Secretary Markets with Local Information

Ning Chen[1], Martin Hoefer[2,⋆], Marvin Künnemann[2,3],
Chengyu Lin[4], and Peihan Miao[5]

[1] Nanyang Technological University, Singapore. ningc@ntu.edu.sg
[2] MPI für Informatik, Saarbrücken, Germany. {mhoefer,marvin}@mpi-inf.mpg.de
[3] Saarbrücken Graduate School of Computer Science, Germany.
[4] Chinese University of Hong Kong, Hong Kong. cylin@cse.cuhk.edu.hk
[5] University of California, Berkeley, USA. peihan@berkeley.edu

Abstract. The secretary model is a popular framework for the analysis of online admission problems beyond the worst case. In many markets, however, decisions about admission have to be made in a decentralized fashion and under competition. We cope with this problem and design algorithms for secretary markets with limited information. In our basic model, there are m firms and each has a job to offer. n applicants arrive iteratively in random order. Upon arrival of an applicant, a value for each job is revealed. Each firm decides whether or not to offer its job to the current applicant without knowing the strategies, actions, or values of other firms. Applicants decide to accept their most preferred offer.
We consider the social welfare of the matching and design a decentralized randomized thresholding-based algorithm with ratio $O(\log n)$ that works in a very general sampling model. It can even be used by firms hiring several applicants based on a local matroid. In contrast, even in the basic model we show a lower bound of $\Omega(\log n/(\log \log n))$ for all thresholding-based algorithms. Moreover, we provide secretary algorithms with constant competitive ratios, e.g., when values of applicants for different firms are stochastically independent. In this case, we can show a constant ratio even when each firm offers several different jobs, and even with respect to its individually optimal assignment. We also analyze several variants with stochastic correlation among applicant values.

1 Introduction

The Voice is a popular reality television singing competition to find new singing talent contested by aspiring singers. The competition employs a panel of coaches; upon the arrival of a singer, every coach critiques the artist's performance and determines in real time if he/she wants the artist to be on his/her team. Among those who express "I want you", the artist selects a favorite coach. What strategy of picking artists should coaches adopt in order to select the best candidates?

This problem is a reminiscent of the classic secretary problem: A firm interviews a set of candidates who arrive in an online fashion. When a candidate

⋆ Supported by DFG Cluster of Excellence MMCI at Saarland University.

arrives, its value is revealed and the firm needs to make an immediate and irrevocable decision on whether to make an offer to the candidate, without knowing the values of future potential candidates. The objective is to maximize the (expected) value of the hired candidate. The secretary problem is well studied in social science and computer science. It is well known that the problem, in the worst case, does not admit an algorithm with any guaranteed competitive ratio. However, if candidates arrive in uniform random order, there is an online algorithm that achieves the optimal competitive ratio $1/e$ [7,21]. For a more detailed discussion on the secretary problem see, e.g., [1].

The scenario of The Voice is a generalization of the secretary problem from one firm to multiple firms and from one hire to multiple hires. Such a generalization yields several fundamental changes to the problem: Firms (i.e., coaches) are independent and compete with each other for candidates. Thus, each firm may determine on their own the strategy to adopt. Firms are decision makers; that is, there is no centralized authority and every firm can choose different strategies on its own (based on observed information). Each firm can only observe information revealed to itself, i.e., it has no knowledge on the values of other (firm, candidate)-pairs and selected strategies of other firms. Hence, adopting a best-response strategy in a game-theoretic sense might require learning other strategies and payoffs. Given the limited feedback this can be hard or even impossible. The same issues occur in many other decentralized markets, e.g., online dating and school admission, where entities behave individually and have to make decisions based on a very limited view on the market, the preferences, and the strategies used by potential competitors.

The objective of the present paper is to design and analyze strategies for all firms in such a decentralized, competitive enviroment to enable efficient allocations. Our algorithms are evaluated both globally and individually: On the one hand, we hope the outcomes achieve good social welfare (i.e., the total value obtained by all firms). Thus, we measure the competitive ratio compared to social welfare given by the optimal centralized online algorithm. On the other hand, considering that firms are self-interested entities, we hope that our algorithms generate a nearly optimal outcome for each individual firm. That is, although given the limited feedback it can be impossible to obtain best-response strategies, we nevertheless hope that (when applied in combination) our algorithms can approximate the outcome of a best response (in hindsight with full information) of every individual firm within a small factor.

We identify several settings that admit algorithms with small constant competitive ratio both globally and individually. This implies that even in decentralized markets with very limited feedback, there are algorithms to obtain a good social solution. For the general case, we provide a strategy to approximate social welfare within a logarithmic factor, and we show almost matching lower bounds on the competitive ratio for a very natural class of algorithms. Thus, in the general case centralized control seems to be necessary in order to achieve good social welfare.

Model and Preliminaries We first outline our *basic model*, a decentralized online scenario for hiring a single applicant per firm with random arrival. There is a complete bipartite graph $G = (U, V, w)$ with sets $U = \{u_1, u_2, \ldots, u_m\}$ and $V = \{v_1, v_2, \ldots, v_n\}$ of firms and applicants, respectively. There is a *value* or *weight function*[6] $w : U \times V \to \mathbb{R}^+$. We assume that each firm can hire at most one applicant and there are more applicants than firms, i.e., $m \leq n$.

The weights describe an implicit preference of each individual to the other side. Each firm $u \in U$ prefers applicants according to the decreasing order of $w(u, \cdot)$ of the edges incident to u; similarly, each applicant $v \in V$ prefers firms according to the decreasing order of $w(\cdot, v)$ of the edges incident to v.[7]

Applicants in V arrive one by one in the market and reveal their edge weights to all firms. Upon the arrival of an applicant, each firm decides on whether to provide an offer for the applicant or not immediately; after collecting all job offers, the applicant then picks one that she prefers most, i.e., the one with the largest weight. Note that each firm can only see its own weights for the applicants and has no information about future applicants; in addition, all decisions cannot be revoked. In this paper, we consider the problem in the random permutation model, i.e., applicants arrive in a uniformly random order.

Our goal is to design decentralized algorithms when each firm makes its decision only based on its own previous information and there is no centralized authority that manages different firms altogether. There are two natural objectives to evaluate the performance of an algorithm, and due to online arrival some performance loss is unavoidable. The standard benchmark is *social welfare*, defined to be the total weight of assigned firm and applicant pairs. For an algorithm $\mathcal{A}$, we say the algorithm has a *competitive ratio* of α if for all instances, we have $\mathbb{E}[w(M^*)] / \mathbb{E}\left[w(M^{\mathcal{A}})\right] \leq \alpha$. Here the expectation is over random permutation, M^* is the maximum weight matching in G, and $M^{\mathcal{A}}$ is the matching returned when every firm runs algorithm $\mathcal{A}$. In addition, we would like to approximate the individual optimum assignment for each firm (i.e., the weight of its best candidate) and strive to obtain a constant competitive ratio for this benchmark.

Contribution and Techniques As a natural first attempt, consider every firm running the classic secretary algorithm [7, 21], which samples the first $r - 1$ applicants, records the best weight seen in the sample, and then offers to every applicant that exceeds this threshold. It turns out that such a strategy fails miserably in a decentralized market, even if each applicant has the same weight for all firms. For spatial reasons, the proof of this statement and many other formal arguments in this extended abstract are deferred to a full version.

Proposition 1. *For any constant $\beta < 1$, when setting $r = \lfloor \beta n \rfloor + 1$, then the classic secretary algorithm has a competitive ratio of $\Omega(n/\log n)$.*

---

[6] To avoid ties, we assume that no two edges have the same weight; this assumption is without loss of generality by using small perturbations or a fixed rule to break ties.

[7] In a more general preference model there are for each pair (u, v) different values obtained by u and v; we will not consider this general case in the present paper.

In contrast, we present in Section 2 a more careful approach based on sampling and thresholds that is $O(\log n)$-competitive. This algorithm can be applied in large generality (well beyond the basic model). In fact, we prove the guarantee in a scenario, where each firm u_i has a private matroid $\mathcal{S}_i$ and can accept any subset of applicants that forms an independent set in $\mathcal{S}_i$. Furthermore, our analysis extends to a general sampling model that encompasses the secretary model (random arrival, worst-case weights), prophet-inequality model (worst-case arrival, stochastic weights), as well as a variety of other mixtures of stochastic and worst-case assumptions [10]. Our main technique to handle decentralized thresholding is to bundle all stochastic decisions and treat correlations using linearity of expectation. The effects of applicant preferences and competition can then be analyzed in a pointwise fashion.

We contrast this result with an almost matching lower bound for thresholding-based algorithms in the basic model. A thresholding-based algorithm samples a number of applicants, determines a threshold, and then offers to every remaining applicant that has a weight above the threshold. Although such algorithms are nearly optimal in the centralized setting, every such algorithm must have a competitive ratio of at least $\Omega(\log n / \log \log n)$ in the decentralized setting. The lower bound carefully constructs a challenge to guess how many firms contribute to social welfare and to avoid overly high concentration of offers on a small number of valuable applicants.

In Section 3 we show that this challenge can be overcome if there is stochastic independence between the weights of an applicant to different firms. We study this property in a generalized model for decentralized k-secretary, where each firm u_i has k_i different jobs to offer. Upon arrival, an applicant reveals k_i weights for each firm u_i, one for each position. If each firm uses a variant of the optimal e-competitive algorithm for bipartite matching [16], independence between weights of different firms allows to show a constant competitive ratio. Moreover, each firm even manages to recover a constant fraction of the individual optimum matching and therefore almost plays a best response strategy.

Finally, in Section 4 we consider two additional variants with stochastically generated weights. In both variants we can show constant competitive ratios, and in one case firms can even hire their best applicant with constant probability.

Related Work The secretary model is a classic approach to stopping problems and online admission [7]. The classic algorithm outlined in the previous section is e-competitive, which is the best possible ratio. In the algorithmic literature, recent work has addressed secretary models for packing problems with random arrival of elements. A prominent case is the matroid secretary problem [2], for which the first general algorithm was $O(\log k)$-competitive, where k is the rank of the matroid. The ratio was very recently reduced to $O(\log \log k)$ [8,20]. Constant-factor competitive algorithms have been obtained for numerous special cases [6, 11, 14, 18, 24]. It remains a fascinating open problem if a general constant-factor competitive algorithm exists.

Another popular domain is bipartite matching in the secretary model, which has many applications in online revenue maximization via ad-auctions. In Section 3 we use a variant of a recent optimal e-competitive algorithm [16], which tightened the ratio and improved it over previous algorithms [2,5,19]. The main idea has recently been extended to construct optimal secretary algorithms for packing linear problems [17], improving over previous approaches [4,23]. Algorithms based on primal-dual techniques are a popular approach, especially for budgeted online matching with different stochastic input assumptions [3,15,22].

Our analysis of the algorithm for the general case applies in a unifying sampling model recently proposed as a framework for online maximum independent set in graphs [10]. It encompasses many stochastic adversarial models for online optimization – the secretary model, the prophet inequality model, and various other mixtures of stochastic and worst-case adversaries.

Closer to our paper are studies of a secretary problem with k queues [9], or game-theoretic approaches with complete knowledge of opponent strategies [12, 13]. These scenarios, however, have significantly different assumptions on the firms and their feedback, and they do not target markets with both decentralized control and restricted feedback that we explore in this paper.

2 General Preferences

For general weights $w : U \times V \to \mathbb{R}^+$, Proposition 1 shows that the classic secretary algorithm may perform poorly in a decentralized market. A reasonable strategy has to be more careful in adopting a threshold to avoid extensive competition for few candidates. Inspired by Babaioff et al. [2], we overcome this obstacle with a randomized thresholding strategy, and analyze it in a very general distributed matroid scenario. We remark that our bounds apply even within a general sampling model [10] that encompasses the secretary model, prophet-inequality model, and many other approaches for stochastic online optimization.

For the combinatorial structure of the scenario, we consider that each firm u_i holds a possibly different matroid $\mathcal{S}_i$ over the set of applicants. Firm u_i can accept an applicant as long as the set of accepted applicants forms an independent set in $\mathcal{S}_i$. Special cases include hiring a single applicant or any subset of at most k_i many applicants. Each firm strives to maximize the sum of weights of hired applicants. The structure of $\mathcal{S}_i$ does not have to be known in advance. u_i only needs an oracle to test if a set of arrived applicants is an independent set in $\mathcal{S}_i$.

Algorithm 1 is executed in parallel by all firms u_i. We first sample a fraction of roughly $n/(c+1)$ applicants, where $c \geq 1$ is a global constant. Then we determine a random threshold based on the maximum weight seen by firm u_i in its sample. Firm u_i then greedily makes an offer only to those candidates whose values are above the threshold.

Theorem 1. *Algorithm 1 is $O(\log n)$-competitive.*

Proof. We denote by V_i^S the set of candidates in the sample and by V_i^I the other candidates. Note that by the choice of sample and the random arrival, we have that $\mathbf{Pr}\left[v_j \in V_i^S\right] = \frac{1}{c+1}$. More broadly, our subsequent arguments will require

Algorithm 1: Thresholding algorithm for u_i with matroids.

Draw a random number $k \sim Binom(n, 1/(c+1))$
Reject the first k applicants, denote this set by V_i^S
$m_i \leftarrow \arg\max_{v_j \in V_i^S} w(u_i, v_j)$
$X_i \leftarrow \text{Uniform}(-1, 0, 1, \ldots, \lceil \log n \rceil)$
$t_i \leftarrow w(u_i, m_i)/2^{X_i}$, $M_i \leftarrow \emptyset$
for *all remaining* v_j *over time* **do**
  if $w(u_i, v_j) \geq t_i$ *and* $M_i \cup \{v_j\}$ *is independent set in* $\mathcal{S}_i$ **then**
    make an offer to v_j
    if v_j *accepts* **then** $M_i \leftarrow M_i \cup \{v_j\}$

only the weaker bounds

$$\mathbf{Pr}\left[v_j \in V_i^I\right] \geq \frac{1}{c+1} \qquad \text{and} \qquad \mathbf{Pr}\left[v_j \in V_i^S\right] \geq \frac{1}{c+1} \ . \tag{1}$$

These sampling inequalities obviously hold for every v_j, independently of $v_{j'} \in V_i^S$ or not for all other candidates $j' \neq j$.

Let $v_i^{\max} = \text{argmax}_j w(u_i, v_j)$ and $v_i^{\text{2nd}} = \text{argmax}_{j \neq v_i^{\max}} w(u_i, v_j)$ be a best and second best applicant for firm u_i, respectively (breaking ties arbitrarily). In addition, we denote by $w_i^{\max} = w(u_i, v_i^{\max})$ and $w_i^{\text{2nd}} = w(u_i, v_i^{\text{2nd}})$ their weights for firm u_i. For most of the analysis, we consider another weight function, the *capped weights* $\tilde{w}(u_i, v_j)$, based on thresholds t_i set by the algorithm as follows

$$\tilde{w}(u_i, v_j) = \begin{cases} w_i^{\max} & \text{if } v_j \in V_i^I, t_i = 2w_i^{\text{2nd}}, \text{ and } w(u_i, v_j) > 2w_i^{\text{2nd}}, \\ t_i & \text{if } v_j \in V_i^I \text{ and } w(u_i, v_j) \geq t_i, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that the definition of $\tilde{w}$ relies on several random events, i.e., $v_j \in V_i^I$ and the choice of thresholds t_i. For any outcome of these events, however, we have that $\tilde{w}(u_i, v_j) \leq w(u_i, v_j)$ for all pairs (u_i, v_j), since if $t_i = 2w_i^{\text{2nd}}$ and $w(u_i, v_j) > 2w_i^{\text{2nd}}$, then $v_j = v_i^{\max}$. By the following lemma, in expectation over all the correlated random events, an optimal offline solution with respect to $\tilde{w}$ still gives an approximation to the optimal offline solution with respect to w.

Lemma 1. *Denote by $w(M)$ and $\tilde{w}(M)$ the weight and capped weight of a solution M. Let $\tilde{M}^*$ and M^* be optimal solutions for $\tilde{w}$ and w, respectively. Then*

$$\mathbb{E}\left[\tilde{w}(\tilde{M}^*)\right] \geq \Omega\left(\frac{1}{\log n}\right) \cdot w(M^*).$$

Proof. Let $(u_i, v_j) \in M^*$ be an arbitrary pair. First, assume that v_j maximizes $w(u_i, v_j)$, i.e., $v_j = v_i^{\max}$. By (1) with probability at least $1/(c+1)^2$, we have $v_j \in V_i^I$ and $v_i^{\text{2nd}} \in V_i^S$. For any such outcome, we have with probability $1/(\lceil \log(n) \rceil + 2)$ that either (1) $t_i = 2w_i^{\text{2nd}}$ and $\tilde{w}(u_i, v_j) = w_i^{\max}$ (if $w_i^{\max} \geq 2w_i^{\text{2nd}}$), or

(2) $t_i = w_i^{\text{2nd}}$ and $t_i \le w_i^{\max} < 2w_i^{\text{2nd}}$ (otherwise). This yields $\mathbb{E}\left[\tilde{w}(u_i, v_j)\right] \ge w(u_i, v_j)/(2(c+1)^2(\lceil\log(n)\rceil + 2))$.

Second, for any $v_j \ne v_i^{\max}$ with $w_i^{\max}/(2n) < w(u_i, v_j) \le w_i^{\max}$, by (1) we know $v_i^{\max} \in V_i^S$ is an independent event which happens with probability at least $1/(c+1)$. Then, there is some $0 \le k' \le \lceil\log n\rceil + 1$, with $w(u_i, v_j) > w_i^{\max}/2^{k'} \ge w(u_i, v_j)/2$. With probability $1/(\lceil\log(n)\rceil+2)$, we have that $X_i = k'$ and $\tilde{w}(u_i, v_j) = t_i \ge w(u_i, v_j)/2$. This yields $\mathbb{E}\left[\tilde{w}(u_i, v_j)\right] \ge w(u_i, v_j)/(2(c+1)^2(\lceil\log(n)\rceil + 2))$, since $v_j \in V_i^I$ with probability at least $1/(c+1)$ by (1).

Finally, we denote by $M^>$ the set of pairs $(u_i, v_j) \in M^*$ for which $w(u_i, v_j) > w_i^{\max}/(2n)$. The expected weight of the best assignment with respect to the threshold values is thus

$$\begin{aligned}
\mathbb{E}\left[\tilde{w}(\tilde{M}^*)\right] &\ge \sum_{(u_i,v_j)\in M^*} \mathbb{E}\left[\tilde{w}(u_i, v_j)\right] \ge \sum_{(u_i,v_j)\in M^>} \frac{w(u_i, v_j)}{2(c+1)^2(\lceil\log(n)\rceil + 2)} \\
&= \frac{1}{2(c+1)^2(\lceil\log(n)\rceil + 2)} \cdot (w(M^*) - w(M^* \setminus M^>)) \\
&\ge \frac{1}{4(c+1)^2(\lceil\log(n)\rceil + 2)} \cdot w(M^*),
\end{aligned}$$

since $\sum_{(u_i,v_j)\in M^*\setminus M^>} w_i^{\max}/(2n) \le \max_i w_i^{\max}/2 \le w(M^*)/2$. □

The previous lemma bounds the weight loss due to (i) all random choices inherent in the process of input generation and threshold selection and (ii) using the capped weights. The next lemma bounds the remaining loss due to adversarial arrival of elements in V_i^I, exploiting that $\tilde{w}$ equalizes equal-threshold firms.

Lemma 2. *Suppose subsets V_i^I and thresholds t_i are fixed arbitrarily and consider the resulting weight function $\tilde{w}$. Let $M^{\mathcal{A}}$ be the feasible solution resulting from Algorithm 1 using the thresholds t_i, for any arbitrary arrival order of applicants in $\bigcup V_i^I$. Then $w(M^{\mathcal{A}}) \ge \tilde{w}(\tilde{M}^*)/2$.*

Combining the insights, we see that that $w(M^*) \le O(\log n) \cdot \mathbb{E}\left[w(M^{\mathcal{A}})\right]$, which proves the theorem. □

Our general upper bound results from a thresholding-based algorithm. We constrast this result with a lower bound for thresholding-based algorithms when every firm wants to hire only a single applicant. It applies even when preferences of all firms over applicants are identical. More formally, an algorithm $\mathcal{A}$ is called *thresholding-based* if during its execution $\mathcal{A}$ rejects applicants for some number of rounds, then determines a threshold T and afterwards enters an *acceptance phase*. In the acceptance phase, it makes an offer to exactly those applicants whose weight exceeds threshold T. Note that the number of rejecting rounds in the beginning and the threshold T can be chosen arbitrarily at random.

The lower bound uses an *identical-firm* instance in which for each applicant v_j all firms have the same weight, i.e., there is $w(v_j) \ge 0$ such that $w(u_i, v_j) = w(v_j)$ for every firm u_i. It applies in the secretary model and the iid model. In the latter we draw the weight $w(v_j)$ for each v_j independently at random from a single distribution. The main difference is that M^* becomes a random variable.

Theorem 2. *Suppose every firm strives to hire a single applicant, and let $\mathcal{A}$ be any thresholding-based algorithm. If every firm adopts $\mathcal{A}$, there is an identical-firm instance I on which $\mathcal{A}$ has a competitive ratio of $\Omega(\log n/\log\log n)$. This lower bound applies in the iid model and the secretary model.*

Proof (Idea). For simplicity, we assume[8]that the thresholding-based algorithm $\mathcal{A}$ does not know the number of firms m. Assume that $n = \sum_{j=2}^{t} t^{2j}$ for some $t \in \mathbb{N}$. We construct a distribution $\mathcal{I}$ on a family of identical-firm instances by drawing the weight $w(v_{j'})$ of each applicant $v_{j'}$ according to $\Pr[w(v_{j'}) = t^{-j}] = t^{2j}/n$ for $j = 2, \ldots, t$. In the secretary model, we may assume that each applicant draws $w(v_{j'})$ at the moment it arrives in the random order, since the order is chosen independently of the weights. Since all applicant weights are identically distributed, we may even completely disregard the random arrival order.

We define classes $C_2, \ldots, C_t$, where each class C_j consists of all applicants $v_{j'}$ with value $w(v_{j'}) = t^{-j}$. Consider how $\mathcal{A}$ performs on $\mathcal{I}$ for some firm u_i. We can assume that $\mathcal{A}$ chooses a threshold among $\{t^{-2}, \ldots, t^{-t}\}$, since all other choices are equivalent concerning the set of applicants receiving an offer from u_i. Let p_j be the probability (over $\mathcal{I}$ and the random choices of $\mathcal{A}$) that threshold t^{-j} is chosen. Clearly, there is some $2 \le k \le t$ with $p_k \le \frac{1}{t-1}$. By setting $m := \sum_{j=2}^{k} t^{2j}$, most firms should choose a threshold of t^{-k} to obtain a competetive solution, but by choice of k few firms do. Hence, the challenge for the firms is to guess m correctly and extract welfare from the right class of applicants. □

3 Independent Preferences

In this section, we show improved results for decentralized matching in the secretary model when preferences are independent among firms. More formally, we assume firm u_i has a set U_i of k_i positions available. An adversary specifies a separate set $\mathcal{P}_i$ of n *applicant profiles* for each firm u_i. An applicant profile $p \in \mathcal{P}_i$ is a function $p : U_i \to \mathbb{R}^+$. In round t, when a new applicant v_t arrives, we pick one remaining profile $p_{it} \in \mathcal{P}_i$ for each $u_i \in U$ independently and uniformly at random. The weight for position $u_{ij} \in U_i$ is then given by $w(u_{ij}, v_t) = p_{it}(u_{ij})$. We pick profiles from $\mathcal{P}_i$ uniformly at random without replacement. Special cases of this model are, e.g., when all weights for all positions are independently sampled from a certain distribution, or for each firm u_i the weights of all applicants are sampled independently from a different distribution for each position.

In contrast to the previous section, we assume that each applicant has k_i weight values for each firm u_i. A straightforward $O(\log n)$-competitive algorithm is to run Algorithm 1 separately for each position of each firm. In contrast, when $n \ge \sum_{i=1}^{m} k_i$ and $k_i \le \alpha n$ for all $i \in [m]$ and some constant $\alpha \in (0, 1)$, we can achieve a constant competitive ratio using Algorithm 2. This algorithm resembles an optimal algorithm for secretary matching with a single firm [16]. Each firm rejects a number of applicants and enters an acceptance phase. In this phase, it maintains two virtual solutions: (1) an individual virtual optimum $M^*_{i,t}$ with

[8] This assumption can be dropped by introducing firms with neglibly small preferences.

Algorithm 2: Matching algorithm for firm u_i for independent weights

Reject the first $r_i - 1$ applicants
$M_i, M'_i \leftarrow \emptyset$
for *applicant* v_t *arriving in round* $t = r_i, \ldots, n$ **do**
  Let $M^*_{i,t}$ be optimum matching for firm u_i and applicants $\{v_1, \ldots, v_t\}$
  if v_t *is matched to position* u_{ij} *in* M^*_i *and* u_{ij} *unmatched in* M'_i **then**
    Make an offer for position u_{ij} to v_t
    $M'_i \leftarrow M'_i \cup \{(u_{ij}, v_t)\}$
    if v_t *accepts* **then**
      $M_i \leftarrow M_i \cup \{(u_{ij}, v_t)\}$

respect to applicants arrived up to and including round t, and (2) a virtual solution M'_i where all applicants are assumed to accept the offers of u_i. If the newly arrived applicant v_t is matched in $M^*_{i,t}$, it is offered the same position unless this position is already filled in M'_i.

Theorem 3. *Algorithm 2 achieves a constant competitive ratio.*

Proof. Fix a firm u_i. The matching M'_i is constructed by assuming that u_i is the only firm in the market, i.e., every applicant accepts the offer of firm u_i. Consider the individual optimum $M^*_{i,n}$ in hindsight. Then, by repeating the analysis of [16], the expected value of M'_i is

$$\mathbb{E}\left[w(M'_i)\right] \geq \frac{r_i - 1}{n} \ln\left(\frac{n}{r_i - 1}\right) \cdot w(M^*_{i,n}) = f(r_i) \cdot w(M^*_{i,n}) \ ,$$

where we denote the ratio by $f(r_i)$. Recall $k_i \leq \alpha n$. Set r_i in the interval $(\beta n, \gamma n)$ for some appropriate constants $\beta, \gamma \in (0,1)$ such that $\beta > \alpha$. This ensures that $f(r_i)$ becomes a constant.

Let us now analyze the performance of the algorithm in the presence of competition. Consider applicant v_t in round t and the following events: (1) $P(u_i, v_t)$ is the event that u_i sends an offer to v_t, and (2) $A(u_i, v_t)$ is the event that u_i sends an offer to v_t and he accepts it. u_i's decision to offer depends only on $M^*_{i,t}$ and M'_i, but not on the acceptance decisions of earlier applicants. v_t for sure accepts an offer from u_i if u_i offers and no other firm offers. Offers from other firms $u_{i'}$ occur only if $u_{i'}$ is matched in $M^*_{i',t}$. More formally, $A(u_i, v_t)$ occurs (at least) if $P(u_i, v_t)$ and none of the $P(u_{i'}, v_t)$ occur. Since the profiles for different firms are combined independently

$$\mathbf{Pr}\left[A(u_i, v_t) \mid P(u_i, v_t)\right] \geq \prod_{i \neq i'} (1 - \mathbf{Pr}\left[P(u_{i'}, v_t)\right])$$

Consider the probability that v_t is matched in $M^*_{i',t}$. Since the order of profiles for $u_{i'}$ is independent of the order for u_i, we can imagine again choosing t

profiles at random. Of those a random profile is chosen to be one of v_t. The t profiles determine $M^*_{i',t}$, which matches $\min(t, k_{i'})$ profiles. Since the last profile is determined at random, the probability that v_t is matched in $M^*_{i',t}$ is at most $\min(1, k_{i'}/t)$. As $t \geq r_{i'} \geq \beta n$, we have

$$\mathbf{Pr}\left[P(u_{i'}, v_t)\right] \leq \begin{cases} 0 & \text{if } t \leq r_{i'} - 1, \\ k_{i'}/(\beta n) & \text{otherwise.} \end{cases}$$

Thus,

$$\begin{aligned} \mathbf{Pr}\left[A(u_i, v_t) \mid P(u_i, v_t)\right] &\geq \prod_{i \neq i'} (1 - \mathbf{Pr}\left[P(u_{i'}, v_t)\right]) \geq \exp\left(\sum_{i=1}^{m} \ln\left(1 - \frac{k_i}{\beta n}\right)\right) \\ &\geq \exp\left(-\sum_{i=1}^{m} \frac{1}{1-(\alpha/\beta)} \cdot \frac{k_i}{\beta n}\right) \geq \exp\left(-\frac{1}{\beta - \alpha}\right). \end{aligned}$$

The third inequality follows from $k_i \leq \alpha n$ by $(1 - k_i/(\beta n)) \geq 1 - \alpha/\beta$. Furthermore, it holds that $\ln(1-x) \geq -\frac{x}{1-x}$ for all $x \in (0,1)$. The last inequality is due to $n \geq \sum_j k_j$.

Consequently, $\mathbb{E}\left[w(M_i)\right]$ recovers at least a constant fraction of $\mathbb{E}\left[w(M_i')\right]$, which represents a constant factor approximation to the individual optimum $M^*_{i,n}$ for i in hindsight. By linearity of expectation, the algorithm achieves a constant competitive ratio for the expected weight of the optimum matching.

□

4 Correlated Preferences

In this section, we treat the basic model where every firm strives to hire one applicant. We consider stochastic input generation which allows correlations on the weights incident to an applicant. Specifically, assume that each applicant v_i has a parameter q_i, measuring his built-in quality, and the weights of edges incident to v_i are generated independently from a distribution D_i with mean q_i. Note that the lower bound for the classical e-competitive algorithm for the secretary problem (Proposition 1) applies to this general setting. As a natural candidate, we consider in particular normal distributions and assume that $D_i \sim N(q_i, \sigma^2)$ where q_i is the quality of applicant v_i and σ is a fixed constant.

We analyze correlations in two regimes: When the random noise is small and the preference lists of each firm are unlikely to differ significantly and when large variance has substantial effects on the preferences.

Small Variance We consider the case of highly correlated preferences of an applicant to all firms with possibly small fluctuations around an applicant's quality. Consider the list-based approach of Algorithm 3 that first samples a linear number $r = \Theta(n)$ of applicants and afterwards maintains a list of the top m candidates observed so far. The key observation we exploit is that Algorithm 3,

Algorithm 3: List-based algorithm for firm u

```
Initialize list L_u = (ℓ_{u,1}, ..., ℓ_{u,m}), initialized with (−∞, ..., −∞)
(maintain L_u to contain the top m weights u observed so far, where
ℓ_{u,1} ≥ ··· ≥ ℓ_{u,m})
Reject the first (r − 1) applicants, denote the set by R
for applicant v_t arriving in round t = r, ..., n do
    if w(u, v_t) ≥ ℓ_{u,m} then
        Update L_u: Push w_{u,v_t} into L_u and pop ℓ_{u,m} out.
        if popped out ℓ_{u,m} = −∞ or corresponds to an applicant in R then
            Make an offer to v_t, stop if v_t accepts
```

in contrast to the classical algorithm for the secretary problem, can cope well with competition, provided that applicants have a global quality that all firms roughly agree on. In particular, each of the top m applicants will be matched to her best firm with constant probability.

Without loss of generality, let $q_1 \geq \cdots \geq q_n$. Formally, define the parameter $\delta_{min} := \min_{i \neq j} |q_i - q_j|$, and $\psi = \frac{\delta_{min}}{\sigma}$.

Theorem 4. *Let $\psi = \omega(n)$ and $r = \Theta(n)$. Algorithm 3 achieves a constant competitive ratio with high probability, i.e., with probability approaching 1 over all possible weights, we have $\mathbb{E}[w(M^*)] \leq c \cdot \mathbb{E}\left[w(M^{\mathcal{A}})\right]$ for some constant $c > 0$.*

Large Variance If weights are perturbed by high-variance normal distributions, this yields a natural situation in which the classic algorithm for the secretary problem achieves a constant competitive ratio. Let $\delta_{max} := \max_{i \neq j} |q_i - q_j|$ denote the largest difference in applicants' qualifications and define the parameter $\varphi := \frac{\delta_{max}}{\sigma}$.

Theorem 5. *The classic secretary algorithm achieves a constant competitive ratio when $\varphi = O(\frac{1}{n^2})$ and $r = \Theta(n)$. Each firm hires its best applicant with constant probability.*

References

1. M. Babaioff, N. Immorlica, D. Kempe, and R. Kleinberg. Online auctions and generalized secretary problems. *SIGecom Exchanges*, 7(2), 2008.
2. M. Babaioff, N. Immorlica, and R. Kleinberg. Matroids, secretary problems, and online mechanisms. In *Proc. 18th Symp. Discrete Algorithms (SODA)*, pages 434–443, 2007.
3. N. Devanur and T. Hayes. The adwords problem: Online keyword matching with budgeted bidders under random permutations. In *Proc. 10th Conf. Electronic Commerce (EC)*, pages 71–78, 2009.
4. N. Devanur, K. Jain, B. Sivan, and C. Wilkens. Near optimal online algorithms and fast approximation algorithms for resource allocation problems. In *Proc. 12th Conf. Electronic Commerce (EC)*, pages 29–38, 2011.

5. N. Dimitrov and G. Plaxton. Competitive weighted matching in transversal matroids. *Algorithmica*, 62(1-2):333–348, 2012.
6. M. Dinitz and G. Kortsarz. Matroid secretary for regular and decomposable matroids. *SIAM J. Comput.*, 43(5):1807–1830, 2014.
7. E. Dynkin. The optimum choice of the instant for stopping a Markov process. In *Sov. Math. Dokl*, volume 4, pages 627–629, 1963.
8. M. Feldman, O. Svensson, and R. Zenklusen. A simple $O(\log\log(\text{rank}))$-competitive algorithm for the matroid secretary problem. In *Proc. 26th Symp. Discrete Algorithms (SODA)*, pages 1189–1201, 2015.
9. M. Feldman and M. Tennenholtz. Interviewing secretaries in parallel. In *Proc. 13th Conf. Electronic Commerce (EC)*, pages 550–567, 2012.
10. O. Göbel, M. Hoefer, T. Kesselheim, T. Schleiden, and B. Vöcking. Online independent set beyond the worst-case: Secretaries, prophets and periods. In *Proc. 41st Intl. Coll. Automata, Languages and Programming (ICALP)*, volume 2, pages 508–519, 2014.
11. S. Im and Y. Wang. Secretary problems: Laminar matroid and interval scheduling. In *Proc. 22nd Symp. Discrete Algorithms (SODA)*, pages 1265–1274, 2011.
12. N. Immorlica, A. Kalai, B. Lucier, A. Moitra, A. Postlewaite, and M. Tennenholtz. Dueling algorithms. In *Proc. 43rd Symp. Theory of Computing (STOC)*, pages 215–224, 2011.
13. N. Immorlica, R. Kleinberg, and M. Mahdian. Secretary problems with competing employers. In *Proc. 2nd Intl. Workshop Internet & Network Economics (WINE)*, pages 389–400, 2006.
14. P. Jaillet, J. Soto, and R. Zenklusen. Advances on matroid secretary problems: Free order model and laminar case. In *Proc. 16th Intl. Conf. Integer Programming and Combinatorial Optimization (IPCO)*, pages 254–265, 2013.
15. C. Karande, A. Mehta, and P. Tripathi. Online bipartite matching with unknown distributions. In *Proc. 43rd Symp. Theory of Computing (STOC)*, pages 587–596, 2011.
16. T. Kesselheim, K. Radke, A. Tönnis, and B. Vöcking. An optimal online algorithm for weighted bipartite matching and extensions to combinatorial auctions. In *Proc. 21st European Symp. Algorithms (ESA)*, pages 589–600, 2013.
17. T. Kesselheim, K. Radke, A. Tönnis, and B. Vöcking. Primal beats dual on online packing LPs in the random-order model. In *Proc. 46th Symp. Theory of Computing (STOC)*, pages 303–312, 2014.
18. R. Kleinberg. A multiple-choice secretary algorithm with applications to online auctions. In *Proc. 16th Symp. Discrete Algorithms (SODA)*, pages 630–631, 2005.
19. N. Korula and M. Pál. Algorithms for secretary problems on graphs and hypergraphs. In *Proc. 36th Intl. Coll. Automata, Languages and Programming (ICALP)*, pages 508–520, 2009.
20. O. Lachish. O(log log rank) competitive ratio for the matroid secretary problem. In *Proc. 55th Symp. Foundations of Computer Science (FOCS)*, pages 326–335, 2014.
21. D. Lindley. Dynamic programming and decision theory. *Applied Statistics*, 10:39–51, 1961.
22. A. Mehta, A. Saberi, U. Vazirani, and V. Vazirani. Adwords and generalized online matching. *J. ACM*, 54(5), 2007.
23. M. Molinaro and R. Ravi. Geometry of online packing linear programs. *Math. Oper. Res.*, 39(1):46–59, 2014.
24. J. Soto. Matroid secretary problem in the random-assignment model. *SIAM J. Comput.*, 42(1):178–211, 2013.

A Useful Facts

Fact 1 *For all $x \in [0, 1)$ it holds that*

$$\ln(1-x) \quad \geq \quad -\frac{x}{1-x} \ .$$

Proof. For $x = 0$ we have equality. The derivative of left- and right-hand sides are $-(1-x)^{-1}$ and $-(1-x)^{-2}$, respectively. Hence, the right-hand side drops faster when $x > 0$ grows towards 1, so the inequality holds for the entire interval. □

B Extension to the Sampling Model

For the sampling model, every time applicant v_j arrives, we assume it reveals a value $w^I(u_i, v_j)$ for each firm u_i. Applicants arrive in adversarial order. The values $w^I(u_i, v_j)$ are drawn from an unknown joint distribution. In addition, a priori each firm u_i is provided with a sample value $w^S(u_i, v_j)$ for each applicant v_j. These sample values are drawn from a (possibly different) unknown joint distribution. For a single v_j, the two distributions can be arbitrarily correlated among different firms and among each other. However, there is no correlation among different applicants in both distributions. The probability that w^I and w^S have value b for pair (u_i, v_j) is similar. We here restrict attention to discrete distributions over integers. It is straightforward that our results hold for general distributions, but this minor extension does not justify the notational and technical overhead in presentation. More formally, we assume

- *Stochastic similarity*: Suppose $c > 1$ is a fixed constant. For every pair (u_i, v_j) and every integer $b > 0$, we assume that $\mathbf{Pr}\left[w^I(u_i, v_j) = b\right] \leq c \cdot \mathbf{Pr}\left[w^S(u_i, v_j) = b\right]$ and $\mathbf{Pr}\left[w^S(u_i, v_j) = b\right] \leq c \cdot \mathbf{Pr}\left[w^I(u_i, v_j) = b\right]$.
- *Stochastic independence*: For every pair (u_i, v_j), the weights $w^I(u_i, v_j)$ and $w^S(u_i, v_j)$ do not depend on the weights w^S and w^I of other candidates $v_{j'} \neq v_j$.

For further discussion of the sampling model and an exposition how to formulate secretary and prophet-inequality cases within this framework, see [**?**].

Algorithm 4 is executed in parallel by all firms u_i. We first simplify the structure of input and sample values by assuming that no candidate has $w^S(u_i, v_j) > 0$ and $w^I(u_i, v_j) > 0$. This loses a factor of at most 2 in the expected value of the solution. Then we determine a random threshold based on the maximum weight seen by firm u_i in its simplified sample. We greedily make an offer only to those candidates whose simplified input values are above the threshold.

Theorem 6. *Algorithm 4 is $O(\log n)$-competitive in the sampling model.*

Proof. The proof follows largely the one presented for the secretary model in the paper. We here only explain how to obtain the required sampling inequalities (1). This allows to apply the remaining argument essentially without modification.

Algorithm 4: Thresholding algorithm for firm u_i for general weights and matroids.

```
For each v_j flip a fair coin: if heads w^I(u_i, v_j) ← 0, if tails w^S(u_i, v_j) ← 0
m_i ← arg max_{v_j} w^S(u_i, v_j)
X_i ← Uniform(−1, 0, 1, ..., ⌈log n⌉)
t_i ← w^S(u_i, m_i)/2^{X_i}
M_i ← ∅
for all v_j over time do
    if w^I(u_i, v_j) ≥ t_i and M_i ∪ {v_j} is independent set in S_i then
        make an offer to v_j
        if v_j accepts then
            M_i ← M_i ∪ {v_j}
```

The first line of our algorithm implements an adjustment of weights, so that at most one of the two weights for a candidate and a firm is positive. Let us assume w.l.o.g. that this condition holds already for the initial weights w^I and w^S. This lowers the expected value of the optimum solution by at most a factor of 2. In addition, it preserves stochastic independence and similarity properties of the sampling model. More formally, we denote

$$\hat{w}(u_i, v_j) = \max\{w^I(u_i, v_j), w^S(u_i, v_j)\}$$

and assume that $(w^I(u_i, v_j), w^S(u_i, v_j)) \in \{(0, \hat{w}(u_i, v_j)), (\hat{w}(u_i, v_j), 0)\}$.

We condition on properties of the candidate with largest and second largest value for firm u_i. To cope with the resulting correlations, we introduce a conditional probability space. For each candidate v_j we assume that $\hat{w}(u_i, v_j)$ is fixed arbitrarily. For simplicity, we drop candidates from consideration for which $\hat{w}(u_i, v_j) = 0$. Let $V_i^I = \{v_j \mid w^I(u_i, v_j) > 0\}$ and $V_i^S = \{v_j \mid w^S(u_i, v_j) > 0\}$. Stochastic similarity implies

$$\mathbf{Pr}\left[w^I(u_i, v_j) = \hat{w}(u_i, v_j)\right] \geq (1/c) \cdot \mathbf{Pr}\left[w^S(u_i, v_j) = \hat{w}(u_i, v_j)\right]$$

and

$$\mathbf{Pr}\left[w^S(u_i, v_j) = \hat{w}(u_i, v_j)\right] \geq (1/c) \cdot \mathbf{Pr}\left[w^I(u_i, v_j) = \hat{w}(u_i, v_j)\right] \quad .$$

By the fact that $V_i^I \cap V_i^S = \emptyset$, we have

$$\mathbf{Pr}\left[v_j \in V_i^I\right] \geq \frac{1}{c+1} \qquad \text{and} \qquad \mathbf{Pr}\left[v_j \in V_i^S\right] \geq \frac{1}{c+1} \tag{2}$$

for each candidate v_j, independent of the outcome of weights of other candidates. The remaining arguments from the proof of Theorem 1 can be applied literally when using $\hat{w}(u_i, v_j)$ in place of $w(u_i, v_j)$. The adjustments outlined above only increase the resulting guarantee by a constant factor. □

C Proofs

C.1 Proof of Proposition 1

Proof. Let $m = \Theta(n)$. There are two types of applicants: $\alpha = \Theta(\log(n))$ 'good' applicants with weight 2 for all edges incident to them and the rest 'bad' applicants with weight 1 for all edges incident to them. (Note that to avoid ties, we can add a small perturbation $\epsilon_{u,v}$ on all pairs).

For any permutation of the applicants, we have $w(M^*) = m + \alpha$. Next, we consider the matching $M^{\mathcal{A}}$ returned by the algorithm and give an upper bound on $\mathbb{E}\left[w(M^{\mathcal{A}})\right]$. In the first $r-1$ rounds, the probability that no firm sees a 'good' applicant is

$$\prod_{i=0}^{\lfloor \beta n \rfloor} \frac{n-\alpha-i}{n-i} \leq \left(1-\frac{\alpha}{n}\right)^{\lfloor \beta n \rfloor} \leq \left(1-\frac{\alpha}{n}\right)^{\beta n - 1}$$

If some firm observes a 'good' applicant, no 'bad' applicant can be hired since the threshold for the firm is set to be 2. Hence,

$$\begin{aligned}\mathbb{E}\left[w(M^{\mathcal{A}})\right] &\leq \left(1-\frac{\alpha}{n}\right)^{\beta n-1}(m+\alpha) + \left(1-\prod_{i=0}^{\lfloor \beta n \rfloor}\frac{n-\alpha-i}{n-i}\right)\cdot 2\alpha \\ &\leq \frac{n}{n-\alpha}\cdot e^{-\alpha\beta}\cdot(m+\alpha)+2\alpha\end{aligned}$$

Therefore,

$$\frac{\mathbb{E}\left[w(M^{\mathcal{A}})\right]}{\mathbb{E}\left[w(M^*)\right]} \leq \frac{n}{n-\alpha}\cdot e^{-\alpha\beta} + \frac{2\alpha}{m+\alpha} = \frac{1}{n^{\Theta(1)}} + \Theta\left(\frac{\log n}{n}\right).$$

□

C.2 Proof of Lemma 2

Proof. We will account for the weight of each edge $(u_i, v_j) \in \tilde{M}^*$ under $\tilde{w}$ by the original weight $w(e)$ of a pair $e \in M^{\mathcal{A}}$, using each pair in $M^{\mathcal{A}}$ at most a constant number of times. Recall the definitions of $\hat{w}$ and w^I in Appendix B. Let $(u_i, v_j) \in \tilde{M}^*$ be arbitrary and w.l.o.g. assume $\tilde{w}(u_i, v_j) > 0$. This implies that $\hat{w}(u_i, v_j) \geq t_i$ and hence v_j will be assigned to u_i by Algorithm 4, except for the cases that either (1) v_j gets an offer from another firm $u_{i'} \neq u_i$ that offers larger weight, or (2) u_i has accepted others applicants before v_j arrives that forbid an offer to v_j.

Assume that case (1) holds and v_j is assigned to another firm $u_{i'}$ with a better offer, then $w^I(u_{i'}, v_j) > w^I(u_i, v_j) = \hat{w}(u_i, v_j) > t_i = \tilde{w}(u_i, v_j)$. Thus we can charge $\tilde{w}(u_i, v_j)$ to $w^I(u_{i'}, v_j)$. Otherwise, consider all remaining $(u_i, v_j) \in \tilde{M}^*$ for firm u_i, which are not yet accounted for, and denote their number by k_i. All these applicants v_j receive their best offer from u_i, along with possibly other applicants. From this pool – due to the exchange property of matroid $\mathcal{S}_i$ – firm u_i accepts in $M^{\mathcal{A}}$ at least k_i applicants, irrespective of the arrival order. Hence, in this step, for each of the k_i pairs $(u_i, v_j) \in M^{\mathcal{A}}$, there is at most one (different or the same) pair $(u_i, v_{j'}) \in \tilde{M}^*$ with the $\tilde{w}(u_i, v_j) = \tilde{w}(u_i, v_{j'}) = t_i$, to which we can charge it. Note that we can match these pairs such that no $(u_i, v_{j'}) \in \tilde{M}^*$ is charged more than once by some $(u_i, v_j) \in M^{\mathcal{A}}$.

Consider an arbitrary edge $(u_{i'}, v_{j'}) \in M^{\mathcal{A}}$. In the above accounting scheme, this edge can only be used to account for $\tilde{w}(u_i, v_j)$ with $(u_i, v_j) \in M^*$ if either $u_i = u_{i'}$ or $v_j = v_{j'}$. Since M^* is a feasible assignment, $(u_{i'}, v_{j'})$ can be used at most twice. This yields $\tilde{w}(M^*) \leq 2w(M^{\mathcal{A}})$. □

C.3 Proof of Theorem 2 (contd.)

Let us denote by T_i the threshold chosen by firm u_i. Consider the firms u_i with $T_i \geq t^{-(k-1)}$. Clearly, these firms can accept only applicants in $C_2, \ldots, C_{k-1}$, with a contribution to social welfare of at most $\sum_{j=2}^{k-1} |C_j| t^{-j}$. Similarly, consider the set of firms F choosing threshold $T_i = t^k$. Their contribution to the social welfare is bounded by $\sum_{j=2}^{k-1} |C_j| t^{-j}$ from applicants among $C_2, \ldots, C_{k-1}$ and at most $|F| \cdot |C_k| t^{-k}$ from applicants in C_k. Let X_i be the indicator variable defined by $X_i = 1$ if and only if $T_i = t^{-k}$. Note that by choice of k, $\mathbb{E}\left[|F|\right] = \mathbb{E}\left[\sum_{i=1}^m X_i\right] \leq m/(t-1)$.

Consider the firms with threshold $T_i = t^{-k'}$ with $k' \geq k+1$. Let $S := \sum_{j=2}^{k'} t^{2j}$. Let W_i be the expected value of an applicant matched to a firm u_i with threshold $T_i = t^{-k'}$. In expectation, W_i is bounded by

$$\sum_{j=2}^{k'} \frac{t^{2j}}{S} \cdot t^{-j} = O\left(\frac{t^{k'}}{t^{2k'}}\right) = O(t^{-(k+1)}) \ ,$$

since the next accepted applicant of value at least $t^{-k'}$ is distributed as $w(v_{j'})$ conditioned on containment in $\{t^{-2}, \ldots, t^{-(k+1)}\}$, except when we run out of applicants, in which case the value is even zero. Suppose that an applicant gets an offer by firm u_i, but decides to go for another firm. Observe that this has no influence on the distribution of accepted applicants for firm u_i. Summarizing the arguments above, the social welfare of the assignment computed by $\mathcal{A}$ is at most

$$2\left(\sum_{j=2}^{k-1} |C_j| t^{-j}\right) + \sum_{i=1}^{m} X_i t^{-k} + \sum_{i=1}^{m} W_i. \tag{3}$$

By Chernoff bounds, we have for each $j = 2, \ldots, t$ that $|C_j| \notin [(1/2)t^{2j}, 2t^{2j}]$ with probability at most $2\exp(-t^{2j}/8)$ over $\mathcal{I}$. Let E denote the event that $|C_j| \in [(1/2)t^{2j}, 2t^{2j}]$ holds for all $2 \leq j \leq t$. By union bound, E occurs with probability at least $1 - 2t\exp(-t^2/8) = 1 - \exp(-\Omega(t^2))$. If E occurs, then $\sum_{j=2}^{k-1} |C_j| t^{-j} \leq \sum_{j=1}^{k-1} 2t^j = \Theta(t^{k-1})$. Additionally, since each firm acts independently, Chernoff bounds give that $\sum_{i=1}^m X_i \geq 2m/(t-1)$ with probability at most

$\exp(-(1/3)m/(t-1))) = \exp(-\Omega(t^{2k-1}))$. Analogously, $\sum_{i=1}^{m} W_i = O(t^{k-1})$ with probability at least $1 - \exp(-\Omega(t^{k-1}))$

By union bound, we obtain that (3) is bounded by $O\left(t^{k-1}\right) + O\left(\frac{m}{t}t^{-k}\right) = O(t^{k-1})$ with probability $1 - \exp(-\Omega(t^{k-1}))$. For any matching $M^{\mathcal{A}}$ resulting from $\mathcal{A}$ it always holds that $w(M^{\mathcal{A}}) \leq m$. Thus, since $k \geq 2$, $m = \Theta(t^k)$ and $t = \Theta(\log n / \log\log n)$, we have

$$\mathbb{E}_{\mathcal{I},\mathcal{A}}[w(M^{\mathcal{A}})] \leq O(t^{k-1}) + \exp(-\Omega(t^{k-1}))m = O(t^{k-1}) \ . \tag{4}$$

Note that under the event E defined earlier, the optimum solution M^* has value at least $w(M^*) \geq \sum_{j=2}^{k}(1/2)t^j = \Theta(t^k)$. This implies that

$$\mathbb{E}_{\mathcal{I}}[w(M^*)] \geq \Theta(t^k) \cdot (1 - \exp(-\Omega(t^2)) = \Theta(t^k) \ .$$

Hence, in the iid model, the ratio of expectations is at least

$$\frac{\mathbb{E}_{\mathcal{I}}[w(M^*)]}{\mathbb{E}_{\mathcal{I},\mathcal{A}}[w(M^{\mathcal{A}})]} = \Omega(t) = \Omega\left(\frac{\log n}{\log\log n}\right) .$$

We extend this analysis to the secretary model by finding a *fixed* instance with a large competitive ratio. By (4), there is an instance in the support of $\mathcal{I}$ such that $\mathbb{E}_{\mathcal{A}}[w(M^{\mathcal{A}})] = O(t^{k-1})$. We need an instance that also has a large offline optimum. Towards this end, we compute

$$\begin{aligned} \mathbb{E}_{\mathcal{I},\mathcal{A}}[w(M^{\mathcal{A}}) \mid w(M^*) = \Omega(t^k)] &\leq \frac{\mathbb{E}_{\mathcal{I},\mathcal{A}}[w(M^{\mathcal{A}})]}{\Pr[w(M^*) = \Omega(t^k)]} \\ &\leq \frac{\mathbb{E}_{\mathcal{I},\mathcal{A}}[w(M^{\mathcal{A}})]}{1 - \exp(-\Omega(t^2))} = O(t^{k-1}) \ . \end{aligned}$$

Consequently, there is an instance I in the support of $\mathcal{I}$ for which $w(M^*) = \Omega(t^k)$ and $\mathbb{E}_{\mathcal{A}}[w(M^{\mathcal{A}})] = O(t^{k-1})$. On this instance I, the competitive ratio is bounded by

$$\frac{w(M^*)}{\mathbb{E}\left[w(M^{\mathcal{A}})\right]} = \Omega(t) = \Omega\left(\frac{\log n}{\log\log n}\right) .$$

This shows the lower bound when $\mathcal{A}$ does not know m. Note that we can introduce auxiliary firms, for which every applicant has value $\epsilon \ll t^{-t}$. These firms do not contribute significantly to the value of any matching. In this way, we can keep the number of firms fixed to, say, n, and provide it to $\mathcal{A}$. The lower bound continues to hold, since the critical challenge remains to estimate the number of firms that contribute significantly to $w(M^*)$. □

C.4 Proof of Theorem 4

Proof. For convenience, we allow each firm to send offers even after it has been matched, i.e., it still sends *virtual offers*, which will always be rejected. Assume that all firms share the same preference list $(v_1, \ldots, v_n)$. Then, if an applicant receives an offer from a firm, every other firm also sends her an offer which might be virtual. Note that in the algorithm every firm sends out offers at most m times, thus no more than m applicants would receive offers from the same firm. It follows that when receiving *some* (potentially virtual) offer, an applicant also sees non-virtual offers and chooses the best from them.

Denote the set of all the applicants who receive offers by S. Analogously to the classical secretary problem and as in the proof of Lemma 5, for each applicant v who is among the best m applicants, we have $\mathbf{Pr}[v \in S] \geq \frac{r-1}{n} \ln(\frac{n}{r-1})$ which is constant for $r = \Theta(n)$.

Lemma 3. *Let $r = \Theta(n)$ and assume that all firms have the same preference list $(v_1, \ldots, v_n)$ and for each applicant v_i, all incident weights $\{w(u, v_i) \mid u \in U\}$ are independently distributed from the same distribution D_i. Then for each $1 \leq k \leq m$, we have that v_k is matched to her best firm with constant probability.*

Proof. Denote the (random) arrival order of applicants by τ and let $s_{\tau,i}$ be the i-th applicant who receives offers. Fix an applicant $v = v_k$, $1 \leq k \leq m$, among the best m applicants.

First, for every τ where $s_{\tau,j} = v$ and $j > 1$, by swapping the position between $s_{\tau,j-1}$ and v we can obtain a new order τ'. In the new arrival order, v becomes the $(j-1)$-th to receive offers, i.e., $s_{\tau',j-1} = v$. Clearly, for two different arrival orders τ_1 and τ_2 with $s_{\tau_1,j} = s_{\tau_2,j} = v$, the corresponding new orders τ_1' and τ_2' are also different. Thus $|\{\tau | s_{\tau,j-1} = v\}| \geq |\{\tau | s_{\tau,j} = v\}|$. Therefore $\Pr_\tau[s_{\tau,j-1} = v \mid v \in S] \geq \Pr_\tau[s_{\tau,j} = v \mid v \in S]$ for all $j > 1$.

Now given that $s_{\tau,j} = v$, among the m offers v has received, $j-1$ of them are virtual and must be rejected. If the best offer for v is among the remaining $m - j + 1$ ones, then v will get her best offer. Since all the weights of edges incident to v are generated independently from the same distribution, this event occurs with probability $\frac{m-j+1}{m}$ and is decreasing in j, therefore

$$\mathbf{Pr}[v_i \text{ gets her best firm} \mid v_i \in S] = \sum_{j=1}^{m} \mathbf{Pr}[s_{\tau,j} = v_i \mid v_i \in S] \cdot \frac{m-j+1}{m} \quad .$$

Since $\sum_{j=1}^{m} \mathbf{Pr}[s_{\tau,j} = v_i \mid v_i \in S] = 1$, by Chebyshev's sum inequality, we have

$$\mathbf{Pr}[v_i \text{ gets her best firm} \mid v_i \in S] \geq \frac{1}{m} \sum_{j=1}^{m} \frac{m-j+1}{m} \geq \frac{1}{2} \quad .$$

Combining this with $\mathbf{Pr}[v \in S] \geq \frac{r}{n} \ln(\frac{n}{r})$, the claim follows. □

Given that the fluctuations in applicants' quality is small enough to have a little effect on the preference lists, it is easy to extend the result to show constant competitive ratio for the case of small variance. We first show that in this regime, indeed the fluctuations keep the same preference list with high probability.

Lemma 4. *When $\psi = \omega(n)$, for any given sequence $q_1 > \cdots > q_n$, with probability approaching 1 it holds that*

(1) for each applicant v_i and firm u, we have $|w(u, v_i) - q_i| < \delta_{min}/2$,
(2) each firm has the same preference list of applicants $(v_1, \dots, v_n)$.

Proof. For some applicant v_i and firm u, note that $w(u, v_i)$ is sampled independently from $N(q_i, \sigma^2)$. By Chebyshev's inequality, we conclude

$$\Pr\left[|w(u, v_i) - q_i| \leq \frac{\delta_{min}}{2}\right] \leq \frac{4\sigma^2}{\delta_{min}} = \frac{4}{\psi^2} \enspace .$$

By union bound, this event holds for all applicants and firms with probability at least $1 - \frac{4nm}{\psi^2}$. Given that $m \leq n$ and $\psi = \omega(n)$, this probability approaches 1 when n goes to infinity. This proves that part (1) of the lemma holds with high probability.

For the second part, assume that (1) holds and fix a particular firm u and write $x_i := w(u, v_i)$. Recall that by assumption $q_1 > q_2 > ... > q_n$. Given that $q_i - q_{i-1} \geq \delta_{min}$, we conclude that

$$x_i - x_{i-1} \geq (q_i - q_{i-1}) - |x_i - q_i| - |x_{i-1} - q_{i-1}| > (q_i - q_{i-1}) - \delta_{min} \geq 0 \enspace .$$

Hence, $x_1 > \cdots > x_n$ and u has the preference list $(v_1, \dots, v_n)$. □

Note that Lemma 4 allows us to apply Lemma 3 by letting D_i be the truncated Gaussian distribution obtained by conditioning $w(u, v_i) \sim N(q_i, \sigma^2)$ to be contained in $(q_i - \delta_{min}/2, q_i + \delta_{min}/2)$.

Finally, for the proof of the theorem denote the set of the best m applicants by T. Clearly, $\mathbb{E}\left[w(M^*)\right] \leq \sum_{v_i \in T} \max_{u \in U} w(u, v_i)$. According to Lemmas 4 and 3, the algorithm guarantees that for every $v_i \in T$, she will be matched to her best firm with constant probability. By linearity of expectation, $\mathbb{E}\left[w(M^{\mathcal{A}})\right] \geq \sum_{v_i \in T} c \cdot \max_{u \in U} w(u, v_i)$ for some constant $c > 0$, concluding the result. □

As a corollary, we obtain the following.

Corollary 1. *Each firm has a probability of $\Omega(\frac{1}{m})$ to obtain the best applicant.*

Proof. By Lemma 4, with probability approaching 1 all firms consider the same applicant as the best. Denote the best applicant by v. By the fact that $\mathbf{Pr}\left[v \in S\right] \geq \frac{r}{n} \ln(\frac{n}{r})$ is a constant, v is matched to some firm with constant probability. Since there is no difference between the firms, each firm has a probability of $\frac{1}{m}$ to be chosen. □

C.5 Proof of Theorem 5

Assume that the arrival order of applicants is $\{v_1, v_2, \dots, v_n\}$. Denote v_j's quality by q_j. For $j \geq r$, $D(u, v_j)$ denotes the event that $w(u, v_j)$ does not exceed the threshold T_u set by u.

Claim. Let $u \in U$ and $j \geq r$. Let $X_1, \dots, X_{r-1} \sim N(q_i - \delta_{max}, \sigma^2)$ and $X_i \sim N(q_i, \sigma^2)$ be independent. Then

$$\mathbf{Pr}\left[D(u, v_i) \mid \{q_i\}_{i=1}^n\right] \geq \mathbf{Pr}\left[\exists 1 \leq j \leq r-1 : X_j > X_i\right] \enspace .$$

Algorithm 5: The classic secretary problem algorithm for firm u.

Reject the first $(r-1)$ applicants, denote the set by R
$T_u \leftarrow \max_{j \in R} w(u, v_j)$
for *applicant* v_t *arriving in round* $t = r, \ldots, n$ **do**
  if $w(u, v_t) \geq T_u$ **then**
    Make an offer to v_t, stop if v_t accepts

Proof. Note that $D(u, v_i)$ is equivalent to the existence of some $1 \leq j \leq r-1$ with $w(u, v_j) > w(u, v_i)$. This probability is minimized if $|q_j - q_i|$ is maximized for all $1 \leq j \leq r-1$. □

Proposition 2. *For each $u \in U$ and $i \geq r$,*

$$\boldsymbol{Pr}[D(u, v_i) \mid \{q_i\}_{i=1}^n] \geq 1 - \frac{1}{r} - \frac{r-1}{\sqrt{2\pi}}\varphi.$$

Proof. By the previous claim, we have

$$\mathbf{Pr}\left[D(u, v_i) \mid \{q_i\}_{i=1}^n\right] \geq 1 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} \left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x_i + \delta_{max}}{\sqrt{2}\sigma}\right)\right)^{r-1} \mathrm{d}x_i.$$

Let $\mathrm{erf}(\cdot)$ be the error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \mathrm{d}t.$$

With

$$\frac{\mathrm{d}(\frac{1}{2} + \frac{1}{2}\mathrm{erf}(\frac{x}{\sqrt{2}\sigma}))^{r-1}}{\mathrm{d}x} = \frac{r-1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x}{\sqrt{2}\sigma}\right)\right)^{r-2} \leq \frac{r-1}{\sqrt{2\pi}\sigma},$$

by the Mean Value Theorem, we have

$$\left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x_i + \delta_{max}}{\sqrt{2}\sigma}\right)\right)^{r-1} \leq \left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x_i}{\sqrt{2}\sigma}\right)\right)^{r-1} + \frac{r-1}{\sqrt{2\pi}\sigma} \cdot \delta_{max}.$$

Therefore,

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} \left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x_i + \delta_{max}}{\sqrt{2}\sigma}\right)\right)^{r-1} \mathrm{d}x_i$$
$$\leq \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} \left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{x_i}{\sqrt{2}\sigma}\right)\right)^{r-1} \mathrm{d}x_i + \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} \frac{(r-1)\delta_{max}}{\sqrt{2\pi}\sigma} \mathrm{d}x_i.$$

Note that the first term is exactly the probability that x_i is the highest value among r random variables drawn independently from the same normal distribution $N(0, \sigma^2)$, which equals $\frac{1}{r}$. And the second term

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} \frac{(r-1)\delta_{max}}{\sqrt{2\pi}\sigma} \mathrm{d}x_i = \frac{(r-1)\delta_{max}}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x_i^2}{2\sigma^2}} \mathrm{d}x_i = \frac{(r-1)\delta_{max}}{\sqrt{2\pi}\sigma},$$

thus,

$$\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{x_i^2}{2\sigma^2}}\left(\frac{1}{2}+\frac{1}{2}\mathrm{erf}\Big(\frac{x_i+\delta_{max}}{\sqrt{2}\sigma}\Big)\right)^{r-1}\mathrm{d}x_i \le \frac{1}{r}+\frac{r-1}{\sqrt{2\pi}}\frac{\delta_{max}}{\sigma}=\frac{1}{r}+\frac{r-1}{\sqrt{2\pi}}\varphi.$$

This concludes the claim. □

Lemma 5. *Let $u \in U$ and $r = \Theta(n)$. With constant probability, u gets its most valuable applicant.*

Proof. Let $B(u, v_i), P(u, v_i)$, and $A(u, v_i)$ be the events that v_i is the best applicant for u, v_i receives an offer from u and v_i receives and accepts an offer from u, respectively. By random arrival, for each $i \ge r$, we have $\Pr[B(u, v_i)] \ge 1/n$ and $\Pr[P(u, v_i) \mid B(u, v_i)] \ge \frac{r-1}{i-1}$ (for the latter, note that if the most valuable applicant among $v_1, \ldots, v_{i-1}$ is among $v_1, \ldots, v_{r-1}$, u cannot have made a previous offer). Hence,

$$\mathbf{Pr}\left[u \text{ gets its best}\right] \ge \sum_{i=r}^{n}\frac{1}{n}\cdot\frac{r-1}{i-1}\cdot\mathbf{Pr}\left[A(u, v_i) \mid B(u, v_i), P(u, v_i)\right]$$

If no other firm sends an offer to v_i, $A(u, v_i)$ must be true if $P(u, v_i)$ holds. For any $u' \ne u$, $P(u', v_i)$ only depends on the value of $\{q_i\}_{i=1}^n$. Thus, given $\{q_i\}_{i=1}^n$, all the events $\{P(u', v_i) \mid u' \ne u\}$ are independent from each other, and are all independent from $B(u, v_i)$ and $P(u, v_i)$. Hence,

$$\begin{aligned}&\mathbf{Pr}\left[A(u, v_i) \mid B(u, v_i), P(u, v_i)\right]\\ &\qquad \ge \mathbb{E}_{\{q_i\}_{i=1}^n}\left[\prod_{u'\ne u}\mathbf{Pr}\left[\overline{P(u', v_i)} \mid \{q_i\}_{i=1}^n\right]\middle| B(u, v_i), P(u, v_i)\right].\end{aligned}$$

Since $D(u', v_i)$ implies $\overline{P(u', v_i)}$ and Proposition 2 is applicable for any $\{q_i\}_{i=1}^n$, we can bound

$$\mathbf{Pr}\left[A(u, v_i) \mid B(u, v_i), P(u, v_i)\right] \ge \left(1-\frac{1}{r}-\frac{r-1}{\sqrt{2\pi}}\varphi\right)^{m-1}.$$

Note that $p := \frac{r}{n}$ is constant and $\varphi \le \frac{c}{r(r-1)}$ for some constant c. Then this term is bounded from below by

$$\left(1-(1+\frac{c}{\sqrt{2\pi}})\frac{1}{r}\right)^{m-1}.$$

Let $c' = 1 + \frac{c}{\sqrt{2\pi}}$ and $\alpha \in (0, 1]$ be a constant such that $m \le \alpha n$. We compute

$$\mathbf{Pr}\left[u \text{ gets its best}\right] \ge \sum_{i=r}^{n}\frac{1}{n}\cdot\frac{r-1}{i-1}\cdot\left(1-\frac{c'}{r}\right)^{m-1}$$

$$\geq \left(\left(1-\frac{c'}{r}\right)^{r}\right)^{\frac{\alpha}{p}} \sum_{i=r}^{n} \frac{1}{n} \cdot \frac{r-1}{i-1},$$

which is bounded from below by a constant depending on α, c', and p.

$\square$

The previous lemma implies that the classic algorithm for the secretary problem is almost an optimal strategy for firms. It guarantees for every single firm the best response with constant probability. By linearity of expectation, the expected social welfare is at least a constant fraction of the optimum, completing the proof of Theorem 5.