omega^{t-1}}
\]
\[
s.t. \sum_{\tau | \tau \leq t - 1; (i, \tau) \in E} x_{it}^{\omega^{t-1}} \cdot \mathbb{1}^{\omega^{t-1}}(i, \tau) + x_{it}^{\omega^{t-1}} \leq 1 \quad \forall (i, t) \in E, \omega^{t-1} \in \Omega
\]
\[
\sum_{i(t) \in E} x_{it}^{\omega^{t-1}} \leq 1 \quad \forall t \in T, \omega^{t-1} \in \Omega
\]
\[
0 \leq x_{it}^{\omega^{t-1}} \leq 1 \quad \forall (i, t) \in E, \omega \in \Omega.
\]

**Dual-NewPBP:**
\[
\min \sum_{t \in T, \omega^{t-1} \in \Omega} \lambda_t^{\omega^{t-1}} + \sum_{(i, t) \in E, \omega^{t-1} \in \Omega} \theta_{it}^{\omega^{t-1}}
\]
\[
s.t. \lambda_t^{\omega^{t-1}} + \theta_{it}^{\omega^{t-1}} + \mathbb{1}^{\omega}(i, t) \cdot \sum_{\tau | \tau \geq t; (i, \tau) \in E} \theta_{it}^{\omega^{t}} \geq p(\omega, t - 1) \cdot p_{it} r_i \quad \forall (i, t) \in E, \omega \in \Omega
\]
\[
\lambda_t^{\omega^{t-1}}, \theta_{it}^{\omega^{t-1}} \geq 0 \quad \forall t \in T, (i, t) \in E, \omega^{t-1} \in \Omega.
\]

It is not clear (to us), if one can always find a setting of the dual variables such that all dual constraints are satisfied within a large enough constant factor, while keeping the dual objective in check. In particular, consider the dual constraints for edge \((i, t)\) on a sample path where the online algorithm under consideration has successfully matched \(i\) in the past, say to some arrival \(t\). Further, suppose that on this sample path \(\mathbb{1}^{\omega}(i, t)\) is zero. To ensure \(\lambda_t^{\omega^{t-1}} + \theta_{it}^{\omega^{t-1}} \geq \gamma \cdot p(\omega, t - 1) \cdot p_{it} r_i\), for some constant \(\gamma\), on all such paths we have to set \(\lambda_t^{\omega^{t-1}}\) and/or \(\theta_{it}^{\omega^{t-1}}\) to non-negative values. Since these variables are not common across
arrivals, we would have to assign sufficiently large values to these variables for a large number of arrivals after \( t' \), if \( i \) has many edges after \( t' \). Therefore, ensuring dual feasibility can cause the dual objective to be much larger than the expected revenue of the online algorithm. Contrast this with the classical case where all probabilities equal one. Since there is a single sample path and \( 1_{\omega'}(i, t) = 1 \), we can simply set \( \theta_{iT} \), where we let \( T_i \) be the last arrival adjacent to \( i \), to a suitably large value. This variable then contributes to the dual constraints for all edges incident on \( i \).

One way to overcome this hurdle is to not satisfy every dual constraint but only some convex combination, such as in Lemma \[ \text{Lemma 1} \]. The resulting dual assignment is not necessarily feasible for the dual program above, unlike the dual fitting in \[ \text{[DJK13]} \], which is easily interpreted to maintain dual feasibility despite the randomization.