Platform Design in Curated Dating Markets

(Authors' names blinded for peer review)

Problem definition: Motivated by online dating apps, we consider the problem of selecting the subset of profiles to show to each user in each period in a two-sided matching platform. Users on each side observe the profiles set by the platform and decide which of them to like. A match occurs if and only if two users mutually like each other, potentially in different periods. We study how platforms should make this decision to maximize the expected number of matches under different platform designs, varying (i) how users interact with each other, i.e., whether one or both sides of the market can initiate an interaction (one and two-directional, resp.), and (ii) the timing of matches, i.e., whether the platform allows non-sequential matches (i.e., when both users see and like each other in the same period) in addition to sequential ones.

Methodology/Results: We focus on the case with two periods and study the performance of different approaches. First, we show that natural approaches in the online matching and assortment literature, including greedy and perfect matching, have worst-case performance arbitrarily close to zero. Second, we show that algorithms that exploit the submodularity of the problem and properties of its feasible region can achieve constant factor approximation guarantees that depend on the platform design, ranging from $1-1/e$ to $1/3$. Finally, we show that the *Dating Heuristic* (DH) (Rios et al. 2023), which is commonly used and achieves good performance in practice, provides an approximation guarantee of $1-1/e$ for all platform designs.

Managerial Implications: We show theoretically and empirically that the performance of the DH is robust to the platform design. Our simulation results—using real data from our industry partner—also show that platforms using a one-directional design should initiate interactions with the side that leads to the smallest expected backlog per profile displayed, balancing size and selectivity. Moreover, we find that a one-directional design can lead to at least half of the matches obtained with a two-directional design. Finally, our results show that avoiding non-sequential matches has no sizable effect, regardless of the platform design.

Key words: Dating, Platform Design, Matching Markets, Submodular Optimization.

1. Introduction

A common feature of many dating platforms is their *curated* approach, presenting users with a limited subset of profiles rather than granting unrestricted access to all the available ones. This strategy, used in platforms such as Bumble, Hinge, Coffee Meets Bagel, and The League, aims to enhance users' experience by promoting more meaningful connections. Moreover, limiting daily profile views helps reduce app fatigue and fosters a sense of anticipation and excitement that can translate into user satisfaction and retention.

Despite their shared motivation, curated dating apps often differ in design and functionalities, catering to various preferences and needs. One design feature that differs across these platforms is *the sequence of interactions*. For instance, the "Ladies Choice" model behind Coffee Meets Bagel uses a one-directional approach, whereby men move first by evaluating a set of profiles. Then, women are shown only profiles of men who liked them in the first place. Bumble follows a similar approach, allowing only women to reach out first. Other platforms, such as Hinge and The League, use a two-directional approach, whereby both sides of the market can screen profiles and initiate the path towards a match.

A second design feature that differentiates some platforms is *the timing of matches*. Most platforms employ a sequential design, whereby users in a match may see each other's profiles in different periods and, thus, a match results from them sequentially liking each other. This design reduces wasteful displays, as one match member sees the other only conditional on a first like. However, some platforms require matches to be non-sequential, i.e., they need both users to see each other in the same period. For instance, Once is a unique dating app that provides users with one "match" per day; thus, both users often see each other in the same period. Other dating apps, such as Bumble and The League, provide video chat or speed-dating services requiring users to see each other simultaneously.

In this paper, we study the problem of deciding the subset of profiles to display to each user in each period to maximize the expected number of matches under different platform designs, whereby a match between two users realizes if, and only if, both users like each other, potentially in different periods. The goal of this paper is twofold. First, we provide a framework that allows for various platform designs, varying two key features to accommodate different functionalities: (i) the sequence of interactions, i.e., whether the platform restricts to one-directional interactions—where one side evaluates first and the second only observes users that previously liked them—or allows two-directional ones—where both sides can start the path towards a match by being the first evaluator; and (ii) the timing of matches, allowing them to be sequential—whereby users see and like each other in different periods conditional on an initial like—or non-sequential—in which users can see and like each other in the same period. Our second goal is to investigate the effectiveness of commonly used algorithms and analyze their performance in theory and practice for the platform designs mentioned above.

1.1. Contributions

As Rios et al. (2023) argue, one-look-ahead policies that consider two periods (current and following) to decide which profiles to show are commonly used and perform well in practice. For this reason, we start by analyzing the two-period version of the problem by extending the model in (Rios et al. 2023) to capture the two design choices mentioned above, i.e., (i) the sequence of interactions and (ii) the timing of matches.

The first variant we analyze is when we only allow for one-directional interactions and sequential matches, i.e., only users on one side of the market can initiate an interaction, and no pair of users see each other in the same period. This case is particularly relevant as (i) it captures the design of several women-oriented platforms (e.g., Coffee Meets Bagel and Bumble) that provide women more control over whom they interact with, and (ii) other two-sided platforms operate in this way, which has motivated recent work, e.g., (Ashlagi et al. 2022, Aouad and Saban 2022, Torrico et al. 2021).

The second variant of the problem extends the first one by allowing sequential matches to happen in both directions, i.e., two-directional interactions with sequential matches. In this case, the platform selects profiles for users on both sides of the market in each period and, thus, both sides of the market can initiate an interaction. However, we assume that no pair of users see each other in the same period, so matches can only occur sequentially in the second period. This design matches dating platforms that update the profiles to be displayed continuously over time (e.g., Hinge and Tinder), responding to users' likes and dislikes decisions.

The third variant expands on the initial one by allowing non-sequential matches in the first period. More precisely, the platform chooses subsets of profiles for the "initiating" side of the market (hence, one-directional interactions) but can also select pairs of users that will see each other in the first period. Thus, matches can also happen if both users in these pairs like each other in the first period. This design is used in some dating apps that combine "standard swiping" with video chat/speed dating, such as Filteroff and The League, where users can evaluate profiles as usual but can also log in at specific times and get matched for short video speed dates.

The last variant of the problem is when interactions are two-directional and matches can happen either (i) sequentially in the second period or (ii) non-sequentially in the first one. This platform design captures the one used by dating apps that precompute (generally in the middle of the night) the profiles to display on a given day and do not adjust them based on users' activity during the same day. Note that this setting is the same as the one originally considered in Rios et al. (2023).

Despite the two-period assumption, we show that many common approaches (such as local greedy and perfect matching) have a worst-case performance arbitrarily close to zero. Moreover, we show that algorithms commonly found in the submodular optimization literature (e.g., the continuous greedy algorithm with dependent randomized rounding (Vondrák 2008, Gandhi et al. 2006) and the local-search algorithm (Lee et al. 2009)) lead to constant factor approximations that differ across platform designs, as summarized in Table 1.

In sharp contrast, we show that the integral version of the Dating Heuristic (DH) introduced by Rios et al. (2023), which involves solving a mixed-integer program with one lookahead period to obtain a solution for the current period accounting for its effect in the following one, achieves a $1 - 1/e$ approximation guarantee for all the platform designs, improving upon the approaches described above. The derivation of this performance guarantee is interesting on its own and relies on two elements. First, we show that we can upper bound the maximum expected number of matches

Table 1 Summary of results.

		Direction of Interactions			
		One-directional		Two-directional	
Timing of Matches	Sequential	$1-1/e$	$\mathbf{1-1/e}$	$1/(2+\epsilon)$	$\mathbf{1-1/e}$
	Non-Sequential	$1/(2+\epsilon)$	$\mathbf{1-1/e}$	$1/(3+\epsilon)$	$\mathbf{1-1/e}$

Note: The results above are for an arbitrary $\epsilon > 0$. These guarantees are relative to the optimal solution of the corresponding DP formulation for each platform design. Guarantees not in bold correspond to those obtained with submodular optimization methods and in bold are those obtained with DH.

that can be obtained in the second period as a result of the profiles displayed in the first period by either adapting DH or using a novel linear program that involves the general distribution of backlogs as decision variables. We refer to the latter as the *distribution problem*. Then, using a duality argument, we establish that the upper bound obtained by the latter is at least that obtained by the former. Our second main ingredient is to show that, given the subsets of profiles displayed in the first period, the distribution of the backlog in the second period is a feasible solution for the distribution problem. Hence, we can use the notion of *correlation gap* (Agrawal et al. 2010), which compares the objective values of our solution and the distribution problem, to establish our guarantee.

To complement our theoretical results, we empirically evaluate the performance of DH and the other benchmarks for different platform designs using real data from a major dating app in the US.[1] In line with the insights in (Rios et al. 2023), we find that allowing for non-sequential matches does not significantly affect the number of matches the platform can generate. More interestingly, DH significantly outperforms the algorithms that rely on submodular optimization techniques. Furthermore, our simulation results demonstrate that platforms using a one-directional design should initiate interactions with the side that leads to the smallest expected backlog per profile displayed, balancing size and selectivity. Lastly, we show that the best one-directional design leads to at least half of the matches generated by the corresponding two-directional design.

We conclude the paper by providing several model extensions. First, we show that DH can be modified to solve the problem with multiple periods and obtain the same $1-1/e$ performance guarantee (i) relative to any semi-adaptive policy for any platform design and (ii) relative to any adaptive policy in the case with one-directional interactions and sequential matches. Second, we analyze the case where the platform can form non-sequential matches in the second period and provide a guarantee of $1/4e$ when the market is sufficiently large and the timing of interactions is one-directional with the most selective side (i.e., with small like probabilities) moving first. Our

[1] This platform provided us with real data to test our proposed algorithms. We keep the app's name undisclosed as part of our NDA.

one-directional assumption on the pickier side aligns with the results in (Kanoria and Saban 2021, Torrico et al. 2021, Shi 2022a), which demonstrate that platforms can reduce market congestion when the more selective side (or the one whose preferences are harder to describe) initiates the matchmaking process. Our large market assumption aligns with previous theoretical work on the online stochastic matching problem, see e.g. (Mehta and Panigrahi 2012, Goyal and Udwani 2023). To prove our guarantee, we show that the gains from allowing non-sequential matches in the second period are relatively small compared to when we allow non-sequential matches in the first period only.

Organization of the paper. In Section 2, we discuss the most related literature. In Section 3, we introduce our model, discuss its two-period version, and introduce DH. In Section 4, we theoretically analyze the two-period version of the problem and provide performance guarantees under different platform designs. In Section 5, we numerically analyze DH and compare it with other benchmarks for different platform designs. In Section 6, we provide several extensions of our baseline model. Finally, in Section 7, we conclude. We defer all the proofs to the Appendix.

2. Related Literature

Our paper is related to several strands of the literature. First, we contribute to the literature on assortment optimization. Most of this literature focuses on one-sided settings, where a retailer must choose the assortment of products to show in order to maximize the expected revenue obtained from a sequence of customers. This model, whose general version was introduced in Talluri and van Ryzin (2004), has been extended to include capacity constraints (Rusmevichientong et al. 2010), different choice models (Davis et al. 2014, Rusmevichientong et al. 2014, Blanchet et al. 2016, Farias et al. 2013), search (Wang and Sahin 2018), learning (Caro and Gallien 2007, Rusmevichientong et al. 2010, Sauré and Zeevi 2013), personalized assortments (Berbeglia and Joret 2015, Golrezaei et al. 2014), reusable products (Rusmevichientong et al. 2020), and also to tackle other problems such as priority-based allocations (Shi 2022b). We refer to Kök et al. (2015) for an extensive review of the current state of the assortment planning literature in one-sided settings.

Over the last couple of years, a new strand of the assortment optimization literature devoted to two-sided markets has emerged. Ashlagi et al. (2022) introduce a model where each customer chooses, simultaneously and independently, to either contact a supplier from their assortment or to remain unmatched. Then, each supplier can either form a match with one of the customers who contacted them in the first place or remain unmatched. The platform's goal is to select the (unconstrained) assortment of suppliers to show to each customer to maximize the expected number of matches. The authors show that the problem is strongly NP-hard, and they provide a constant factor approximation algorithm. Torrico et al. (2021) study the same problem and significantly improve the approximation factor obtained by Ashlagi et al. (2022), which also applies to more general settings. Aouad and Saban (2022) introduce the online version of the model in (Ashlagi et al. 2022), i.e., customers arrive and make their choices sequentially over time and, after some time, each supplier can choose to match with at most one of the customers that chose them. The authors show that when suppliers do not accept/reject requests immediately, then a simple greedy policy achieves a 1/2-factor approximation. Aouad and Saban (2022) also propose balancing algorithms that perform relatively well under the Multinomial and Nested Logit models. Notice that all these papers analyze sequential two-sided matching markets, where only one side can start the path toward a match. Users on the other side only respond by deciding which user to match with among those who contacted them in the first place. Moreover, customers are limited to choosing only one supplier in their assortments, and suppliers can only choose one customer among those who initially contacted them. Hence, we contribute to this literature by studying different platform designs with users making multiple selections, enabling two-directional interactions and non-sequential matches.

Within the emerging assortment optimization literature in two-sided markets, the closest paper to ours is (Rios et al. 2023). The authors introduce a finite horizon model where a platform chooses a subset of profiles for each user (on both sides of the market), and users can like/dislike as many of the profiles as they want, forming a match if they like each other. Hence, their model is equivalent to ours when we allow two-directional interactions and non-sequential matches. The authors introduce a family of heuristics (*Dating Heuristics*) that rely on a one-period lookahead optimization problem to decide which subsets of profiles to show in each period, and they show through a field experiment that their heuristics perform well in practice. However, Rios et al. (2023) provide no theoretical guarantees for their proposed algorithms, as they focus on estimating like probabilities and showing how these depend on the number of matches recently obtained. Hence, our paper complements this work in that we provide the first performance guarantee for the problem involving two-directional interactions and non-sequential matches. Moreover, we extend their model and the analysis to cover other relevant platform designs.

The second stream of literature related to our paper is on the design of matching platforms. Starting with the seminal work of Rochet and Tirole (2003), this literature has focused on participation, competition, and pricing, highlighting the role of cross-side externalities in different settings, including ridesharing (Besbes et al. 2021), labor markets (Aouad and Saban 2022, Besbes et al. 2023), crowdsourcing (Manshadi and Rodilitz 2022), public housing (Arnosti and Shi 2020), and volunteering platforms (Manshadi et al. 2022). In the dating context, Halaburda et al. (2018) show that two platforms can successfully coexist charging different prices by limiting the set of options offered to their users. Cui and Hamilton (2024) analyze subscription pricing strategies and show that platforms can achieve higher revenue and welfare by varying the term of subscriptions. Kanoria and Saban (2021) study how the search environment can impact users' welfare and the performance of the platform. They find that simple interventions, such as limiting what side of the market reaches out first or hiding quality information, can considerably improve the platform's outcomes. Immorlica et al. (2022) study settings where the platform can use its knowledge about agents' preferences to guide their search process, and show that the platform can induce an equilibrium with approximately optimal welfare and simplify the agents' decision problem by limiting choice. Finally, Celdir et al. (2024) analyze the popularity bias present in some recommendation systems and show that platforms can increase the number of matches and their revenue by recommending popular profiles as long as they do not become "out of reach". All these models consider a stylized matching market, where users interact with the other side of the market and leave the platform upon getting a match. Hence, we contribute to this literature by allowing agents to like multiple profiles and potentially match with many of them within the time horizon.

The last stream of the literature related to our work, motivated by applications in kidney exchange, is the stochastic matching problem in the query-commit model, also known as stochastic probing with commitment. In this problem, the matchmaker can query the edges of a general graph (e.g., to assess the compatibility between a pair donor-patient) to form a match of maximum cardinality using the accepted edges. Starting with Chen et al. (2009), who introduce the problem (with patience constraints) and provide the first performance guarantee, most of this literature has focused on settings where the matchmaker queries only one edge at a time, see e.g. (Adamczyk 2011, Costello et al. 2012, Bansal et al. 2012, Gamlath et al. 2019, Hikima et al. 2021, Brubach et al. 2021, Jeloudar et al. 2021). Nevertheless, Chen et al. (2009) also study a case closer to ours, where the planner can query a matching in each period. The authors show that a greedy algorithm that selects the edges with the highest success rate in decreasing order provides a 1/4-approximation to the optimal online algorithm when forced to commit. Jeloudar et al. (2021) study the case when there is no such commitment (i.e., the matchmaker can choose not to use an accepted edge) and show that a similar greedy algorithm achieves a 0.316-approximation guarantee. Our problem is similar in that the edges are of uncertain reward (given that like decisions are stochastic). However, we focus on selecting (or in other words, probing) subsets of profiles and each edge realization (a match) depends on the result of two outcomes. Moreover, we show that the greedy approaches used in this strand of the literature perform arbitrarily badly when adapted to our setting.

3. Model

In this section, we introduce a model of a two-sided market mediated by a platform that decides, in each period, which subset of profiles (if any) to show to each user to maximize the expected number of matches. Users can like as many profiles as they want, and if two users mutually like each others' profiles (potentially in different periods), a match between them materializes. In Section 3.1, we formalize each model component. In Section 3.2, we focus on the two-period version of the problem. Finally, in Section 3.3, we introduce the Dating Heuristic that we later analyze in Section 4.

3.1. Problem formulation

We now describe the dating market, the matching process, and the platform's goal and design.

Dating Market. Consider a dating platform that faces a discrete-time problem over a finite horizon of T periods, where the set of periods is denoted by $[T] = \{1, \dots, T\}$. Let I and J be the sets of users participating on this platform, which are known at the beginning and remain fixed throughout the horizon, i.e., no users enter or leave the platform. This assumption captures settings involving short-term horizons, e.g., platforms operating daily. To simplify the exposition, we focus on a heterosexual market; thus, we can model the users and their interactions through a bipartite graph.

At the beginning of the horizon, each user $\ell \in I \cup J$ reports their profile information (e.g., age, height, race, etc.) and their preferences regarding each of these dimensions (e.g., preferred age and height ranges, preferred races, etc.), which remain fixed throughout the horizon. The platform uses this information to compute the initial set of potential partners $\mathcal{P}_\ell^1$—or simply *potentials*—for each user ℓ, i.e., the set of users that ℓ prefers over being single and for whom ℓ satisfies their preferences. Since our market is a bipartite graph, then $\mathcal{P}_i^1 \subseteq J$ for each $i \in I$ and $\mathcal{P}_j^1 \subseteq I$ for each $j \in J$.

Matching Process. In each period $t \in [T]$, the platform selects a subset of potentials $S_\ell^t \subseteq \mathcal{P}_\ell^t$ to be displayed to each user $\ell \in I \cup J$, where $\mathcal{P}_\ell^t$ is the set of potentials for user ℓ at the beginning of period t. As we later discuss, the sets of potentials are updated at the end of each period to capture users' decisions and prevent users from evaluating a profile they have already seen in the past or someone who has already disliked them. To mimic our industry partner's practice, we assume that the maximum number of profiles a user can see in a given period is fixed and equal to K_ℓ, i.e., $|S_\ell^t| \le K_\ell$ for all $\ell \in I \cup J$ and $t \in [T]$.

For each user $\ell \in I \cup J$ and a profile $\ell' \in S_\ell^t$, let $\Phi_{\ell,\ell'}^t$ be the binary random variable that indicates whether ℓ likes ℓ' in period t, i.e., $\Phi_{\ell,\ell'}^t = 1$ if ℓ likes ℓ' and $\Phi_{\ell,\ell'}^t = 0$ otherwise.[2] In addition, let $\phi_{\ell,\ell'}^t = \mathbb{P}(\Phi_{\ell,\ell'}^t = 1)$ be the probability that ℓ likes ℓ' in period t; in the remainder of the paper, we refer to it as the *like probability*.[3] These probabilities are known to the platform and are independent across users and periods. Independence across users is reasonable because individuals do not know whether other users liked them in the past, while independence across periods simplifies the analysis. Indeed, Rios et al. (2023) find that the number of matches obtained in the recent past affects users' future

[2] We focus on cases where users can either like or not like a profile, ruling out the skip option that is part of our partner's platform. As discussed in (Rios et al. 2023), this is without loss of generality because 5% of profiles are skipped. Hence, throughout this paper, dislike and non-like are equivalent.

[3] Note that $\phi_{\ell,\ell'}^t$ is not necessarily symmetric, i.e., $\phi_{\ell,\ell'}^t$ might be different from $\phi_{\ell'\ell}^t$.

behavior, so they assume that like probabilities depend on past matches. Since deriving provable guarantees is substantially more challenging in the past-dependent setting and the magnitude of the *history* effect identified in (Rios et al. 2023) is small relative to the overall like probabilities, we omit correlations across periods. Moreover, note that there is no dependence on the subset of profiles displayed on the like probabilities. This assumption simplifies the analysis, allowing users to potentially like several profiles without having to model preferences over sets. This assumption is without major practical loss, as Rios et al. (2023) observe that like probabilities are almost independent of the rest of the profiles displayed. Finally, let $\beta^t_{\ell,\ell'} = \phi^t_{\ell,\ell'} \cdot \phi^t_{\ell'\ell}$ be the probability of a match between users ℓ and ℓ' conditional on them seeing each other in period t.

Let $\mathcal{B}^t_\ell$ be the backlog of user $\ell \in I \cup J$ at the beginning of period t, i.e., the subset of users that liked ℓ's profile before period t, but have not been shown to ℓ yet. Then, the set of potentials and the backlog of user ℓ in period t can be computed as:

$$\mathcal{P}^t_\ell = \mathcal{P}^{t-1}_\ell \setminus (S^{t-1}_\ell \cup D^{t-1}_\ell), \quad \mathcal{B}^t_\ell = \left(\mathcal{B}^{t-1}_\ell \cup L^{t-1}_\ell\right) \setminus S^{t-1}_\ell, \tag{1}$$

where the sets $L^{t-1}_\ell = \left\{\ell' : \ell \in S^{t-1}_{\ell'} \text{ and } \Phi^{t-1}_{\ell',\ell} = 1\right\}$ and $D^{t-1}_\ell = \{\ell' : \ell \in S^{t-1}_{\ell'} \text{ and } \Phi^{t-1}_{\ell'\ell} = 0\}$ correspond to the sets of users that liked and disliked ℓ in period $t-1$, respectively. In words, user ℓ's set of potentials in period t can be obtained by excluding from ℓ's set of potentials in the previous period (i) the set of users displayed to ℓ in period $t-1$ and (ii) the set of users who disliked ℓ in $t-1$. Similarly, the backlog of user ℓ in period t corresponds to their backlog in period $t-1$ adding the set of users who liked ℓ in $t-1$ and removing those profiles displayed to ℓ in period $t-1$. Note that, for any user $\ell \in I \cup J$ and period $t \in [T]$, (i) $\mathcal{B}^t_\ell \subseteq \mathcal{P}^t_\ell$, i.e., the backlog is a subset of the set of potentials, and (ii) that $\mathcal{P}^t_\ell$ can only decrease as t increases since the market's composition does not vary over time, i.e., no users enter or leave the platform.

A match between users ℓ and ℓ' occurs if both users see and like each other. Let $\mu^t_{\ell,\ell'}$ be the random variable whose value is one if a match between users ℓ and ℓ' happens in period t, and zero otherwise. Then, we know that $\mu^t_{\ell,\ell'} = 1$ if and only if one of the next two events holds: (i) users see and like each other in different periods, i.e., $\left\{\Phi^t_{\ell,\ell'} = 1,\, \ell' \in \mathcal{B}^t_\ell\right\}$ or $\left\{\Phi^t_{\ell',\ell} = 1,\, \ell \in \mathcal{B}^t_{\ell'}\right\}$; or (ii) users see and like each other in period t, i.e., $\ell \in S^t_{\ell'}$, $\ell' \in S^t_\ell$, and $\Phi^t_{\ell,\ell'} = \Phi^t_{\ell'\ell} = 1$. In the former case, we say that the match happens *sequentially*, while in the latter, we say it happens *non-sequentially*. Notice that these two events are disjoint since users see each other at most once and, thus, we cannot simultaneously have that $\Phi^t_{\ell',\ell} = 1$ and $\ell' \in \mathcal{B}^t_\ell$.

Platform's Goal and Design. The platform aims to find a dynamic policy that selects a feasible subset of profiles for each user in each period to maximize the total expected number of matches throughout the entire horizon, as formalized in Problem 1.[4]

[4] We can easily adapt our model to capture settings with a utility value per match; the properties and insights we prove in this work carry over.

A policy $\pi \in \Pi$ for this problem prescribes a sequence of feasible subsets $\mathbf{S}^{t,\pi} = \left\{S_\ell^{t,\pi}\right\}_{\ell \in I \cup J}$ for $t = 1, \ldots, T$ that depends on the initial sets of potentials, the history of profiles shown, the realized like/dislike decisions, and the platform's design. As discussed in Section 1, we consider two design choices: (i) the timing of matches (only sequential or adding non-sequential), and (ii) the sequence of interactions (one or two-directional). We now define these design choices formally.

(i) *Timing of Matches.* We say that a policy π restricts to sequential matches if no pair of users see each other in the same period, i.e., for any pair $(i,j) \in I \times J$ and period t, $i \in S_j^{t,\pi}$ implies that $j \notin S_i^{t,\pi}$ and $j \in S_i^{t,\pi}$ implies that $i \notin S_j^{t,\pi}$. As a result, matches can only happen if users see and like each other in different periods. In contrast, a policy π allows non-sequential matches if no such constraint exists.

(ii) *Sequence of interactions.* We say that a policy π is one-directional if only one market side can initiate the path towards a match. Formally, suppose that π is a one-directional policy and that I is the initiating side. Then, in any period t, π is such that $S_i^{t,\pi} \subseteq \mathcal{P}_i^t$ for each $i \in I$, while $S_j^{t,\pi} \subseteq \mathcal{B}_j^t \cup \left\{i \in \mathcal{P}_j^t \,:\, j \in S_i^{t,\pi}\right\}$ for $j \in J$.[5] In words, users in I can see any profile from their set of potentials, while users in J can only see profiles in their backlog or from users who see them in period t (if non-sequential matches are allowed). In contrast, a policy π is two-directional if any side can initiate an interaction, i.e., π is such that $S_\ell^{t,\pi} \subseteq \mathcal{P}_\ell^t$ for each $\ell \in I \cup J$.

Given a platform design, let Π be the set of all admissible policies satisfying the design requirements and the additional constraints described above (e.g., cardinality of subsets, feasibility based on potentials, etc.). Then, the problem faced by the platform can be formulated as:

Problem 1. Given a set of admissible policies Π, the problem faced by the platform is the following:

$$\max_{\pi \in \Pi} \; \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i \in I} \sum_{j \in J} \mu_{i,j}^{t,\pi}\right]$$

where the expectation is over the probabilistic choices made by the users and possibly the policy's randomization.

Remark 1. Problem 1 can be formulated as an exponentially-sized dynamic program (DP) where the sets of potentials and backlogs fully characterize the state of the system; see Appendix A.1 for a general version of this DP. As previously noted (in Table 1), we use the optimal solution to this DP as the benchmark to compute our performance guarantees. One may consider a stronger benchmark that knows the realizations of all likes/dislikes in advance and computes the maximum matching in the realized graph. However, this benchmark (known as *omniscient* or *offline* optimum) is too

[5] Note that, by the updating formulas in (1), $i \in \mathcal{P}_j^t$ and $j \in S_i^\tau$ for some $\tau < t$ imply that $i \in \mathcal{B}_j^t$, since otherwise i would have been removed from the set of potentials of j at the end of period τ.

strong in our setting since no dynamic policy can achieve a meaningful guarantee relative to it (even in a two-period setting). Similar observations have been previously done in the literature, e.g., Chen et al. (2009), Jeloudar et al. (2021). We include an example in the Appendix A.2 for completeness.

Notation. Let $E = \{\{i,j\} : i \in I,\ j \in J\}$ be the set of all possible undirected edges between I and J, and let $\vec{E} = \vec{E}_I \cup \vec{E}_J$ be the set of all possible directed edges, where $\vec{E}_I = \{(i,j) : i \in I,\ j \in J\}$ and $\vec{E}_J = \{(j,i) : i \in I,\ j \in J\}$ are the sets of all directed arcs between I and J and vice versa, respectively. In the remainder of the paper, we use bold notation to denote vectors and families of subsets, while standard italic notation for their components. Finally, when clear from the context, we remove the dependency on the policy π to simplify the notation.

3.2. Two-Period Model

As discussed in (Rios et al. 2023), one-lookahead policies are commonly used and perform well in practice.[6] For this reason, we focus for now on the two-period version of the problem and return to the general case in Section 6.1.

Given that $T = 2$, we can simplify the notation and denote by $\mathcal{B}_\ell = \mathcal{B}_\ell^2$ the backlog of user ℓ in the second period, i.e., the subset of users that liked ℓ in the first one and that ℓ has not seen yet. Note that $\mathcal{B} = \{\mathcal{B}_\ell\}_{\ell \in I \cup J}$ is a random subset since it depends on the profiles displayed $\mathbf{S}^1$ and the random realizations of the like decisions in the first period. Hereafter, we denote by $\mathbf{B} = \{B_\ell\}_{\ell \in I \cup J}$ the realizations of $\mathcal{B}$. Note that the family of subsets of profiles to display in period $t \in \{1,2\}$, $\mathbf{S}^t = \{S_\ell^t\}_{\ell \in I \cup J}$, can be fully represented by two binary vectors: (i) $\mathbf{x}^t \in \{0,1\}^{\vec{E}}$, which captures sequential interactions; and (ii) $\mathbf{w}^t \in \{0,1\}^E$, which captures non-sequential interactions, i.e., that both users see each other in period t. Then, $\ell' \in S_\ell^t$ if and only if $x_{\ell,\ell'}^t = 1$ or $w_e^t = 1$ with $e = \{\ell, \ell'\}$. As we discuss later, we impose additional constraints to make the latter two events mutually exclusive so that users see each other sequentially or non-sequentially, but not both. To avoid confusion, note that we index the components of $\mathbf{w}^t$ with $e \in E$ instead of ℓ, ℓ' (as we do for the components of $\mathbf{x}^t$) because an edge $e \in E$ is not an ordered pair; instead, we interpret $e \in E$ as a subset of size 2. When clear from context, we omit the clarification $e = \{\ell, \ell'\}$.

Since like/dislike decisions are independent across users, we can characterize the distribution of the random backlog $\mathcal{B}$ using the vector of sequential interactions $\mathbf{x}^1$ as

$$\mathbb{P}_{\mathbf{x}^1}(\mathcal{B} = \mathbf{B}) = \prod_{\ell \in I \cup J} \left[\prod_{\ell' \in B_\ell} \phi_{\ell',\ell}^1 \cdot x_{\ell',\ell}^1 \prod_{\ell' \notin B_\ell} \left(1 - \phi_{\ell',\ell}^1 \cdot x_{\ell',\ell}^1\right) \right]. \tag{2}$$

In words, the backlog of user ℓ at the beginning of the second period is equal to B_ℓ if all users $\ell' \in B_\ell$ saw (i.e., $x_{\ell',\ell}^1 = 1$) and liked user ℓ in the first period, while all users not in B_ℓ either (i) did not see

[6] A one-lookahead policy is a policy that, in every period t, optimizes over the current and the next period, i.e., it considers as horizon $\{t, t+1\}$.

ℓ in their first period (i.e., $x^1_{\ell',\ell}=0$), (ii) ℓ and ℓ' saw each other in the first period (i.e., $w^1_e=1$), or (iii) saw ℓ's profile but did not like it.

Given a backlog $\mathbf{B}=\{B_\ell\}_{\ell\in I\cup J}$ realized after the first period, the problem faced by the platform in the second period is to select a subset of profiles to display to each user to maximize the expected number of matches, which can be formalized as:

$$\begin{aligned}
\max \quad & \sum_{\ell\in I\cup J}\sum_{\ell'\in\mathcal{P}^2_\ell}\phi^2_{\ell,\ell'}\cdot x^2_{\ell,\ell'}+\sum_{e\in E}\beta^2_e\cdot w^2_e \\
\text{s.t.} \quad & \sum_{\ell'\in\mathcal{P}^2_\ell}x^2_{\ell,\ell'}+\sum_{e\in E:\,\ell\in e}w^2_e\le K_\ell, && \forall \ell\in I\cup J, \\
& w^2_e\le \mathbb{1}_{\{\ell'\in\mathcal{P}^2_\ell\setminus B_\ell\}}, && \forall\ell\in I\cup J,\ \ell'\in\mathcal{P}^2_\ell,\ e=\{\ell,\ell'\} \\
& x^2_{\ell,\ell'}\le\mathbb{1}_{\{\ell'\in B_\ell\}}, && \forall\ell\in I\cup J,\ \ell'\in\mathcal{P}^2_\ell \\
& \mathbf{x}^2\in\{0,1\}^{\vec{E}},\ \mathbf{w}^2\in\{0,1\}^E
\end{aligned} \tag{3}$$

where $\mathcal{P}^2_\ell$ is the set of potentials of user ℓ in the second period, updated according to (1). The first family of constraints captures the cardinality requirements, while the second ensures that users can only see each other in the second period if none of them have seen the other before. Finally, the third family of constraints guarantees that sequential interactions, captured by $x^2_{\ell,\ell'}$, involve only profiles in the backlog. Note that we do not need to impose that $x^2_{\ell,\ell'}+w^2_e\le 1$ since it is redundant from the second and third families of constraints.

In Proposition 1, we show that the set function that returns the optimal expected number of matches from (3) given a realized backlog is not submodular in $\mathbf{B}$. Although we can efficiently evaluate this function by solving a linear program (see Proposition 8 in Appendix A), the lack of submodularity poses an extra technical barrier to obtaining provable guarantees.

Proposition 1. *The set function that returns the optimal expected number of matches in Problem* (3) *is not submodular in* $\mathbf{B}$.

The lack of submodularity in Problem (3) follows from allowing non-sequential matches in the second period to compensate for the lack of profiles in a user's backlog. Hence, in the remainder of the paper, we focus on cases where the platform forbids non-sequential matches in the last period. This assumption simplifies the analysis and allows us to leverage submodular optimization techniques and other technical tools, such as the correlation gap (Agrawal et al. 2010), to solve the problem. Moreover, as we empirically show in Section 5, this is without practical loss because most matches materialize sequentially. In Section 6.2, we return to the general case, discuss its challenges, and provide a guarantee for when like probabilities are small.

Let $f(\mathbf{B})$ be the function that returns the maximum expected number of matches obtained from the backlog $\mathbf{B}$ realized after the first period. Since we do not allow non-sequential matches in the second period, we can set $\mathbf{w}^2 = 0$ in (3) and, thus, f corresponds to the maximum expected number of *sequential* matches that can be generated, which can be obtained as follows:

$$f(\mathbf{B}) = \max_{\mathbf{x}^2 \in R^2(\mathbf{B})} \left\{ \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}_\ell^2} \phi^2_{\ell,\ell'} \cdot x^2_{\ell,\ell'} \right\}, \tag{4}$$

where the feasible region $R^2(\mathbf{B})$ is defined as

$$R^2(\mathbf{B}) = \left\{ \mathbf{x}^2 \in \{0,1\}^{\vec{E}} : \sum_{\ell' \in \mathcal{P}_\ell^2} x^2_{\ell,\ell'} \le K_\ell, \quad x^2_{\ell,\ell'} \le \mathbb{1}_{\{\ell' \in B_\ell\}}, \quad \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^2 \right\}. \tag{5}$$

Note that (4) and (5) hold regardless of the platform design. As the sequence of interactions only restricts who moves first, (4) is unaffected because it only entails profiles from the backlog (by (5)) and, thus, interactions that started in the first period. Similarly, allowing non-sequential matches in the first period does not affect the definition of f since the latter corresponds to the maximum expected number of matches the platform can generate from the backlog $\mathbf{B}$.

The fact that we can characterize the maximum expected number of matches from the backlog using (4) and (5) facilitates the problem since $R^2(\mathbf{B})$ is a matroid and the function f is monotone and submodular, as we formalize in Lemma 1.[7]

LEMMA 1. *The function $f(\mathbf{B})$ defined in (4) is monotone and submodular.*

Although the function $f(\mathbf{B})$ can be efficiently evaluated using a greedy algorithm, the two-period problem is significantly more challenging since backlogs depend on the first-period decisions. Formally, let $\mathcal{M}^2(\mathbf{x}^1, \mathbf{w}^1)$ be the maximum expected number of matches that can be produced in the second period given the subsets of profiles displayed in the first period captured by $(\mathbf{x}^1, \mathbf{w}^1)$, i.e.,

$$\mathcal{M}^2(\mathbf{x}^1, \mathbf{w}^1) := \mathbb{E}_{\mathcal{B} \sim \phi^1 \cdot \mathbf{x}^1} \left[f(\mathcal{B}) \right] = \sum_{\mathbf{B} = \{B_\ell\}_{\ell \in I \cup J}} f(\mathbf{B}) \cdot \mathbb{P}_{\mathbf{x}^1} \left(\mathcal{B} = \mathbf{B} \right). \tag{6}$$

where $\phi^1 \cdot \mathbf{x}^1$ denotes the vector with components $\phi^1_{\ell,\ell'} \cdot x^1_{\ell,\ell'}$ and $\mathcal{B} \sim \phi^1 \cdot \mathbf{x}^1$ represents the (random) backlog sampled from vector $\phi^1 \cdot \mathbf{x}^1$. Note that the summation in (6) is over all possible backlog realizations, i.e., all possible $\mathbf{B} = \{B_\ell\}_{\ell \in I \cup J}$ with $B_\ell \subseteq \mathcal{P}^1_\ell$ for each $\ell \in I \cup J$. Moreover, note that $\mathcal{M}^2(\mathbf{x}, \mathbf{w})$ is also monotone and submodular in $\mathbf{x}$, as we show in Lemma 4. Finally, let $\mathcal{M}^1(\cdot)$ be the function that returns the total expected number of non-sequential matches generated in the first period given the first-period decisions $(\mathbf{x}^1, \mathbf{w}^1)$, i.e.,

$$\mathcal{M}^1(\mathbf{x}^1, \mathbf{w}^1) = \sum_{e \in E} \beta^1_e \cdot w^1_e. \tag{7}$$

Given the definitions of $\mathcal{M}^1(\cdot)$ and $\mathcal{M}^2(\cdot)$, we can formalize our two-period problem.

[7] In Appendix A.3, we provide definitions for the concepts of monotonicity, submodularity, and also for the multilinear extension of set functions.

Problem 2. Given a set of admissible policies Π, the two-period version of Problem 1 is:

$$\max\left\{\mathcal{M}^1(\mathbf{x}^1,\mathbf{w}^1)+\mathcal{M}^2(\mathbf{x}^1,\mathbf{w}^1): \quad \left(\mathbf{x}^1,\mathbf{w}^1\right)\in R^{1,\Pi}\right\}, \tag{8}$$

where $R^{1,\Pi}$ is the feasible region determined by Π.

Note that, given our re-formulation of the problem, $R^{1,\Pi}$ is the only element affected by the platform design. For the general case with non-sequential matches in the first period and two-directional interactions, $R^{1,\Pi}$ can be written as

$$R^{1,\Pi}=\Bigg\{\mathbf{x}^1\in\{0,1\}^{\vec{E}},\ \mathbf{w}^1\in\{0,1\}^{E}: x^1_{\ell,\ell'}+x^1_{\ell',\ell}+w^1_e\leq 1,\ \forall\ell\in I\cup J,\ \ell'\in\mathcal{P}^1_\ell,\ e=\{\ell,\ell'\}$$

$$\sum_{\ell'\in\mathcal{P}_\ell} x^1_{\ell,\ell'}+\sum_{e\in E:\,\ell\in e} w^1_e\leq K_\ell,\ \forall\ell\in I\cup J\Bigg\}, \tag{9}$$

where the first family of constraints guarantees that profiles are displayed either sequential or non-sequential (but not both) and the second family of constraints enforces the cardinality constraints. In Section 4, we describe how each platform design translates into a different $R^{1,\Pi}$ and $\mathcal{M}^1$.

In the remainder of the paper, we denote the optimal value of Problem 2 as OPT^Π for each platform design Π and say that a policy achieves an *α-approximation* if it implements a feasible solution leading to an expected number of matches that is at least an $\alpha\in[0,1]$ of OPT^Π in polynomial time. As previously mentioned, we remove the dependency on Π when clear from the context.

3.3. Dating Heuristics.

The Dating Heuristic (DH) (Rios et al. 2023) is a method to solve Problem 2 that has been used in practice due to its simplicity and effectiveness. The idea is to solve a linear program that provides an upper bound for Problem 1 and round its solution, prioritizing profiles that may produce sequential matches. We now describe an integral version of DH for our two-period setting. We focus on the integral version for three reasons: (i) it allows us to derive our guarantees in a simpler way as it does not need the rounding step, which may lead to inefficiencies; (ii) its connection to the two-period model described in Section 3.2 is direct; and (iii) the search for integral solutions does not generate an extra computational burden. Moreover, to simplify the exposition, we assume that users have no backlog in the first period. For completeness, we formally describe the original DH in Algorithm 3 in Appendix A.4. In the remainder of this paper, we omit "integral" to ease exposition.

In the first period, DH solves the following mixed integer program:

$$\begin{aligned}
\max \quad & \sum_{\ell \in I\cup J} \sum_{\ell' \in \mathcal{P}^1_\ell} \frac{1}{2} \cdot w_{\ell,\ell'} \cdot \beta^1_{\ell,\ell'} + \sum_{\ell \in I\cup J} \sum_{\ell' \in \mathcal{P}^1_\ell} y_{\ell,\ell'} \cdot \phi^2_{\ell,\ell'} \\
\text{s.t.} \quad & y_{\ell,\ell'} \le (x_{\ell',\ell} - w_{\ell,\ell'}) \cdot \phi^1_{\ell',\ell}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^1_\ell, \\
& \sum_{\ell' \in \mathcal{P}^1_\ell} x_{\ell,\ell'} \le K_\ell, && \forall \ell \in I \cup J, \\
& \sum_{\ell' \in \mathcal{P}^1_\ell} y_{\ell,\ell'} \le K_\ell, && \forall \ell \in I \cup J, \\
& w_{\ell,\ell'} \le x_{\ell,\ell'},\ w_{\ell,\ell'} \le x_{\ell',\ell},\ w_{\ell,\ell'} = w_{\ell',\ell}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^1_\ell, \\
& x_{\ell,\ell'},\ w_{\ell,\ell'} \in \{0,1\},\ y_{\ell,\ell'} \in [0,1], && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^1_\ell.
\end{aligned} \tag{10}$$

In Problem (10), the decision variables $x_{\ell,\ell'}$ indicate whether or not user ℓ sees ℓ''s profile in the first period. The decision variables $y_{\ell,\ell'}$ represents the probability that user ℓ sees ℓ''s profile in the second period as the follower of the interaction, i.e., ℓ' liked ℓ in the first period and, thus, ℓ' is in ℓ's backlog. The decision variables $w_{\ell,\ell'}$ represent whether or not ℓ and ℓ' both see each other in the first period. The objective in (10) is to maximize the expected number of matches obtained in the two-period horizon, including non-sequential (first term in the objective) and sequential matches (second term in the objective). The first set of constraints defines $\mathbf{y}$ and captures the evolution of the backlog. The second and third sets of constraints limit the number of profiles to display to each user. Finally, the last set of constraints captures the definition of $w_{\ell,\ell'}$.

Based on the optimal solution $(\mathbf{x}^1, \mathbf{y}^1, \mathbf{w}^1)$ of Problem (10), DH (i) constructs the subsets to display in the first period, i.e., $S^1_\ell = \{\ell' \in \mathcal{P}^1_\ell :\ x_{\ell,\ell'} = 1\}$ for each $\ell \in I \cup J$, (ii) updates the potentials and backlogs according to (1) based on the realized like/dislike decisions; and (iii) obtains $\mathbf{x}^2$ by solving (4), displaying the subset of profiles $S^2_\ell = \{\ell' \in \mathcal{P}^2_\ell :\ x^2_{\ell,\ell'} = 1\}$ for each $\ell \in I \cup J$ in the second period. We formalize this procedure in Algorithm 1.

Algorithm 1 Integral Dating Heuristic (DH)

Input: An instance of Problem 2.

Output: A feasible subset of profiles to display to each user in each period.

1: Solve Problem (10) and let $(\mathbf{x}^1, \mathbf{y}^1\, \mathbf{w}^1)$ be the optimal solution.
2: For each user $\ell \in I \cup J$, display the subset of profiles $S^1_\ell = \{\ell' \in \mathcal{P}^1_\ell :\ x^1_{\ell,\ell'} = 1\}$.
3: Observe like/dislike decisions. Update potentials and backlogs following (1).
4: Solve Problem (4) and let $\mathbf{x}^2$ be the optimal solution.
5: For each user $\ell \in I \cup J$, display the subset of profiles $S^2_\ell = \{\ell' \in \mathcal{P}^2_\ell :\ x^2_{\ell,\ell'} = 1\}$.

Note that, although the decision variables in (10) do not correspond exactly to the variables in (9), these two problems are closely related. Indeed, as we show in Lemma 2, we can re-formulate (10) extending the variable space in (2) as follows:

$$\begin{aligned} \max \quad & \sum_{e \in E} w_e \cdot \beta_e^1 + \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}_\ell^1} y_{\ell,\ell'} \cdot \phi_{\ell,\ell'}^2 \\ \text{s.t.} \quad & y_{\ell,\ell'} \leq x_{\ell',\ell} \cdot \phi_{\ell',\ell}^1, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^1, \\ & x_{\ell,\ell'} + x_{\ell',\ell} + w_e \leq 1, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^1,\ e = \{\ell,\ell'\}, \\ & \sum_{\ell' \in \mathcal{P}_\ell^1} x_{\ell,\ell'} + \sum_{e \in E:\, \ell \in e} w_e \leq K_\ell, && \forall \ell \in I \cup J, \\ & \sum_{\ell' \in \mathcal{P}_\ell^1} y_{\ell,\ell'} \leq K_\ell, && \forall \ell \in I \cup J, \\ & \mathbf{x} \in \{0,1\}^{\vec{E}},\ \mathbf{w} \in \{0,1\}^{E},\ \mathbf{y} \in [0,1]^{\vec{E}}. \end{aligned} \tag{11}$$

Lemma 2. *Problem* (10) *is equivalent to Problem* (11).

Based on this reformulation, the connection with Problem 2 becomes clear. The first and second terms in the objective of (11) are analogous to $\mathcal{M}^1(\cdot)$ and $\mathcal{M}^2(\cdot)$, respectively. Moreover, the second and third families of constraints in (11) coincide with $R^{1,\Pi}$ in (9). In Section 4.5, we will exploit this connection to derive performance guarantees for DH.

4. Analysis for Different Platform Designs

Our goal in this section is to provide performance guarantees for various approaches to solve Problem 2 under different platform designs. In Sections 4.1 and 4.2, we consider settings where the platform only allows sequential matches, and we vary whether interactions are one or two-directional, respectively. Then, in Sections 4.3 and 4.4, we relax the sequential-only restriction and allow for non-sequential matches in the first period. Finally, in Section 4.5, we provide a performance guarantee for DH that is common to all platform designs. The omitted proofs from this section can be found in Appendix B.

4.1. One-directional Interactions and Sequential Matches

As discussed in Section 1, many dating platforms restrict which side of the market can reach first to improve users' welfare (Kanoria and Saban 2021, Immorlica et al. 2022). Given our horizon of two periods, the simplest approach to accomplish this is to assume that the platform displays profiles to only one side of the market in the first period and, conditional on the realized like and dislike decisions, the platform decides which subset of profiles to show to the other side of the market in the second period. As a result, matches can only happen sequentially in the second period. This setting is similar to that studied in the recent two-sided assortment optimization literature (Ashlagi et al.

2022, Torrico et al. 2021, Aouad and Saban 2022), where customers first select a supplier from their assortment and, later, the suppliers observe all the customers that chose them and decide whom to serve. Our model departs from this literature in two key aspects. First, we assume that users can like as many profiles as they want and, consequently, they can potentially match with multiple users on the other side. Second, we assume that like probabilities are independent of the subset of profiles displayed, while these papers consider an underlying choice model (e.g., MNL) to compute them. As mentioned above, this independence assumption greatly simplifies the analysis, and it is practical given that assortments minimally affect users' like probabilities (Rios et al. 2023).

Without loss of generality, we assume that interactions can only be initiated by agents in I. Hence, in the first period, the platform must choose a subset of profiles for each user in I, i.e., $\mathbf{x}^1 = \{x^1_{i,j}\}_{i\in I,\, j\in\mathcal{P}^1_i}$, that satisfies the constraints defined by the platform, namely, that $\sum_{j\in\mathcal{P}^1_i} x^1_{i,j} \le K_i$ for each $i\in I$. Hence, the feasible region for the first-period decisions can be formulated as:

$$R^1 = \left\{ \mathbf{x}^1 \in \{0,1\}^{\vec{E}_I} : \sum_{j\in\mathcal{P}^1_i} x^1_{i,j} \le K_i,\ \forall i\in I \right\}. \tag{12}$$

Note that the only difference between R^1 and $R^{1,\Pi}$ in (9) is that the former involves $\mathbf{w}^1 = 0$ and $\vec{E}_I$ instead of $\vec{E} = \vec{E}_I \cup \vec{E}_J$, ensuring that only users in I can see profiles in the first period.

A natural approach to solve Problem 2 under feasible region (12) is to adapt commonly used algorithms in the online matching and assortment optimization literature. One such algorithm is the greedy policy, which provides a performance guarantee of $1/2$ for both the online matching problem (Karp et al. 1990) and the online two-sided assortment problem (Aouad and Saban 2022). In our setting, such a (local) greedy policy would select the subset of profiles that maximizes the expected number of matches obtained by each user in isolation (see Appendix B.1.1 for the formal definition). In Proposition 2, we show that this policy achieves a worst-case performance arbitrarily close to zero, as it does not account for the potential congestion that some users may cause on others.

PROPOSITION 2. *The worst-case approximation guarantee for the local greedy policy is $O(1/n)$, where n is the size of the market.*

An alternative approach is to find a maximum weight perfect matching in each period, where the weight of each edge is the probability of having a match between the users (see Appendix B.1.2 for the formal definition). Chen et al. (2009) and Jeloudar et al. (2021) consider a similar approach in the probing problem with and without commitment and show that it achieves a performance guarantee of $1/4$ and 0.43, respectively. Nevertheless, as we show in Proposition 3, this policy has also a worst-case performance arbitrarily close to zero. This result is not surprising, as the perfect matching policy does not exploit the information provided by the realized like decisions.

Proposition 3. *The worst-case approximation guarantee for the perfect matching policy is $O(1/n)$, where n is the size of the market*

Given that our model requires the realization of two random variables to generate a match, the approaches above perform poorly because of their *non-adaptive* nature. To address this challenge, we exploit the structural properties of $f(\mathbf{B})$ defined in (4) and the feasible region R^1. As discussed in Section 3.2, the function $f(\mathbf{B})$ that returns the maximum expected number of matches that can be achieved given a realized backlog $\mathbf{B}$ is monotone and submodular in $\mathbf{B}$. Moreover, it is easy to see that the feasible region R^1 in (12) is a partition matroid. Therefore, we can show a $(1-1/e)$-approximation guarantee by combining the continuous greedy algorithm introduced by Vondrák (2008) (for submodular maximization under matroid constraints) with the dependent randomized rounding algorithm by Gandhi et al. (2006). We formalize this in Proposition 4.

Proposition 4. *When Π is restricted to one-directional policies with sequential matches, there exists a feasible solution $\mathbf{x}^1$ of Problem 2 whose objective value is such that $\mathcal{M}^2(\mathbf{x}^1) \geq (1-1/e)\cdot OPT$.*

Remark 2. Note that the worst-case performance of the local greedy and perfect matching policies also holds for the other platform designs. Hence, in the remainder of this section, we focus on devising performance guarantees for algorithms exploiting the submodularity of the objective function and, in Section 4.5, we provide a guarantee for DH that holds for all platform designs.

4.2. Two-directional Interactions and Sequential Matches

As in the previous case, we assume that the platform only allows sequential matches. However, we now consider the case where both sides of the market can initiate an interaction in the first period. This assumption holds in platforms that dynamically compute the profiles to show within each day based on their users' most recent evaluations. For instance, Tinder describes their algorithm as "[...] a dynamic system that continuously factors in how you're engaging with others on Tinder through Likes, Nopes, and what's on members' profiles [...]".[8] Tinder, and other platforms, have additional features that promote sequential matches by enhancing profile visibility and allowing users to signal a heightened level of interest, such as Tinder's "Super Like" or Bumble's "SuperSwipe".

To capture this platform design, we now assume that the first period decisions, represented by $\mathbf{x}^1 \in \{0,1\}^{\vec{E}}$, include arcs in both directions, i.e., $\vec{E} = \vec{E}_I \cup \vec{E}_J$. Then, the feasible region for the first-period decisions can be characterized as:

$$R^1 = \left\{ \mathbf{x}^1 \in \{0,1\}^{\vec{E}} : \sum_{\ell' \in \mathcal{P}_\ell^1} x^1_{\ell,\ell'} \leq K_\ell, \ \forall \ell \in I \cup J, \text{ and } x^1_{\ell,\ell'} + x^1_{\ell',\ell} \leq 1, \ \forall \ell \in I \cup J, \ \ell' \in \mathcal{P}_\ell^1 \right\}. \quad (13)$$

[8] You can find more details on Tinder's website.

The first family of constraints ensures our cardinality requirements, while the second ensures that no pair of users see each other in the first period, preventing non-sequential matches. Note that we set $\mathbf{w}^1 = 0$ since non-sequential matches are not allowed and, consequently, $\mathcal{M}^1 \equiv 0$.

Following a similar strategy as for the one-directional case, we can exploit the submodularity of the objective function and the structure of R^1 to find an approximation guarantee. Specifically, the feasible region R^1 corresponds to the intersection of two matroids, so we can use Fisher et al. (1978) to devise a feasible solution with an approximation factor of 1/3, or we can apply the local search algorithm in (Lee et al. 2009) for an improved guarantee of $1/(2+\epsilon)$[9]. We formalize this result in Proposition 5.

PROPOSITION 5. *When* Π *is restricted to two-directional policies with sequential matches, there exists a feasible solution* $\mathbf{x}^1$ *of Problem 2 whose objective value is such that* $\mathcal{M}^2(\mathbf{x}^1) \geq \frac{1}{2+\epsilon} \cdot OPT$*, for any* $\epsilon > 0$.

REMARK 3. Note that the approximation factor worsens relative to Proposition 4 because the family of constraints that prevents non-sequential matches correlates the decisions for each pair of users. An alternative approach would be to remove these constraints and penalize non-sequential matches. However, this approach may affect the submodularity of f in the second period. Another possibility is to consider the approach used in Proposition 4. However, the correlation between the decision variables prevents from using the dependent randomized rounding for each user.

4.3. One-directional Interactions and Non-Sequential Matches

So far, we have assumed that matches can only happen sequentially. However, many platforms have recently launched games/features that involve non-sequential matches. For instance, in 2021, Tinder introduced the "Hot Takes" game, where users who want to participate must log in within some time window (e.g., from 6 pm to midnight) and, after answering some questions, they are paired with another user and allowed to start a chat conversation before they both like each other and form a match. Other platforms, such as Filteroff and The League, have implemented similar speed-dating formats, allowing users to video chat for some time (e.g., 3 minutes) and giving them a chance to form a match and extend their time. Finally, platforms like Once provide users with only one match per day and, consequently, must ensure that users see each other in the same period. Thus, some platforms may consider generating non-sequential matches as part of their design.

As in Section 4.1, we assume (without loss of generality) that interactions are one-directional with I as the initiating side. However, we now allow the platform to select pairs of users that will simultaneously see each other in the first period and, thus, can potentially generate a non-sequential

[9] Both methods were originally designed for submodular maximization under the intersection of matroids.

match. To accomplish this, we use the variables $\mathbf{w}^1 \in \{0,1\}^E$ defined over the set of undirected edges between I and J, and we also set $x^1_{\ell,\ell'} = 0$ for all $\ell \in J$, $\ell' \in \mathcal{P}^1_\ell$ since interactions are only initiated by users in I, i.e., $\mathbf{x}^1 \in \{0,1\}^{\vec{E}_I}$. Then, the feasible region for the first-period decisions becomes:

$$R^1 = \Bigg\{ \mathbf{x}^1 \in \{0,1\}^{\vec{E}_I},\ \mathbf{w}^1 \in \{0,1\}^E : x^1_{i,j} + w^1_e \leq 1,\ \forall i \in I,\ j \in \mathcal{P}^1_i,\ e = \{i,j\} \\ \sum_{j \in \mathcal{P}^1_i} x^1_{i,j} + \sum_{e \in E: i \in e} w^1_e \leq K_i,\ \forall i \in I, \\ \sum_{e \in E: j \in e} w^1_{i,j} \leq K_j,\ \forall j \in J. \Bigg\} \tag{14}$$

The first family of constraints guarantees that no profile targets both a sequential (i.e., $x^1_{i,j} = 1$) and a non-sequential match (i.e., $w^1_e = 1$), while the second and third families of constraints ensure that the subsets of profiles to display satisfy the cardinality requirements for sides I and J, respectively. Note that each user $j \in J$ only sees profiles involving users $i \in I$ for which $w^1_e = 1$ and, thus, users in J cannot initiate sequential matches.

Similar to the previous case, we can show that the feasible region R^1 corresponds to the intersection of a partition with a laminar matroid. Therefore, since the objective function is still monotone and submodular, we can use the local search algorithm proposed by Lee et al. (2009) to find a feasible solution with an approximation factor of $1/(2+\epsilon)$, as we formalize in Proposition 6.

Proposition 6. *When* Π *is restricted to one-directional policies with non-sequential matches in the first period, there exists a feasible solution* $(\mathbf{x}^1, \mathbf{w}^1)$ *of Problem 2 whose objective value is such that* $\mathcal{M}^1(\mathbf{x}^1,\mathbf{w}^1) + \mathcal{M}^2(\mathbf{x}^1,\mathbf{w}^1) \geq \frac{1}{2+\epsilon} \cdot OPT$*, for any* $\epsilon > 0$.

4.4. Two-directional Interactions and Non-Sequential Matches

The last platform design we consider is when interactions are two-directional and matches can happen sequentially (in the second period) or non-sequentially (in the first period). This setting captures the one used by many platforms (including our industry partner) that daily "pre-compute" the subsets to display the next day and allow some users to see each other in the same period, especially those with a limited pool of potential partners.

As discussed in Section 3.2, the feasible region for the first-period decisions can be represented by (9), which corresponds to the intersection of three partition matroids. Hence, we can use the local search algorithm in Lee et al. (2009) to find a feasible solution with an approximation factor of $1/(3+\epsilon)$, as we formalize in Proposition 7.

Proposition 7. *When* Π *is restricted to two-directional policies with non-sequential matches in the first period, there exists a feasible solution* $(\mathbf{x}^1, \mathbf{w}^1)$ *of Problem 2 whose objective value is such that* $\mathcal{M}^1(\mathbf{x}^1,\mathbf{w}^1) + \mathcal{M}^2(\mathbf{x}^1,\mathbf{w}^1) \geq \frac{1}{3+\epsilon} \cdot OPT$*, for any* $\epsilon > 0$.

4.5. Improved Guarantee for All Platform Designs using DH

The results so far show that exploiting the feasible region's structure and the objective function's monotonicity and submodularity can lead to feasible solutions with constant-factor approximation guarantees. However, these guarantees are specific to each platform design and worsen as we enable more features. In this section, we show that DH achieves a $(1-1/e)$-approximation guarantee for Problem 2 for any platform design, as we formalize in Theorem 1.

THEOREM 1. *DH achieves a $(1-1/e)$-approximation guarantee for Problem 2 for any platform design.*

The proof of Theorem 1 (in Appendix B.3) relies on a novel connection between the reformulation (11) of the relaxation used by DH and a different relaxation that aims to find the distribution of backlogs that maximizes the number of matches while satisfying constraints on the marginal probabilities. Specifically, given the optimal solution $(\mathbf{x}^\star, \mathbf{w}^\star, \mathbf{y}^\star)$ of (11), we first observe that $\mathbf{y}^\star$ can be recovered by solving the following problem:

$$\begin{aligned} F(\mathbf{x}^\star, \mathbf{w}^\star) := \max \quad & \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}^1_\ell} \phi^2_{\ell,\ell'} \cdot y_{\ell,\ell'} && && (15) \\ s.t. \quad & y_{\ell,\ell'} \le \phi^1_{\ell',\ell} \cdot x^\star_{\ell',\ell}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^1_\ell \\ & \sum_{\ell' \in \mathcal{P}^1_\ell} y_{\ell,\ell'} \le K_\ell, && \forall \ell \in I \cup J \\ & y_{\ell,\ell'} \ge 0, && \forall \ell \in I \cup J, \ell' \in \mathcal{P}^1_\ell. \end{aligned}$$

We then show that a second upper bound for the expected number of matches in the second period can be obtained for any vectors of first-period decisions $(\mathbf{x}, \mathbf{w}) \in \{0,1\}^{\vec{E}} \times \{0,1\}^{E}$ by solving:

$$\begin{aligned} G(\mathbf{x}, \mathbf{w}) := \max \quad & \sum_{\ell \in I \cup J} \sum_{B \subseteq \mathcal{P}^1_\ell} f_\ell(B) \cdot \lambda_{\ell,B} && && (16) \\ s.t. \quad & \sum_{B \subseteq \mathcal{P}^1_\ell} \lambda_{\ell,B} = 1 && \forall \ell \in I \cup J \\ & \sum_{B \subseteq \mathcal{P}^1_\ell : \ell' \in B_\ell} \lambda_{\ell,B} = \phi^1_{\ell,\ell'} \cdot x_{\ell,\ell'}, && \forall \ell \in I \cup J, \ell' \in \mathcal{P}^1_\ell \\ & \lambda_{\ell,B} \ge 0, && \forall \ell \in I \cup J,\ B \subseteq \mathcal{P}^1_\ell. \end{aligned}$$

In this formulation, which we refer to as the *distribution problem*, the decision variables $\lambda_{\ell,B}$ represent the probability that the set $B \subseteq \mathcal{P}^1_\ell$ is the backlog of user ℓ, and $f_\ell(B)$ is the maximum expected number of matches that user ℓ can achieve given a backlog B (as defined in (4)). The first and last families of constraints ensure that the variables $\{\lambda_{\ell,B}\}_{B \subseteq \mathcal{P}^1_\ell}$ correctly define a probability distribution, while the second family of constraints guarantees that the marginal probabilities are consistent with the first-period decisions.

Based on these two formulations, the next step in the proof is to show that $G(\mathbf{x}^\star, \mathbf{w}^\star)$ is at least $F(\mathbf{x}^\star, \mathbf{w}^\star)$ and that a feasible solution for the latter can be constructed as $\lambda_{\ell,B} = \mathbb{P}_{\mathbf{x}^\star_\ell}(B)$ (as defined in 2 but only considering ℓ). Noticing that the expected number of matches in the second period produced by this feasible solution coincides with $\mathcal{M}^2(\mathbf{x}^\star, \mathbf{w}^\star)$, we combine strong-duality with the correlation gap in Agrawal et al. (2010) to show that

$$\mathcal{M}^2(\mathbf{x}^\star, \mathbf{w}^\star) \geq (1 - 1/e) \cdot G(\mathbf{x}^\star, \mathbf{w}^\star) \geq (1 - 1/e) \cdot F(\mathbf{x}^\star, \mathbf{w}^\star).$$

Finally, recalling that $\mathcal{M}^1(\mathbf{x}^\star, \mathbf{w}^\star) + F(\mathbf{x}^\star, \mathbf{w}^\star) \geq \text{OPT}$ (since (11) provides an upper bound of OPT and the left-hand side coincides with its optimal solution), we conclude that

$$\mathcal{M}(\mathbf{x}^\star, \mathbf{w}^\star) = \mathcal{M}^1(\mathbf{x}^\star, \mathbf{w}^\star) + \mathcal{M}^2(\mathbf{x}^\star, \mathbf{w}^\star) \geq (1 - 1/e) \cdot \text{OPT}.$$

5. Experiments

In this section, we evaluate the performance of DH for different platform designs and compare it with the other benchmarks analyzed in Section 4.

5.1. Data

We use a dataset obtained from our industry partner to perform our experiments. This dataset includes all heterosexual users from Houston, TX that logged in between February 14 and August 14, 2020, and includes all the observable characteristics displayed in their profiles for each user in the sample, namely, their age, height, location, race, and religion. It also includes an attractiveness score—or simply score—that depends on the number of likes received and evaluations received in the past.[10] Finally, the dataset includes all the profiles that each user evaluated between February 14 and August 14, 2020, including the decisions made (like or dislike), which other profiles were displayed, and relevant timestamps. As a result, we have a panel of observations, and we can fully characterize each profile evaluated by each user in the sample.

Using this dataset, for any pair of users $\ell \in I \cup J$ and $\ell' \in \mathcal{P}^1_\ell$, we compute the probability that ℓ likes ℓ' using the panel regression model described in detail in Appendix C.1. Since we assume that like probabilities are independent across periods (i.e., we assume that there is no effect of the history on the like behavior of users),[11] we estimate the like probabilities considering a logit model with fixed effects at the user level. In addition, we control for all the observable characteristics available in the data, namely, the characteristics of the profile evaluated and the interaction with those of the

[10] This score is measured on a scale from 0 to 10, where 10 represents the most attractive profiles and 0 the least attractive ones.

[11] Rios et al. (2023) find that the number of matches obtained in the recent past has a significant negative effect on users' future like/dislike decisions. However, the authors show that the primary source of improvement comes from taking into account the market's two-sidedness and choosing better assortments. Hence, we decided to focus on the latter and simplify the estimation of probabilities.

user evaluating. In Table 2 (see Appendix C.1), we report the estimation results. Then, by using the estimated coefficients and users' observable characteristics, we predict the probabilities $\phi_{\ell,\ell'}$ for all $\ell \in I \cup J$ and $\ell' \in \mathcal{P}_\ell^1$. We perform our simulations considering a random sample of the dataset described above to reduce the computational time.[12]

5.2. Benchmarks

To assess the performance of DH under different platform designs, we compare it with the benchmarks studied in Section 4:[13]

1. Local Greedy: for each user, select the subset of profiles that maximizes their expected number of matches, i.e., $S_\ell^t = \arg\max_{S \subseteq \mathcal{P}_\ell \setminus \cup_{\tau=1}^{t-1} S_\ell^\tau : |S| \leq K_\ell} \left\{ \sum_{\ell' \in S} \phi_{\ell,\ell'} \cdot \left(\mathbb{1}_{\{\ell' \in B_\ell^t\}} + \phi_{\ell'\ell} \cdot \mathbb{1}_{\{\ell' \notin B_\ell^t\}} \right) \right\}$.
2. Perfect Matching (PM): in each period $t \in [T]$, solve the perfect matching problem (including possible initial backlogs). We formalize this method in Appendix B.1.2.
3. Global Greedy: for each user, select the subset of profiles based on Algorithm 5 (using the standard greedy method for submodular maximization (Fisher et al. 1978) as ALG). We consider two special cases of this algorithm: (i) None and (ii) First. In the former, we restrict to sequential matches when making the first-period decisions (i.e., we use Algorithm 5 directly). In the latter, we extend this algorithm to allow non-sequential matches in the first period.

Moreover, we consider three variants of DH that differ on whether we allow for non-sequential matches when solving Problem (10) to obtain the first period decisions: (i) None forces $w_e^t = 0 \ \forall t \in \{1,2\}$, $e \in E$, i.e., excludes non-sequential matches in both periods; (ii) First allows non-sequential matches in the first period, i.e., forces $w_e^2 = 0$ for all $e \in E$; and (iii) Both allows non-sequential matches in both periods. Finally, we compare all these methods with the upper bound (UB) obtained from solving Problem (10).

5.3. Results

For each platform design and benchmark, we perform 100 simulations where, in each period, (i) we choose the profiles to display to each user considering $K_\ell = 3$ for all $\ell \in I \cup J$, (ii) we simulate the decisions of the users based on their like probabilities, and (iii) we update the state of the system before moving on to the next period. In all these simulations, we assume that each user starts with no backlog (i.e., $\mathcal{B}_\ell^1 = \emptyset$ for all $\ell \in I \cup J$).[14]

In Figure 1, we report the average number of matches generated by each benchmark for each platform design.

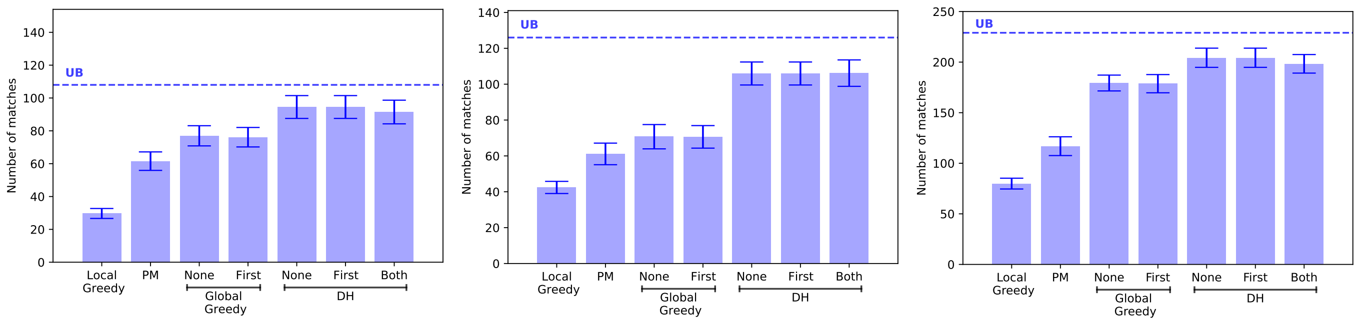
[12] In Table 3 in Appendix C.1, we report several summary statistics of the sample, including the number of users, their average score, the average number of potentials available, their average backlog size, and their average like probabilities.

[13] We also tested other benchmarks such as Naive, Random, and our partner's algorithm. However, since the results reported here significantly outperform these other benchmarks, we decided to omit them and focus on the results of the algorithms proposed above.

[14] The results are similar if we assume an initial backlog.

Figure 1 Matches per Benchmark by Platform Design

(a) Men Initiating (b) Women Initiating (c) Two-Directional

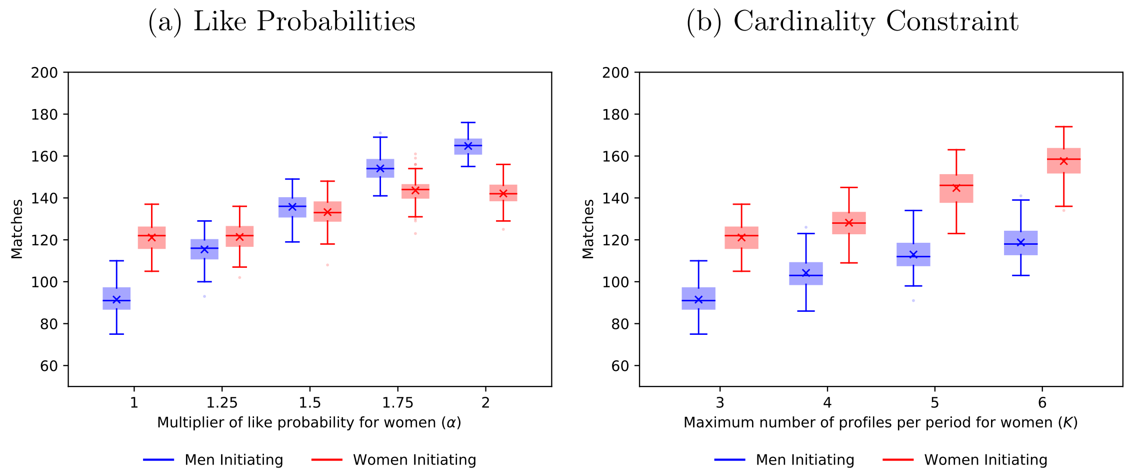


5.3.1. One-Directional Interactions. Figures 1a and 1b report the results for one-directional platform designs, with men and women initiating interactions, respectively. In both cases, we observe that the performance of PM and Local Greedy is considerably better than their worst-case performance. Moreover, we observe that Global Greedy outperforms these benchmarks, but allowing non-sequential matches in the first period leads to no improvement. Finally, we find that DH and its variants significantly outperform all the other benchmarks and that allowing for non-sequential matches plays almost no role, consistent with the results for Global Greedy.

More interestingly, comparing Figures 1a and 1b, we observe that the initiating side matters. In particular, we find that DH leads to significantly more matches when women initiate interactions. As theoretically suggested by (Kanoria and Saban 2021, Torrico et al. 2021, Shi 2022a), one potential reason is that women have a significantly lower like probability (20.3% vs. 53.6%; see Table 3 in Appendix C.1). To test this conjecture, in Figure 2, we compare the results for men and women initiating interactions varying the like probabilities for women (Figure 2a) and the maximum number of profiles that women can see in each period (Figure 2b). Specifically, for each women j, we either multiply their like probabilities $\phi^t_{j,i}$ by a constant $\alpha \in \{1, 1.25, 1.5, 1.75, 2\}$ for each $i \in \mathcal{P}^1_j$ and $t \in \{1,2\}$ or we modify the right-hand side of the cardinality constraint for women $K \in \{3,\dots,6\}$ while keeping that for men fixed at three, and repeat the simulation procedure described above.

From Figure 2a we observe that the number of matches is strictly increasing in α when men move first, while it is weakly increasing when women do so. Moreover, there is a threshold $\bar{\alpha}$ (approximately at 1.5) such that the number of matches obtained when men move first is strictly larger than the number of matches obtained if $\alpha > \bar{\alpha}$, while the opposite holds when $\alpha < \bar{\alpha}$. Note that this switching point is below the ratio of the average like probabilities for men and women (equal to $2.115 = 0.571/0.227$), suggesting that the difference in selectivity cannot fully explain the effect of the initiating side. Indeed, we observe that the threshold is closer to the ratio of the expected backlog size produced by men and women, i.e., the ratio of like probabilities times the

Figure 2 **Sensitivity to Problem Parameters for Women**



ratio of the size of each side of the market, which is equal to $1.381 = 2.115 \cdot 113/173$.[15] Moreover, from Figure 2b, we observe that the number of matches is strictly increasing in the cardinality of the set of profiles that women can see in each period, and the gap on the number of matches obtained by each initiating side is relatively stable. These results indicate that the balance between size and selectivity is the primary driver of the differences in the performance by initiating side.

Overall, the results in this section suggest that the initiating side matters and that the performance of the algorithms is sensitive to the differences in the expected backlog per profile displayed on each side of the market. Hence, platforms that use a one-directional design and aim to maximize the expected number of matches should consider as initiator the side leading to the smallest expected backlog per profile displayed, balancing size and selectivity.

5.3.2. Two-Directional Interactions. In Figure 1c, we report the results for a two-directional platform design. First, consistent with the results for one-directional interactions, we find that DH outperforms all the benchmarks considered. Second, allowing non-sequential matches in the first period leads to no improvement. This result, which also holds in the one-directional case, aligns with the results in Rios et al. (2023), which show that most matches occur sequentially. Finally, comparing the results in Figure 1c with those in Figure 1b, we observe that the one-directional design with women initiating leads to at least 50% of the matches obtained with two-directional interactions.

[15] This holds since the number of profiles each user on each side of the market sees is the same (equal to three). If this were not the case, the ratio of backlogs would be equal to the product of the ratios of (i) like probabilities, (ii) side sizes, and (iii) average profiles that each user sees on each side.

6. Extensions

In this section, we analyze several extensions of our model. In Section 6.1, we analyze the problem with a general time horizon. In Section 6.2, we extend the model in Section 4.3 to allow for non-sequential matches in the second period. We defer the proofs to Appendix D.

6.1. Multiple Periods: One-Directional Interactions and Sequential Matches

So far, we have focused on a two-period version of the problem. As previously discussed, this assumption is borrowed from practice since most dating platforms either (i) precompute the subsets of profiles to show to each user daily or (ii) compute them dynamically over time based on the state of the system, namely, the set of potentials and the backlogs. Hence, considering a longer time horizon in the computation does not lead to substantial benefits. Nevertheless, whether the guarantees discussed in Section 4 extend to multiple periods is an interesting theoretical question.

In this section, we show that we can extend DH (specifically, the relaxation in (11)) to obtain guarantees for any time horizon when we restrict to either *semi-adaptive* or *adaptive* policies (as defined next), provided that we exclude non-sequential matches in the last period.

Definition 1. We say that a policy is *semi-adaptive* if (i) it non-adaptively selects profiles for the initiating side, i.e., when selecting profiles to start a sequential or a non-sequential interaction; and (ii) it adaptively selects profiles for the responding side, i.e., when selecting profiles from the backlog. In contrast, we say that a policy is *adaptive* if it selects the profiles to display in both sides (initiating and responding) using information from previous stages.

To simplify the exposition and analysis, we focus on the case with time-homogeneous probabilities, i.e., $\phi^t_{\ell,\ell'} = \phi_{\ell,\ell'}$ for all $t \in [T]$, $\ell \in I \cup J$, $\ell' \in \mathcal{P}^1_\ell$. Then, consider the following extension of the relaxation in (11) used by DH:

$$
\begin{aligned}
\max \quad & \sum_{e\in E} w_e \cdot \beta_e + \sum_{\ell \in I\cup J} \sum_{\ell' \in \mathcal{P}^1_\ell} y_{\ell,\ell'} \cdot \phi_{\ell,\ell'} \\
\text{s.t.} \quad & y_{\ell,\ell'} \le x_{\ell',\ell} \cdot \phi^1_{\ell',\ell}, && \forall \ell \in I\cup J,\ \ell' \in \mathcal{P}^1_\ell, \\
& x_{\ell,\ell'} + x_{\ell',\ell} + w_e \le 1, && \forall \ell \in I\cup J,\ \ell' \in \mathcal{P}^1_\ell,\ e = \{\ell,\ell'\}, \\
& \sum_{\ell' \in \mathcal{P}^1_\ell} x_{\ell,\ell'} + \sum_{e\in E:\, \ell\in e} w_e + \sum_{\ell'\in\mathcal{P}^1_\ell} y_{\ell,\ell'} \le K_\ell \cdot T, && \forall \ell \in I \cup J, \\
& \mathbf{x} \in \{0,1\}^{\vec{E}},\ \mathbf{w} \in \{0,1\}^{E},\ \mathbf{y} \in [0,1]^{\vec{E}}.
\end{aligned}
\tag{17}
$$

The decision variables $x_{\ell,\ell'}$ indicate whether or not user ℓ sees ℓ''s profile throughout the T-period horizon, initiating their interaction. The decision variables $y_{\ell,\ell'}$ represent the probability that ℓ sees ℓ' as the respondent/follower of the interaction, which happens only conditional on ℓ' liking ℓ in the first place. Finally, the decision variables w_e with $e = \{\ell,\ell'\}$ capture whether ℓ and ℓ' both see each other in the same period. Note that when $T = 2$ the third constraint in (17) is equivalent to the aggregation of the third and fourth constraints in (11). As we show in Lemma 3, (17) provides an upper bound for any semi-adaptive policy.

LEMMA 3. *Problem* (17) *is an upper bound for Problem 1 when* Π *is restricted to semi-adaptive policies.*

Based on this result, we can design a T-period version of DH, formalized in Algorithm 2, that achieves a $1-1/e$ approximation factor for Problem 1 when Π is restricted to semi-adaptive policies, as we show in Theorem 2. Algorithm 2 starts by solving Problem (17), whose optimal solution $(\mathbf{x}^\star, \mathbf{w}^\star, \mathbf{y}^\star)$ provides an upper bound for the maximum expected number of matches that any semi-adaptive policy can obtain in T periods (by Lemma 3). Then, in each period t, the algorithm iterates over users and decides the profiles to display based on $\mathbf{x}^\star$ and $\mathbf{w}^\star$. The algorithm first exhausts the profiles initiating an interaction, i.e., $x_{\ell,\ell'} = 1$ (Steps 6-8). If there is still space left in S_ℓ^t, the algorithm displays profiles that are part of non-sequential interactions (i.e., $w_e^\star = 1$) as long as the cardinality constraint is not binding for both users (Steps 9-12). If the cardinality constraint is still non-binding, the algorithm computes the optimal subset of profiles to display to ℓ based on the profiles in their backlog (Step 14). Finally, the algorithm updates the potentials and backlogs based on the observed like/dislike decisions (Step 17).

THEOREM 2. *Algorithm 2 achieves a* $(1-1/e)$*-approximation guarantee for Problem 1 for any platform design and when* Π *is restricted to semi-adaptive policies.*

Note that we could obtain a similar guarantee for the best adaptive policy by solving the linear relaxation of Problem (17), as this would provide an upper bound for Problem 1 when restricted to adaptive policies. However, the analysis of this case is more complex since the optimal solution may be fractional and, thus, deriving the guarantee would require rounding techniques. Specifically, the second set of constraints creates a negative correlation among the variables $(\mathbf{x}, \mathbf{w}, \mathbf{y})$, and it is unclear how to design a randomized rounding method with meaningful guarantees under this setting. However, this is possible when policies are restricted to one-directional interactions and sequential matches, as we show in Theorem 3.

THEOREM 3. *There exists a semi-adaptive policy that achieves a* $(1-1/e)$*-approximation guarantee for Problem 1 when* Π *allows for adaptive policies, but it is restricted to one-directional interactions and sequential matches.*

REMARK 4. The guarantee in Theorem 3 is with respect to the best possible adaptive policy, i.e., the solution derived from solving the DP presented in Appendix A.1. Our non-adaptive policy consists of solving an adaptation of Formulation (17) with the integrality constraints relaxed. Once a fractional solution is obtained, we construct a feasible solution with the dependent randomized rounding method proposed in (Gandhi et al. 2006). Then, we feed this solution to the analogous of Algorithm 2 to this setting.

Algorithm 2 T-periods DH

Input: An instance of Problem 1.

Output: A subset of profiles to display in each period.

```
Solve (17) and let (x*, w*, y*) be the optimal solution.
Define X_ℓ = {ℓ' ∈ P^1_ℓ : x*_{ℓ,ℓ'} = 1} for each ℓ ∈ I ∪ J and W_ℓ = {e ∈ E : ℓ ∈ e, w*_e = 1}.
for t ∈ [T] do
    Initialize S^t_ℓ = ∅ for each ℓ ∈ I ∪ J.
    for ℓ ∈ I ∪ J do
        while |S^t_ℓ| < K_ℓ and X_ℓ ≠ ∅ do
            Let ℓ' be φ_{ℓ',ℓ} ∈ arg max{φ_{a,ℓ} : a ∈ X_ℓ} breaking ties arbitrarily.
            S^t_ℓ = S^t_ℓ ∪ {ℓ'}, X_ℓ = X_ℓ \ {ℓ'}.
        while |S^t_ℓ| < K_ℓ and W_ℓ ≠ ∅ do
            Choose any ℓ' ∈ W_ℓ.
            if |S^t_ℓ'| < K_ℓ' then
                S^t_ℓ = S^t_ℓ ∪ {ℓ'}, S^t_ℓ' = S^t_ℓ' ∪ {ℓ}, W_ℓ = W_ℓ \ {e}, W_ℓ' = W_ℓ' \ {e}
        if |S^t_ℓ| < K_ℓ then
            Solve f^t_ℓ(B^t_ℓ) defined as in (4) considering K_ℓ − |S^t_ℓ| as the right-hand side of the
            cardinality constraint, where B^t_ℓ is ℓ's backlog at the beginning of period t. Let z be
            the optimal solution.
            S^t_ℓ = S^t_ℓ ∪ {ℓ' ∈ B^t_ℓ : z_{ℓ,ℓ'} = 1}.
    For each ℓ ∈ I ∪ J, display the subset of profiles S^t_ℓ.
    Observe like/dislike decisions. Update the sets of potentials and the backlogs following (1).
```

6.2. One-Directional Interactions and Non-Sequential Matches in Both Periods

As we show in Proposition 1, the function that returns the optimal number of matches attained in the second period (from sequential and non-sequential interactions) given a family of backlogs $\mathbf{B}$ is not submodular when we enable non-sequential matches in the second period. Thus, the analysis is significantly more complex as we can no longer rely on submodular optimization techniques and the correlation gap. Nevertheless, we can still provide a performance guarantee for the case when non-sequential matches are allowed in both periods and when like probabilities on the initiating side are small, as we formalize in Theorem 4.

Theorem 4. *Consider the case with one-directional sequential matches starting from side I and non-sequential matches in both periods. Suppose that probabilities are time-independent, i.e., $\phi^1_{\ell\ell'} = \phi^2_{\ell\ell'} = \phi_{\ell\ell'}$ and that $\phi_{ij} \le 1/n$ for any $i \in I, j \in J$, where $n = |I|$. Denote by OPT_1 the optimal solution when Π is restricted to policies with non-sequential matches only in the first period and by OPT_2*

the optimal value when Π *allows for policies to implement non-sequential matches in both periods. Then, we have*

$$OPT_1 \geq \left(\frac{1}{2e} - o(1)\right) \cdot OPT_2.$$

More importantly, any γ*-approximation algorithm for* OPT_1 *leads to a* $\gamma\left(\frac{1}{2e} - o(1)\right)$ *approximation for* OPT_2*.*

Theorem 4 implies that allowing non-sequential matches in the second period does not arbitrarily improve the optimal expected number of matches when the initiating side is sufficiently picky and the market is sufficiently large. In other words, the majority of matches come from sequential interactions. As a result, we conjecture that the guarantees in Section 4 may extend to the more general case. We emphasize that analyses on large markets with small probabilities are common in the literature; see Mehta and Panigrahi (2012), Goyal and Udwani (2023) for some examples from the online matching literature.

7. Conclusions

In this paper, we study the design and optimization of curated dating platforms, focusing on two key design features: (i) the sequence of interactions and (ii) the timing of matches.

Our theoretical analysis demonstrates that the Dating Heuristic (DH) introduced by Rios et al. (2023) provides a simple and robust framework to solve the problem, achieving a $1-1/e$ constant-factor approximations for all platform designs, in sharp contrast to other common approaches. Moreover, our empirical analysis, which uses real data from our industry partner, confirms the theoretical results, generates new insights into the performance of the DH in practice, and provides a series of guidelines for designing these platforms. First, we show that the performance of the DH is robust to the different platform designs, outperforming all the other benchmarks considered. Second, our results confirm that avoiding non-sequential matches does not significantly affect the number of matches the platform can generate, regardless of the platform design. More interestingly, our simulation results demonstrate that platforms using a one-directional design should initiate interactions with the side that leads to the smallest expected backlog per profile displayed, balancing size and selectivity. Lastly, we show that a one-directional design can lead to at least half of the matches as those obtained with a two-directional design.

Overall, our work contributes to advancing the understanding of curated dating platforms' optimization and provides actionable insights for improving their design depending on their desired goals and functionalities. Moreover, our model can be adapted to capture a general market (i.e., not only heterosexual individuals) by appropriately duplicating nodes on each side and imposing constraints on the set of potentials of each user. As we introduce more constraints to avoid conflict between duplicated nodes, we expect DH to perform well in practice but likely with lower worst-case provable guarantees. Finally, we remark that our insights may be relevant to other two-sided assortment problems, such as those faced by freelancing, ride-sharing, and accommodation platforms.

References

Adamczyk M (2011) Improved analysis of the greedy algorithm for stochastic matching. *Information Processing Letters* 111(15):731–737.

Agrawal S, Ding Y, Saberi A, Ye Y (2010) Correlation robust stochastic optimization. *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, 1087–1096 (SIAM).

Aouad A, Saban D (2022) Online assortment optimization for two-sided matching platforms. *Management Science* 0(0).

Arnosti N, Shi P (2020) Design of lotteries and wait-lists for affordable housing allocation. *Management Science* 66(6):2291–2307.

Ashlagi I, Krishnaswamy A, Makhijani R, Saban D, Shiragur K (2022) Technical note - assortment planning for two-sided sequential matching markets. *Operations Research* 70(5):2784–2803.

Bansal N, Gupta A, Li J, Mestre J, Nagarajan V, Rudra A (2012) When lp is the cure for your matching woes: Improved bounds for stochastic matchings. *Algorithmica* 63:733–762.

Berbeglia G, Joret G (2015) Assortment Optimisation Under a General Discrete Choice Model: A Tight Analysis of Revenue-Ordered Assortments.

Besbes O, Castro F, Lobel I (2021) Surge pricing and its spatial supply response. *Management Science* 67(3):1350–1367.

Besbes O, Fonseca Y, Lobel I, Zheng F (2023) Signaling competition in two-sided markets.

Blanchet J, Gallego G, Goyal V (2016) A Markov chain approximation to choice modeling. *Operations Research* 64(4):886–905.

Brubach B, Grammel N, Ma W, Srinivasan A (2021) Improved guarantees for offline stochastic matching via new ordered contention resolution schemes. *Advances in Neural Information Processing Systems* 34:27184–27195.

Caro F, Gallien J (2007) Dynamic assortment with demand learning for seasonal consumer goods. *Management Science* 53(2):276–292.

Celdir ME, Cho SH, Hwang EH (2024) Popularity bias in online dating platforms: Theory and empirical evidence. *Manufacturing & Service Operations Management* 26(2):537–553.

Chen N, Immorlica N, Karlin AR, Mahdian M, Rudra A (2009) Approximating matches made in heaven. *Automata, Languages and Programming*, 266–278 (Springer Berlin Heidelberg).

Costello KP, Tetali P, Tripathi P (2012) Stochastic matching with commitment. *Automata, Languages, and Programming: 39th International Colloquium, ICALP 2012, Warwick, UK, July 9-13, 2012, Proceedings, Part I 39*, 822–833 (Springer).

Cui T, Hamilton M (2024) Pricing strategies for online dating platforms.

Davis JM, Gallego G, Topaloglu H (2014) Assortment Optimization Under Variants of the Nested Logit Model. *Operations Research* 62(2):250–273.

Erdős P, Rényi A (1964) On random matrices. *Publ. Math. Inst. Hungarian Academy of Sciences* 8:455–461.

Erdős P, Rényi A (1968) On random matrices ii. *Studia Sci. Math. Hungar* 3:459–464.

Farias VF, Jagabathula S, Shah D (2013) A Nonparametric Approach to Modeling Choice with Limited Data. *Management Science* 59(2):305–322.

Fisher ML, Nemhauser GL, Wolsey L (1978) An analysis of approximations for maximizing submodular set functions–ii. *Polyhedral combinatorics* .

Gamlath B, Kale S, Svensson O (2019) Beating greedy for stochastic bipartite matching. *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2841–2854 (SIAM).

Gandhi R, Khuller S, Parthasarathy S, Srinivasan A (2006) Dependent rounding and its applications to approximation algorithms. *Journal of the ACM (JACM)* 53(3):324–360.

Golrezaei N, Nazerzadeh H, Rusmevichientong P (2014) Real-Time Optimization of Personalized Assortments. *Management Science* 60(6):1532–1551.

Goyal V, Udwani R (2023) Online matching with stochastic rewards: Optimal competitive ratio via path-based formulation. *Operations Research* 71(2):563–580.

Halaburda H, Piskorski MJ, Yildirim P (2018) Competing by Restricting Choice: The Case of Search Platforms. *Management Science* 64(8):3574–3594.

Hikima Y, Akagi Y, Kim H, Kohjima M, Kurashima T, Toda H (2021) Integrated optimization of bipartite matching and its stochastic behavior: New formulation and approximation algorithm via min-cost flow optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 3796–3805.

Immorlica N, Lucier B, Manshadi V, Wei A (2022) Designing approximately optimal search on matching platforms. *Management Science* Forthcoming.

Jeloudar M, Lo I, Pollner T, Saberi A (2021) Decentralized matching in a probabilistic environment. *Proceedings of the 22nd ACM Conference on Economics and Computation*, 635–653.

Kanoria Y, Saban D (2021) Facilitating the search for partners on matching platforms. *Management Science* 67(10):5990–6029.

Karp R, Vazirani U, Vazirani V (1990) An optimal algorithm for online bipartite matching. *Proceedings of the 22nd Annual ACM Symposium on Theory of Computinh.*

Kök A, Fisher M, Vaidyanathan R (2015) *Retail Supply Chain Management: Quantitative Models and Empirical Studies*, 175–236 (Springer US).

Lee J, Sviridenko M, Vondrák J (2009) Submodular maximization over multiple matroids via generalized exchange properties. *International Workshop on Approximation Algorithms for Combinatorial Optimization*, 244–257 (Springer).

Manshadi V, Rodilitz S (2022) Online policies for efficient volunteer crowdsourcing. *Management Science* 68(9):6572–6590.

Manshadi V, Rodilitz S, Saban D, Suresh A (2022) Online algorithms for matching platforms with multi-channel traffic.

Mehta A, Panigrahi D (2012) Online matching with stochastic rewards. *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, 728–737 (IEEE).

Rios I, Saban D, Zheng F (2023) Improving match rates in dating markets through assortment optimization. *Manufacturing & Service Operations Management* 25(4):1304–1323.

Rochet JC, Tirole J (2003) Two-Sided Markets. *Journal of the European Economic Association* 990–1029.

Rusmevichientong P, Shen ZJM, Shmoys DB (2010) Dynamic Assortment Optimization with a Multinomial Logit Choice Model and Capacity Constraint. *Operations Research* 58(6):1666–1680.

Rusmevichientong P, Shmoys D, Tong C, Topaloglu H (2014) Assortment optimization under the multinomial logit model with random choice parameters. *Production and Operations Management* 23(11):2023–2039.

Rusmevichientong P, Sumida M, Topaloglu H (2020) Dynamic assortment optimization for reusable products with random usage durations. *Management Science* 66(7):2820–2844.

Sauré D, Zeevi A (2013) Optimal Dynamic Assortment Planning with Demand Learning. *Manufacturing & Service Operations Management* 15(3):387–404.

Shi P (2022a) Optimal matchmaking strategy in two-sided marketplaces. *Management Science* 0(0):null.

Shi P (2022b) Optimal priority-based allocation mechanisms. *Management Science* 68(1):171–188.

Talluri K, van Ryzin G (2004) Revenue Management Under a General Discrete Choice Model of Consumer Behavior. *Management Science* 50(1):15–33.

Torrico A, Carvalho M, Lodi A (2021) Multi-agent assortment optimization in sequential matching markets.

Vondrák J (2008) Optimal approximation for the submodular welfare problem in the value oracle model. *Proceedings of the 40th Annual ACM Symposium on the Theory of Computing (STOC)*, 67–74.

Wang R, Sahin O (2018) The Impact of Consumer Search Cost on Assortment Planning and Pricing. *Management Science* 64(8):3649–3666.

Notation. In the remainder of the Appendix, with a slight abuse of notation, we write the family of backlogs $\mathbf{B}$ as a collection of directed edges, e.g., $(\ell,\ell') \in \mathbf{B}$ means that ℓ is in the backlog of ℓ', i.e., $\ell \in B_{\ell'}$. For a set function f we write $f(\mathbf{B} \cup e)$ for $f(\mathbf{B} \cup \{e\})$.

Appendix A: Appendix to Section 3

A.1. Dynamic programming formulation

Let $\mu^t_{\ell,\ell'} = 1$ if a match between users ℓ and ℓ' happens in period t, and $\mu^t_{\ell,\ell'} = 0$ otherwise. As discussed in Section 3, $\mu^t_{\ell,\ell'} = 1$ if and only if one of the following three (disjoint) events takes place: (i) $\left\{\Phi^t_{\ell,\ell'} = 1, \ell' \in \mathcal{B}^t_\ell\right\}$, (ii) $\{\Phi^t_{\ell'\ell} = 1, \ell \in \mathcal{B}^t_{\ell'}\}$, or (iii) $\left\{\Phi^t_{\ell,\ell'} = \Phi^t_{\ell'\ell} = 1\right\}$.

Given a set of potentials $\vec{\mathcal{P}} = \{\mathcal{P}_\ell\}_{\ell \in I \cup J}$ and a set of backlogs $\vec{\mathcal{B}} = \{\mathcal{B}_\ell\}_{\ell \in I \cup J}$, Problem 1 can be formulated as the following dynamic program:[16]

$$V^t(\vec{\mathcal{P}},\vec{\mathcal{B}}) = \max_{\substack{S_\ell \in \mathcal{P}_\ell \\ |S_\ell| \le K_\ell \forall \ell \in I\cup J}} \left\{ \mathbb{E}\left[\sum_{\ell \in I \cup J} \sum_{\ell' \in S_\ell} \mu^t_{\ell\ell'} \cdot \mathbb{1}_{\{\ell' \in \mathcal{B}_\ell\}} + \frac{1}{2} \cdot \mu^t_{\ell,\ell'} \cdot \mathbb{1}_{\{\ell \notin \mathcal{B}_\ell\}} + V^{t+1}\left(\vec{\mathcal{P}} \setminus \left(\vec{S} \cup \vec{D}\right), \vec{\mathcal{B}} \cup \vec{L} \setminus \vec{S}\right)\right]\right\}$$

$$V^{T+1}(\vec{\mathcal{P}},\vec{\mathcal{B}}) = 0$$

where $\vec{L} = \{L_\ell\}_{\ell \in I \cup J}$ and $\vec{D} = \{D_\ell\}_{\ell \in I \cup J}$ are such that L_ℓ and D_ℓ represent the sets of users that liked and disliked ℓ in that period, i.e.,

$$L_\ell = \{\ell' \in \mathcal{P}_\ell : \ell \in S'_\ell, \Phi_{\ell'\ell} = 1\}$$

$$D_\ell = \{\ell' \in \mathcal{P}_\ell : \ell \in S'_\ell, \Phi_{\ell'\ell} = 0\}$$

Note that the first term in the summation ($\mu^t_{\ell\ell'} \cdot \mathbb{1}_{\{\ell' \in \mathcal{B}_\ell\}}$) captures sequential matches, i.e., matches produced by cases (i) and (ii) mentioned above, while the second term ($\mu^t_{\ell,\ell'} \mathbb{1}_{\{\ell \notin \mathcal{B}_\ell\}}$) captures non-sequential matches generated produced by case (iii).[17] The latter term is multiplied by $1/2$ to avoid double-counting non-sequential matches.

A.2. Offline Optimum versus Online Optimum

Consider the following instance taken from Chen et al. (2009). Let I and J be the two sides of the market, with n users on each side. In addition, let $\phi_{i,j} = (2\log n)/n$ and $\phi_{j,i} = 1$ for all $j \in J,\ i \in I$, so that $\beta_e = 2\log n/n$ for each edge. Finally, suppose that $T = 1$ and $K_\ell = 1$ for all $\ell \in I \cup J$. According to Erdős and Rényi (1964, 1968), with constant probability there exists a perfect matching which means that the value of the omniscient optimum is $c \cdot n$ for some constant $c > 0$. On the other hand, the optimal one-directional policy shows $S_{i_k} = \{j_k\}$ and produces a matching of expected size $2\log n$. Therefore, the gap is $O(\frac{\log n}{n})$.

A.3. Two-Period Model

A.3.1. Definitions.

Definition 2 (Monotonicity). A non-negative set function $f : \{0,1\}^{\vec{E}} \to \mathbb{R}_+$ is *monotone* if, for every $\mathbf{x} \in \{0,1\}^{\vec{E}}$ and $e \in \vec{E}$ such that $x_e = 0$, we have $f(\mathbf{x} + \mathbb{1}_{\{e\}}) \ge f(\mathbf{x})$.

[16] In a slight abuse of notation, we assume that the set operations in $V^{t+1}(\cdot)$ are component-wise, i.e., $\vec{\mathcal{P}} \setminus \left(\vec{S} \cup \vec{D}\right) = \{\mathcal{P}_\ell \setminus (S_\ell \cup D_\ell)\}_{\ell \in I \cup J}$ and $\vec{\mathcal{B}} \cup \vec{L} \setminus \vec{S} = \{\mathcal{B}_\ell \cup L_\ell \setminus S_\ell\}_{\ell \in I \cup J}$.

[17] This holds because it cannot happen that $\ell \in \mathcal{B}_{\ell'}$ and $\ell' \in \mathcal{B}_\ell$.

Definition 3 (Submodularity). A non-negative set function $f:\{0,1\}^{\vec{E}}\to\mathbb{R}_+$ is *submodular* if, for every $\mathbf{x}\in\{0,1\}^{\vec{E}}$ and for every $e,e'\in\vec{E}$ such that $x_e=x_{e'}=0$, we have $f(\mathbf{x}+\mathbb{1}_{\{e\}}+\mathbb{1}_{\{e'\}})-f(\mathbf{x}+\mathbb{1}_{\{e'\}})\le f(\mathbf{x}+\mathbb{1}_{\{e\}})-f(\mathbf{x})$, where $\mathbb{1}_{\{e\}}\in\{0,1\}^{\vec{E}}$ is the indicator vector with value 1 in component e and zero elsewhere.

A common approach in submodular optimization is to extend the set function $f:\{0,1\}^{\vec{E}}\to\mathbb{R}_+$ into a continuous domain $[0,1]^{\vec{E}}$. Although there are different continuous extensions, the most useful in terms of algorithmic applications is the *multilinear extension*.

Definition 4 (Multilinear Extension). For a set function $f:\{0,1\}^{\vec{E}}\to\mathbb{R}_+$, we define its *multilinear extension* $F:[0,1]^{\vec{E}}\to\mathbb{R}_+$ by

$$F(x)=\sum_{S\subseteq\vec{E}} f(S)\prod_{e\in S}x_e\prod_{e\in\vec{E}\setminus S}(1-x_e).$$

A.3.2. Missing Proofs. With a slight abuse of notation, let $\tilde{f}(\mathbf{B})$ be the expected number of matches generated in the second period in the general case (i.e., obtained from solving (3)).

Proof of Proposition 1. Let $\mathcal{X}(\mathbf{B})$ and $\mathcal{W}(\mathbf{B})$ be the sets of backlog and non-backlog pairs shown in the optimal solution of the second period problem. Consider the following example. Let $I=\{i_1,i_2\}$, $J=\{j_1,j_2\}$, and the following probabilities:

$$\phi_{i_1,j_1}=1,\phi_{j_1,i_1}=\epsilon,\phi_{i_2,j_2}=\epsilon,\phi_{j_2,i_2}=1,\beta_{i_1,j_2}=1/2,\beta_{i_2,j_1}=1/2.$$

If $\mathbf{B}=\emptyset$, then $\mathcal{W}(\mathbf{B})=\{(i_1,j_2),(i_2,j_1)\}$. If we add (i_2,j_2) to the backlog, then $\mathcal{W}(\mathbf{B}\cup(i_2,j_2))=\{(i_2,j_1)\}$ and $\mathcal{X}(\mathbf{B}\cup(i_2,j_2))=\{(j_2,i_2)\}$. Hence,

$$\tilde{f}(\mathbf{B}\cup(i_2,j_2))-\tilde{f}(\mathbf{B})=1+1/2-(1/2+1/2)=1/2.$$

On the other hand, if $\mathbf{B}'=\{(j_1,i_1)\}$, $\mathcal{W}(\mathbf{B}')=\{(i_2,j_1)\}$ and $\mathcal{X}(\mathbf{B}')=\{(i_1,j_1)\}$. If we add (i_2,j_2) to $\mathbf{B}'$, then $\mathcal{W}(\mathbf{B}'\cup(i_2,j_2))=\{(i_2,j_1)\}$ and $\mathcal{X}(\mathbf{B}'\cup(i_2,j_2))=\{(i_1,j_1),(j_2,i_2)\}$. Then,

$$\tilde{f}(\mathbf{B}'\cup(i_2,j_2))-\tilde{f}(\mathbf{B}')=1+1+1/2-(1+1/2)=1.$$

Hence, we have that

$$\mathbf{B}\subset\mathbf{B}' \quad\text{and}\quad \tilde{f}(\mathbf{B}\cup(i_2,j_2))-\tilde{f}(\mathbf{B})<\tilde{f}(\mathbf{B}'\cup(i_2,j_2))-\tilde{f}(\mathbf{B}'),$$

so we conclude that $\tilde{f}(\mathbf{B})$ is not submodular. □

Proposition 8. *The function $\tilde{f}(\mathbf{B})$ can be efficiently evaluated by solving a linear program.*

Proof of Proposition 8. Given a realized backlog $\mathbf{B}=\{B_\ell\}_{\ell\in I\cup J}$ and a set of potentials $\mathcal{P}$, define a bipartite graph with two sides U,V with $U=I\cup J$ and $V=B\cup\{(i,j)\in I\times J: j\in\mathcal{P}_i, i\in\mathcal{P}_j\}$, i.e., U contains the set of users and V the set of arcs that could be displayed in the second period. Let $E\subseteq U\times V$ be the set of edges. Then, a pair $(\ell,(\ell',\ell''))\in U\times V$ belongs to E if and only if

$$(\ell',\ell'')\in B_\ell \quad\text{or}\quad [(\ell',\ell'')\notin B_\ell, \ell\in\{\ell',\ell''\}].$$

In words, an edge between $\ell\in U$ and $(\ell',\ell'')\in V$ exists if and only if the edge (ℓ',ℓ'') is either in the backlog of ℓ or both users ℓ and ℓ' can see each other simultaneously. Now, for any pair of nodes $(u,v)\in E$ such that $v \in \mathbf{B}$, we define a variable $y_{u,v}$ that is equal to 1 if $v = (u, u') \in B_u$ and u sees u', and zero otherwise. Similarly, for any pair $(u, v) \in E$ such that $v = (u, u') \in V \setminus \mathbf{B}$, we define a variable $x_{u,v}$ that is equal to 1 if u sees u', and zero otherwise. Note that here we do a slight abuse of notation and assume that if $x_{i,(i,j)} = 1$, then $i \in I$ sees profile $j \in P_i \setminus B_i$ and, similarly, if $x_{j,(i,j)} = 1$, then $j \in J$ sees $i \in P_j \setminus B_j$. For convenience, we will use in these cases that $i \in (i, j)$ and $j \in (i, j)$. Finally, for any pair $(u, v) \in E$ such that $u \in I$ and $v \in V \setminus \mathbf{B}$, let $w_{u,v} = 1$ if both users involved in v see each other simultaneously, i.e., if $v = (i, j) \in V \setminus \mathbf{B}$, then $x_{i,(i,j)} = x_{j,(i,j)} = 1$.

Using these variables, we can formulate the second period problem as follows:

$$\tilde{f}(\mathbf{B}) := \max \quad \sum_{u \in U} \sum_{v \in \mathbf{B}} y_{u,v} \cdot \phi_v^2 + \sum_{i \in I} \sum_{\delta(i) \cap V \setminus \mathbf{B}} w_{i,v} \cdot \beta_v^2 \tag{18a}$$

$$s.t. \quad \sum_{v \in \delta(u) \cap \mathbf{B}} y_{u,v} + \sum_{v \in \delta(u) \cap V \setminus \mathbf{B}} x_{u,v} \leq K_u, \quad \forall u \in U \tag{18b}$$

$$w_{i,(i,j)} - x_{u,(i,j)} \leq 0, \quad \forall (i,j) \in V \setminus \mathbf{B}, i \in I, u \in (i,j) \tag{18c}$$

$$w_{i,(i,j)} - x_{u,(i,j)} \leq 0, \quad \forall (i,j) \in V \setminus \mathbf{B}, i \in I, u \in (i,j) \tag{18d}$$

$$x_{u,v} \in \{0,1\}, \quad v \in V \setminus \mathbf{B}, u \in v \tag{18e}$$

$$y_{u,v} \in \{0,1\}, \quad \forall u \in U, v \in \delta(u) \cap \mathbf{B} \tag{18f}$$

$$w_{i,v} \in \{0,1\}, \quad \forall v \in V \setminus \mathbf{B}, i \in v \tag{18g}$$

where $\beta_v^2 = \phi_v^2 \cdot \phi_{\bar{v}}^2$ is the match probability between the users in the pair v and $\delta(u)$ is the set of edges incident to node $u \in U$. Let Q^2 be the set of constraints in (18b), (18c) and (18d). Note that each variable appears at most twice in Q^2, and that every time they appear they are multiplied by either 1 or -1. Thus, to show that the matrix of constraints if totally unimodular, it remains to show that the constraints in Q^2 can be separated in two subsets such that (i) if a variable appears twice with different signs, then the constraints belong to the same subset, and (ii) if a variable appears twice with the same sign, then the constraints belong to different subsets. Let Q_I^2 and Q_J^2 be the subsets of constraints of Q^2 involving $u \in I$ and $u \in J$, respectively. Then, observe that

- each $w_{i,(i,j)}$ appears in two constraints with the same sign ($w_{i,(i,j)} - x_{i,(i,j)} \leq 0$ and $w_{i,(i,j)} - x_{j,(i,j)} \leq 0$), but these constraints belongs to Q_I^2 and Q_J^2, respectively.
- each $x_{i,(i,j)}$ appears in two constraints with different signs ($\sum_{v \in \delta(i) \cap B_i} y_{i,v} + \sum_{v \in \delta(i) \cap V \setminus B_i} x_{i,v} \leq K_i$ and $w_{i,(i,j)} - x_{i,(i,j)} \leq 0$), but these constraints belong both to the same subset Q_I^2. Similarly, $x_{j,(i,j)}$ appears in two constraints with different signs, but both constraints belong to Q_J^2.

Hence, using Hoffman's sufficient condition, we conclude that the constraints in Q^2 can be written as the product of a totally unimodular matrix and our vector of decisions variables. Finally, since the right-hand sides of the constraints are integral, we conclude that the feasible region of the problem is an integral polyhedron, and thus we can solve its linear relaxation. □

Proof of Lemma 1. Let us construct a partition of the edges $\vec{E}$: Define part $\mathcal{E}_\ell$ for each $\ell \in I \cup J$ as the set $\{(\ell, \ell') : \ell' \in \mathcal{B}_\ell\}$. The budget for each part is K_ℓ. This shows that $R^2(\mathbf{B})$ corresponds to a partition matroid. Then, using Proposition 3.1 in Fisher et al. (1978), we conclude that $f(\mathbf{B})$ is monotone and submodular. □

LEMMA 4. *The function $\mathcal{M}^2(\mathbf{x},\mathbf{w})$ is non-negative monotone and submodular in $\mathbf{x}$.*

Proof of Lemma 4. Since $\mathcal{M}^2(\mathbf{x},\mathbf{w})$ does not directly depend on $\mathbf{w}$, we will drop it from the notation to ease exposition and simply write $\mathcal{M}^2(\mathbf{x})$.

Recall that f is a non-negative, monotone submodular function due to Lemma 1. Since f is non-negative, then $\mathcal{M}^2$ is also non-negative. For $\mathbf{x}\in\{0,1\}^{\vec{E}}$, let $\vec{E}(\mathbf{x})=\{e\in\vec{E}:\ x_e=1, x_{\overline{e}}=0\}$ where $\overline{e}$ denotes the inverted arc e, i.e., if $e=(\ell,\ell')$ then $\overline{e}=(\ell',\ell)$. [18] Then, we can re-write $\mathcal{M}^2$ as follows

$$\mathcal{M}^2(\mathbf{x})=\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{B})\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)$$

On the other hand, for any $e\notin\vec{E}(\mathbf{x})$ we have

$$\begin{aligned}\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}})=&\ \phi_e^1\cdot\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{B}\cup e)\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)\\&+(1-\phi_e^1)\cdot\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{B})\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)\end{aligned}$$

Therefore,

$$\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}})-\mathcal{M}^2(\mathbf{x})=\phi_e^1\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})}[f(\mathbf{B}\cup e)-f(\mathbf{B})]\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)$$

Since f is monotone, then $f(\mathbf{B}\cup e)-f(\mathbf{B})\geq 0$ for all $\mathbf{B}\subseteq\vec{E}(\mathbf{x})$ and $e\notin\vec{E}(\mathbf{x})$, which implies $\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}})-\mathcal{M}^2(\mathbf{x})\geq 0$.

Now, let us prove that $\mathcal{M}^2$ is submodular. Consider any $\mathbf{x}\in\{0,1\}^{\vec{E}}$ and $e,e'\notin\vec{E}(\mathbf{x})$. Our goal is to show an alternative characterization of submodularity: $\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}})-\mathcal{M}^2(\mathbf{x})\geq\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}}+\mathbb{1}_{\{e'\}})-\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e'\}})$. Note that we have the following expression for $\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}}+\mathbb{1}_{\{e'\}})$

$$\begin{aligned}\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}}+\mathbb{1}_{\{e'\}})=&\ \phi_e^1\phi_{e'}^1\cdot\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{B}\cup e,e')\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)\\&+\phi_{e'}^1(1-\phi_e^1)\cdot\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{B}\cup e')\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)\\&+\phi_e(1-\phi_{e'}^1)\cdot\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{B}\cup e)\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)\\&+(1-\phi_{e'}^1)(1-\phi_e^1)\cdot\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{B})\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1)\end{aligned}$$

Analogously, we can compute $\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}})$ and $\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e'\}})$. By deleting common terms, we can obtain the following

$$\begin{aligned}&\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}})-\mathcal{M}^2(\mathbf{x})-\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e\}}+\mathbb{1}_{\{e'\}})+\mathcal{M}^2(\mathbf{x}+\mathbb{1}_{\{e'\}})\\&=\phi_e^1\phi_{e'}^1\cdot\sum_{\mathbf{B}\subseteq\vec{E}(\mathbf{x})}[f(\mathbf{B}\cup e)-f(\mathbf{B})-f(\mathbf{B}\cup e,e')+f(\mathbf{B}\cup e')]\prod_{e\in\mathbf{B}}\phi_e^1\prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{B}}(1-\phi_e^1),\end{aligned}$$

from which submodularity follows due to submodularity of f. □

[18] Note that, when $\mathbf{x}\in R^{1,\Pi}$, for any pair $i\in I$, $j\in\mathcal{P}_i^1$ we cannot have $(i,j),(j,i)\in\vec{E}(\mathbf{x})$ because of one of the constraint of the problem.

A.4. Dating Heuristic (DH)

In this section, we formalize DH as initially introduced in (Rios et al. 2023). For each period $t \in [T]$, DH considers two steps:

1. Optimization: this step involves solving the following linear program:

$$
\begin{aligned}
\max \quad & \sum_{\tau=t}^{t+1} \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}_\ell^t} y_{\ell,\ell'}^\tau \phi_{\ell,\ell'}^\tau + \frac{1}{2} \cdot w_{\ell,\ell'}^\tau \phi_{\ell,\ell'}^\tau \phi_{\ell'\ell}^\tau \\
\text{s.t.} \quad & \sum_{\tau=t}^{t'} y_{\ell,\ell'}^\tau \leq \mathbb{1}_{\left\{\ell' \in \mathcal{B}_\ell^t\right\}} + \sum_{\tau=t}^{t'-1} (x_{\ell'\ell}^\tau - w_{\ell,\ell'}^\tau) \phi_{\ell'\ell}^\tau, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^t,\ t' \in [t, t+1], \\
& \sum_{\tau=t}^{t+1} x_{\ell,\ell'}^\tau + y_{\ell,\ell'}^\tau \leq 1, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^t, \\
& \sum_{\ell'} x_{\ell,\ell'}^\tau + y_{\ell,\ell'}^\tau \leq K_\ell, && \forall \ell \in I \cup J, t \in [t, t+1], \\
& w_{\ell,\ell'}^\tau \leq x_{\ell,\ell'}^\tau,\ w_{\ell,\ell'}^\tau \leq x_{\ell'\ell}^\tau,\ w_{\ell,\ell'}^\tau = w_{\ell'\ell}^\tau, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^t,\ \tau \in [t, t+1], \\
& x_{\ell,\ell'}^\tau,\ y_{\ell,\ell'}^\tau, w_{\ell,\ell'}^\tau \in [0,1], && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^t,\ \tau \in [t, t+1].
\end{aligned}
\tag{19}
$$

The decision variables $y_{\ell,\ell'}^t$ and $x_{\ell,\ell'}^t$ represent whether ℓ sees profile ℓ' in period t as part of a backlog and to initiate a sequential match, respectively. The objective is to maximize the expected number of matches obtained in periods $\{t, t+1\}$, including sequential (first term in the objective) and non-sequential matches (second term in the objective). The first family of constraints defines y and captures the evolution of the backlog. The second family of constraints captures that a profile can be shown at most once, while the third considers the cardinality constraints. Finally, the last family of constraints captures the definition of $w_{\ell,\ell'}^t$, which accounts for non-sequential matches between ℓ and ℓ' in period t.

2. Rounding: since the optimal decisions $x^{*,t}, y^{*,t}, w^{*,t}$ of (19) may be fractional, this step involves rounding them in order to decide the profiles to show in the current period. Specifically, the rounding process starts by adding to S_ℓ^t the profiles for which $y_{\ell,\ell'}^t > 0$ (in decreasing order). Then, if there is space left, the rounding procedure adds to S_ℓ^t the profiles for which $x_{\ell,\ell'}^t > 0$ (in decreasing order), making sure that the cardinality constraints are satisfied.

Notice that these two steps consider the current set of potentials $\mathcal{P}_\ell^t$ and backlog $\mathcal{B}_\ell^t$ for each user $\ell \in I \cup J$. Then, at the end of each period, the sets of potentials and the backlogs are updated considering the profiles shown and the like/dislike decisions, as shown in (1). This procedure is formally described in Algortihm 3.

A.4.1. Proof of Lemma 2. Let $(\mathbf{x}, \mathbf{w}, \mathbf{y})$ be a solution in Problem (10). Then, define $(\hat{\mathbf{x}}, \hat{\mathbf{w}}, \hat{\mathbf{y}})$ as follows: $\hat{\mathbf{y}} = \mathbf{y}$ and

- If $x_{\ell,\ell'} = 1$, $x_{\ell',\ell} = 0$ and $w_{\ell,\ell'} = w_{\ell',\ell} = 0$, then set $\hat{x}_{\ell,\ell'} = 1$ and $\hat{x}_{\ell',\ell} = 0$ and $w_e = 0$.
- If $x_{\ell,\ell'} = x_{\ell',\ell} = 1$ and $w_{\ell,\ell'} = w_{\ell',\ell} = 1$, then set $\hat{x}_{\ell,\ell'} = \hat{x}_{\ell',\ell} = 0$ and $\hat{w}_e = 1$.
- If $x_{\ell,\ell'} = x_{\ell',\ell} = 0$ and $w_{\ell,\ell'} = w_{\ell',\ell} = 0$, then set $\hat{x}_{\ell,\ell'} = \hat{x}_{\ell',\ell} = \hat{w}_e = 0$

Clearly $(\hat{\mathbf{x}}, \hat{\mathbf{w}}, \hat{\mathbf{y}})$ is feasible in Problem (11) and it has the same objective value. To show the opposite direction, the construction of variables is analogous. □

Algorithm 3 Dating Heuristic (DH), (Rios et al. 2023)

Input: An instance of the problem.

Output: A feasible subset of profiles for each user in each period.

1: **for** $t \in [T]$ **do**
2: Solve (19) and keep $x^{*,t}$, $y^{*,t}$.
3: For each user ℓ, sequentially add profiles ℓ' for which $y^{*,t}_{\ell,\ell'} > 0$ until the cardinality constraint is binding. If the latter is not binding, add profiles ℓ' for which $x^{*,t}_{\ell,\ell'} > 0$ until the constraint becomes binding.
4: Update potentials and backlogs following (1).

Appendix B: Appendix to Section 4

B.1. Missing Proofs in Section 4.1

B.1.1. Local Greedy. The Local Greedy policy selects, for each user and period, the subset of profiles that maximizes their expected number of matches, i.e.,

$$S^t_\ell = \underset{S \subseteq \mathcal{P}^t_\ell : |S| \le K_\ell}{\arg\max} \left\{ \sum_{\ell' \in S} \phi_{\ell,\ell'} \cdot \left(\mathbb{1}_{\{\ell' \in B^t_\ell\}} + \phi_{\ell'\ell} \cdot \mathbb{1}_{\{\ell' \notin B^t_\ell\}} \right) \right\}.$$

Proof of Proposition 2. Suppose that there are n users on each side of the market, i.e., $I = \{i_1, \dots, i_n\}$ and $J = \{j_1, \dots, j_n\}$. In addition, suppose that $\mathcal{P}^1_i = J$ for every $i \in I$, $\mathcal{P}^1_j = I$ for every $j \in J$ and $K_\ell = 1$ for all $\ell \in I \cup J$. Let us set the probabilities: $\beta^1_{i,j} = 1$ for $j = j_1$ and for all $i \in I$, $\beta^1_{i,j} = 1 - \varepsilon$ for all $i \in I$ and $j \neq j_1$, while $\beta^2_{i,j} = 0$ for all $i \in I \cup J$, $j \in \mathcal{P}^1_i$. In this setting, the Local Greedy policy will choose $S^1_i = \{j_1\}$ for every user i, and therefore only one match will take place in expectation. In contrast, an optimal solution is to assign $S^1_{i_k} = \{j_k\}$, which leads to $1 + (n-1)(1-\varepsilon)$ matches in expectation. Then, the performance of the greedy policy is given by $1/(1 + (n-1)(1-\varepsilon)) \to 0$, as $n \to \infty$ for ε sufficiently small. $\square$

B.1.2. Perfect Matching. The Perfect Matching heuristic solves, in each period $t \in [T]$, the following problem:

$$\begin{aligned}
\max \quad & \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}^1_\ell} y^t_{\ell,\ell'} \phi^t_{\ell,\ell'} + \frac{1}{2} w^t_{\ell,\ell'} \beta^t_{\ell,\ell'} \\
st. \quad & y^t_{\ell,\ell'} \le \mathbb{1}_{\{\ell' \in \mathcal{B}^t_\ell\}}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^t_\ell \\
& x^t_{\ell,\ell'} + y^t_{\ell,\ell'} \le 1, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^t_\ell \\
& \sum_{\ell' \in \mathcal{P}^t_\ell} x^t_{\ell,\ell'} + y^t_{\ell,\ell'} \le K_\ell, && \forall \ell \in I \cup J \\
& w^t_{\ell,\ell'} \le x^t_{\ell,\ell'},\ w^t_{\ell,\ell'} \le x^t_{\ell'\ell},\ w^t_{\ell,\ell'} = w^t_{\ell'\ell}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^t_\ell \\
& x^t_{\ell,\ell'}, y^t_{\ell,\ell'}, w^t_{\ell,\ell'} \in \{0,1\}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^t_\ell.
\end{aligned} \tag{20}$$

Then, set $S_\ell^t = \left\{\ell' \in \mathcal{P}_\ell^t : x_{\ell,\ell'}^t = 1 \text{ or } y_{\ell,\ell'}^t = 1\right\}$ for each $\ell \in I \cup J$. Note that, if there is no initial backlog, then the problem can be re-formulated as:

$$\begin{aligned} \max \quad & \sum_{e \in E} \beta_e^t \cdot w_e^t \\ st. \quad & \sum_{e \in E : \ell \in e} w_e^t \le K_\ell, && \forall \ell \in I \cup J \\ & \sum_{t \in [T]} w_e^t \le 1, && \forall e \in E \\ & w_e^t \in \{0,1\}, && \forall e \in E,\ t \in [T], \end{aligned} \tag{21}$$

where the second constraint ensures that each edge is used at most once, i.e., that no two users see each other more than once. This is captured in the previous formulation through the set of potentials $\mathcal{P}_\ell^t$.

Proof of Proposition 3. Suppose that $|I| = 2n$, $|J| = 2$, that $\mathcal{P}_i^1 = J$ for every $i \in I$, $\mathcal{P}_j^1 = I$ for every $j \in J$, $K_\ell = 1$ for all $\ell \in I \cup J$ and that $\phi_{i,j}^t = p$ while $\phi_{j,i}^t = q$ for all $i \in I$, $j \in J$, and $t \in \{1,2\}$. Then, it is easy to see that the sequential perfect match policy leads to $4pq$ matches in expectation. On the other hand, consider the policy where:

(i) In $t = 1$, $\{i_1, \ldots, i_n\}$ see j_1, $\{i_{n+1}, \ldots, i_{2n}\}$ see j_2, j_1 sees i_{2n} and j_2 sees i_1.

(ii) In $t = 2$, $\{i_1, \ldots, i_n\}$ see j_2, $\{i_{n+1}, \ldots, i_{2n}\}$ see j_1, j_1 sees any profile that liked her in $t = 1$, and same for $j = 2$.

Given this policy, the matches (i_{2n}, j_1) and (i_1, j_2) happen with probability pq each. On the other hand, j_1 matches with someone in $\{i_1, \ldots, i_n\}$ with probability $(1-(1-p)^n)q$, and the same for j_2 matching with someone in $\{i_{n+1}, \ldots, i_{2n}\}$. Then, the total expected number of matches is $2pq + 2q(1-(1-p)^n)$, which is optimal for this instance. Then, the sequential perfect match policy achieves a performance of $4pq/(2q(p+1-(1-p)^n)) \to 2/(1+n)$ when $p \to 0$, and since $2/(1+n) \to 0$ when $n \to \infty$, we conclude the proof. $\square$

B.1.3. Proof of Proposition 4. We use appropriate submodular optimization and randomized rounding tools to theoretically analyze Problem 2 under one-directional interactions and sequential matches. To formalize this analysis, let F be the multilinear extension of the set function f defined in (4), i.e.,

$$F(\mathbf{x}) = \sum_{\mathbf{B} \subseteq \vec{E}_I} f(\mathbf{B}) \prod_{e \in \mathbf{B}} x_e \prod_{e \notin \mathbf{B}} (1 - x_e).$$

Note that we define this extension over arcs in $\vec{E}_I$ since this is a one-directional setting. Also, observe that for any $\mathbf{x} \in \{0,1\}^{\vec{E}_I}$, we have $F(\phi^1 \cdot \mathbf{x}^1) = \mathcal{M}^2(\mathbf{x}^1, \mathbf{w}^1)$, where $\phi^1 \cdot \mathbf{x}^1$ denotes the vector with components $\phi_{i,j}^1 \cdot x_{i,j}^1$ for all $i \in I,\ j \in J$. As previously discussed, representing $\mathcal{M}^2(\mathbf{x}^1, \mathbf{w}^1)$ through the multilinear extension of f has the advantage that the latter can be evaluated in $[0,1]^{\vec{E}_I}$ rather than only in $\{0,1\}^{\vec{E}_I}$. Then, consider the following optimization problem:

$$\max \quad F(\mathbf{z}) \tag{22a}$$

$$\text{s.t.} \quad \sum_{j \in J : \phi_{i,j}^1 > 0} \frac{z_{i,j}}{\phi_{i,j}^1} \le K_i \qquad \text{for every } i \in I, \tag{22b}$$

$$0 \le z_e \le \phi_e^1 \qquad \text{for every } e \in \vec{E}_I. \tag{22c}$$

First, we show the following:

Lemma 5. *The optimal value of* (22a)-(22c) *is an upper bound on the optimal value of Problem 2 under one-directional interactions and sequential matches.*

Proof of Lemma 5. Consider a feasible solution $\mathbf{x}^1 \in \{0,1\}^{\vec{E}_I}$ of Problem 2 under one-directional interactions and sequential matches, that is, $\sum_{j\in J} x^1_{i,j} \le K$ for every $i \in I$. Let $\mathbf{z} = \phi^1 \cdot \mathbf{x}^1$. Note that, for each $i \in I$,

$$\sum_{j\in J:\phi^1_{i,j}>0} z_{i,j}/\phi^1_{i,j} = \sum_{j\in J:\phi^1_{i,j}>0} x^1_{i,j} \le K_i,$$

and $z_e = x^1_e \phi^1_e \le \phi^1_e$ for every $e \in \vec{E}_I$. Therefore, $\mathbf{z}$ is a feasible solution for the problem (22a)-(22c). Since the objective value of $\mathbf{x}^1$ in (7) is equal to the objective of $\mathbf{z}$ in (22a)-(22c), the proof follows.

Since f is monotone and submodular and F inherits all its properties, we can use Lemma 4.2 in Vondrák (2008) (see Corollary 1) to find our desired performance guarantee. Formally,

Corollary 1 **(Vondrák (2008))**. *There exists an efficient algorithm that computes a point* $\mathbf{z}$ *that satisfies* (22b) *and* (22c) *such that* $F(\mathbf{z}) \ge (1-1/e)\cdot F(\mathbf{z}^*)$*, where* $\mathbf{z}^*$ *is an optimal solution of* (22a)-(22c).

We emphasize that the solution $\mathbf{z}$ in Corollary 1 might be fractional, so we need to use a rounding procedure to construct the final solution of our problem. To find our feasible point we use: (i) the continuous greedy algorithm proposed by (Vondrák 2008), and (ii) the dependent randomized rounding algorithm by Gandhi et al. (2006). Specifically, the feasible solution can be obtained as follows:

1. Compute a solution $\mathbf{z}$ for the problem (22a)-(22c) using the algorithm from Corollary 1.
2. For each $i \in I$ and $j \in J$, set $\tilde{x}_{i,j} = z_{i,j}/\phi^1_{i,j}$ when $\phi^1_{i,j} > 0$ and zero otherwise.
3. Independently for each user $i \in I$, run the dependent randomized rounding algorithm Gandhi et al. (2006) on the fractional vector $\tilde{\mathbf{x}}_i \in \mathbb{R}^J$ to compute an integral random vector $\mathbf{x}_i \in \{0,1\}^J$.

Algorithm 4 Approximation Algorithm for One-Directional Interactions and Sequential Matches

Input: An instance of the problem.

Output: A feasible collection of subsets $\mathbf{S}^1$.

1: Compute a solution $\mathbf{z}$ for the problem (22a)-(22c) using the algorithm from Corollary 1.
2: For each $i \in I$ and $j \in J$, set $y_{i,j} = z_{i,j}/\phi^1_{i,j}$ when $\phi^1_{i,j} > 0$ and zero otherwise.
3: Independently for each user $i \in I$, run the dependent randomized rounding algorithm Gandhi et al. (2006) on the fractional vector $\mathbf{y}_i$ to compute an integral random vector $\mathbf{x}_i \in \{0,1\}^J$.
4: For each $i \in I$ return $S^1_i = \{j \in J : x_{i,j} = 1\}$.

Given the fractional solution $\mathbf{z}$ satisfying the guarantee in Corollary 1, let $\tilde{\mathbf{x}}_i \in [0,1]^J$ be the fractional vector such that the j-th entry is equal to $z_{i,j}/\phi^1_{i,j}$ when $\phi^1_{i,j} > 0$ and zero otherwise. Observe that thanks to constraint (22c) we have $\tilde{\mathbf{x}}_i \in [0,1]^J$ for each $i \in I$. Then, independently for each user $i \in I$, by the algorithm in Gandhi et al. (2006) it is possible to efficiently compute an integral random vector $\mathbf{x}_i \in \{0,1\}^J$ satisfying the following conditions:

1. $\sum_{j\in J} x_{i,j} \le \lceil \sum_{j\in J} \tilde{x}_{i,j} \rceil$, and

2. $\mathbb{E}[x_{i,j}] = \tilde{x}_{i,j}$ for each $i \in I$ and $j \in J$.

Thanks to condition (1) of the randomized rounding algorithm, for each $i \in I$ we have $\sum_{j\in J} x_{ij} \leq \left\lceil \sum_{j\in J} \tilde{x}_{ij} \right\rceil = \lceil \sum_{j\in J:\phi_{ij}>0} z_{ij}/\phi^1_{ij} \rceil \leq K_i$, where the last inequality holds since K_i is integral and $\tilde{\mathbf{z}}$ satisfies constraint (22b). Therefore, our algorithm gives a feasible solution for Problem 2. We now analyze the approximation guarantee.

$$\begin{aligned}
\mathbb{E}_{\mathcal{B}\sim\phi^1\mathbf{x}}[f(\mathcal{B})] &= \sum_{\mathbf{B}\subseteq\vec{E}_I} f(\mathbf{B}) \cdot \mathbb{P}_{\mathbf{x}}(\mathcal{B} = B) \\
&= \sum_{\mathbf{B}\subseteq\vec{E}_I} f(\mathbf{B}) \cdot \mathbb{E}_{\mathbf{x}}\left[\prod_{e\in\mathbf{B}} \phi^1_e x_e \prod_{e\notin\mathbf{B}} (1-\phi^1_e x_e)\right] \\
&= \sum_{\mathbf{B}\subseteq\vec{E}_I} f(\mathbf{B}) \cdot \prod_{e\in\mathbf{B}} \phi^1_e \mathbb{E}_{\mathbf{x}}[x_e] \prod_{e\notin\mathbf{B}} (1-\phi^1_e \mathbb{E}_{\mathbf{x}}[x_e]) \\
&= \sum_{\mathbf{B}\subseteq\vec{E}_I} f(\mathbf{B}) \cdot \prod_{e\in\mathbf{B}} \phi^1_e \cdot \frac{z_e}{\phi^1_e} \prod_{e\notin\mathbf{B}} (1-\phi^1_e \cdot \frac{z_e}{\phi^1_e}) = \mathcal{M}^2(\mathbf{z}),
\end{aligned}$$

where the second equality comes from the fact that $\mathbf{x}_i$ is independent from $\mathbf{x}_{i'}$ for every $i, i' \in I$ with $i \neq i'$, and the third equality comes from condition 2 of the randomized rounding procedure. Finally, Lemma 5 states that $\text{OPT}' \geq \text{OPT}$, where OPT' is the optimal value of Problem (22a)-(22c), so we conclude the proof by using Corollary (1). $\square$

B.2. Approaches Based on Submodular Maximization

B.2.1. Missing Proofs in Section 4.2 . We first show that for this setting, the feasible region corresponds to the intersection of two matroids.

LEMMA 6. *The feasible region $R^{1,\Pi}$ corresponds to the intersection of two matroids when Π is restricted to two-directional interactions with sequential matches.*

Proof. Our ground set of elements is $\mathcal{E} = \vec{E}_I \cup \vec{E}_J$. The first partition consists of the following parts: $\mathcal{E}_\ell = \{e : e = (\ell, \ell') \text{ for every } \ell' \in \mathcal{P}^1_\ell\}$ for all $\ell \in I \cup J$. It is easy to check that $\mathcal{E} = \cup_{\ell\in I\cup J}\mathcal{E}_\ell$ and $\mathcal{E}_\ell \cap \mathcal{E}_{\ell'} = \emptyset$ for every ℓ, ℓ' such that $\ell \neq \ell'$. Finally, the budget for each part $\mathcal{E}_\ell$ is K_ℓ. Now, let us construct the second partition matroid. For every pair $i \in I$ and $j \in J$, we define a part $\mathcal{E}_{i,j}$ as the set $\{(i,j),(j,i)\}$. Indeed this forms a partition of $\mathcal{E}$. Finally, the budget for each part $\mathcal{E}_{i,j}$ is 1.

Proof of Proposition 5. Consider Problem 2 with two-directional interactions and sequential matches. Since $\mathbf{x}^1$ is a 0-1 vector and $\mathcal{M}^2(\cdot)$ is monotone submodular over elements in $\vec{E}$, we know that a vanilla greedy algorithm achieves a $1/(1+r)$-approximation for the problem of maximizing a monotone submodular function over the intersection of r matroids Fisher et al. (1978) and local-search guarantees a factor of $1/(r+\epsilon)$ for any fixed $\epsilon > 0$ (Lee et al. 2009). By Lemma 6, we know that $r = 2$, and, thus, the guarantee is $1/3$ for greedy and $1/(2+\epsilon)$ for local-search.

B.2.2. Missing Proofs in Section 4.3 Before proving the next result, let us consider the following definition:

DEFINITION 5 (LAMINAR MATROID). A family $\mathcal{X} \subseteq 2^{\mathcal{E}}$ over a ground set of elements $\mathcal{E}$ is called laminar if for any $X, Y \in \mathcal{X}$ we either have $X \cap Y = \emptyset$, $X \subseteq Y$ or $X \subseteq Y$. Assume that for each element $u \in \mathcal{E}$ there exists some $A \in \mathcal{X}$ such that $A \ni u$. For each $A \in \mathcal{X}$ let $c(A)$ a positive integer associated with it. A laminar matroid $\mathcal{I}$ is defined as $\mathcal{I} = \{A \subseteq \mathcal{E} : |A \cap X| \leq c(X) \ \forall X \in \mathcal{X}\}$.

Lemma 7. *The feasible region $R^{1,\Pi}$ corresponds to the intersection of a partition and a laminar matroid when Π is restricted to one-directional interactions allowing non-sequential matches in the first period.*

Proof. Our ground set of elements is $\mathcal{E} = \vec{E}_I \cup E$. First, let us define our laminar family $\mathcal{X}$. For every pair $i \in I, j \in J$ consider $X_{i,j} = \{(i,j), \{i,j\}\}$, also for every $i \in I$ consider $Y_i = \{(i,\ell) : \ell \in \mathcal{P}_i^1\} \cup \{\{i,\ell\} : \ell \in \mathcal{P}_i^1\}$. Indeed, this is a laminar family, two sets of type X do not intersect and two sets of type Y also do not intersect. Sets of type X and Y intersect only if they correspond to the same $i \in I$ in which case $X_{i,j} \subseteq Y_i$. Finally, for every $X_{i,j}$ we have $c(X_{i,j}) = 1$ and for each Y_i we have $c(Y_i) = K_i$. Therefore, for this laminar family $\mathcal{X}$ and values $c(\cdot)$ we have that $\mathcal{I} = \{A \subseteq \mathcal{E} : |A \cap X| \le c(X) \ \forall X \in \mathcal{X}\}$ coincides with feasible region Q^1. The partition matroid is defined by the following parts: for each $j \in J$ consider $\mathcal{E}_j = \{\{i,j\} : i \in \mathcal{P}_j\}$ and $\mathcal{E}_0 = \mathcal{E} \setminus \bigcup_{j\in J} \mathcal{E}_j$. The budget for each part $\mathcal{E}_j$ is K_j and for $\mathcal{E}_0$ is $|\mathcal{E}_0|$.

Proof of Proposition 6. Similar to the proof of Proposition 5.

B.2.3. Missing Proofs in Section 4.4

Lemma 8. *The feasible region $R^{1,\Pi}$ corresponds to the intersection of three partition matroids when Π is restricted to two-directional interactions allowing non-sequential matches in the first period.*

Proof of Lemma 8. Our ground set of elements is $\mathcal{E} = \vec{E}_I \cup \vec{E}_J \cup E$. As in Lemma 6, the first partition consists of the following parts: $\mathcal{E}_i = \{\{(i,j), \{i,j\}\} :$ for every $j \in \mathcal{P}_i^1\}$ for all $i \in I$ and $\mathcal{E}_0 = \mathcal{E} \setminus \cup_{i\in I} \mathcal{E}_i$. It is easy to check that $\mathcal{E} = \mathcal{E}_0 \cup \bigcup_{i\in I} \mathcal{E}_\ell$ and $\mathcal{E}_\ell \cap \mathcal{E}_{\ell'} = \emptyset$ for every $\ell, \ell' \in I \cup \{0\}$ such that $\ell \neq \ell'$. Finally, the budget for each part $\mathcal{E}_i$ is K_i for every $i \in I$ and $|\mathcal{E}_0|$ for $\mathcal{E}_0$. Analogously, we can construct the second partition matroid by considering parts: $\mathcal{E}_{i,j} = \{\{(j,i), \{i,j\}\} :$ for every $i \in \mathcal{P}_j^1\}$ for all $j \in J$ and $\mathcal{E}_0 = \mathcal{E} \setminus \cup_{j\in J} \mathcal{E}_j$. The final partition is composed by the following parts: For every pair $i \in I$ and $j \in J$, we define a part $\mathcal{E}_{i,j}$ as the set $\{(i,j), (j,i), \{i,j\}\}$. Indeed this forms a partition of $\mathcal{E}$. Finally, the budget for each part $\mathcal{E}_{i,j}$ is 1.

Proof of Proposition 7. Similar to the proof of Proposition 5.

B.2.4. Policy Based on Submodular Maximization Algorithms. Our policy based on submodular optimization approaches is formalized in Algorithm 5.

Algorithm 5 Policy based on submodular maximization

Input: An instance of the problem and an algorithm ALG for submodular maximization subject to the intersection of r matroids.

Output: Feasible subsets: $\mathbf{x}^1$, $\mathbf{x}^2$

1: Use ALG to obtain an approximate solution $\mathbf{x}^1$ of Problem 2.

2: Observe the backlogs $\mathbf{B}$ generated by $\mathbf{x}^1$ according to (2)

3: Obtain $\mathbf{x}^2$ by solving to optimality (4).

B.3. Missing Proofs in Section 4.5

Consider an optimal solution $(\mathbf{x}^\star, \mathbf{w}^\star, \mathbf{y}^\star)$ of (11) and the following formulation

$$\begin{aligned} F(\mathbf{x}^\star, \mathbf{w}^\star) := \max \quad & \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}^1_\ell} \phi^2_{\ell,\ell'} \cdot y_{\ell,\ell'} && && (23)\\ s.t. \quad & y_{\ell,\ell'} \leq \phi^1_{\ell',\ell} \cdot x^\star_{\ell',\ell}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^1_\ell \\ & \sum_{\ell' \in \mathcal{P}^1_\ell} y_{\ell,\ell'} \leq K_\ell, && \forall \ell \in I \cup J \\ & y_{\ell,\ell'} \geq 0, && \forall \ell \in I \cup J, \ell' \in \mathcal{P}^1_\ell. \end{aligned}$$

Clearly, $\mathbf{y}^\star$ is an optimal solution in $F(\mathbf{x}^\star, \mathbf{w}^\star)$. Define $\mathcal{X}_\ell = \{\ell' \in \mathcal{P}^1_\ell :\ x^\star_{\ell',\ell} = 1\}$ and observe that, in the problem above, we can restrict variables $y_{\ell,\ell'}$ to $\ell' \in \mathcal{X}_\ell$ (the others are zero). Also, note that (23) is separable, i.e., $F(\mathbf{x}^\star, \mathbf{w}^\star) = \sum_{\ell \in I \cup J} F_\ell(\mathbf{x}^\star, \mathbf{w}^\star)$, where

$$F_\ell(\mathbf{x}^\star, \mathbf{w}^\star) := \max \left\{ \sum_{\ell' \in \mathcal{X}_\ell} \phi^2_{\ell,\ell'} \cdot y_{\ell'} :\ \sum_{\ell' \in \mathcal{X}_\ell} y_{\ell'} \leq K_\ell,\ y_{\ell'} \leq \phi^1_{\ell',\ell},\ \forall\, \ell' \in \mathcal{X}_\ell \right\}. \tag{24}$$

Our goal is to compare $F(\mathbf{x}^\star, \mathbf{w}^\star)$ with the *distribution problem*, which considers a larger space of distributions over the backlogs. For $\mathbf{x}^\star, \mathbf{w}^\star$, the distribution problem is defined as:

$$\begin{aligned} G(\mathbf{x}^\star, \mathbf{w}^\star) := \max \quad & \sum_{\ell \in I \cup J} \sum_{B \subseteq \mathcal{P}^1_\ell} f_\ell(B) \cdot \lambda_{\ell,B} \\ s.t. \quad & \sum_{B \subseteq \mathcal{P}^1_\ell} \lambda_{\ell,B} = 1, && \forall \ell \in I \cup J \\ & \sum_{\substack{B \subseteq \mathcal{P}^1_\ell \\ \ell' \in B}} \lambda_{\ell,B} = \phi_{\ell',\ell} \cdot x^\star_{\ell',\ell}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}^1_\ell \\ & \lambda_{\ell,B} \geq 0, && \forall \ell \in I \cup J,\ B \subseteq \mathcal{P}^1_\ell. \end{aligned}$$

where $\lambda_{\ell,B}$ can be interpreted as the probability that the backlog of user ℓ is B. The second family of constraints states that the probability of the backlog of ℓ containing ℓ' equals the marginal probability that ℓ' saw and liked ℓ in the first period. Similar to $F(\mathbf{x}^\star, \mathbf{w}^\star)$, note that $G(\mathbf{x}^\star, \mathbf{w}^\star)$ is separable since $f(\mathbf{B}) = \sum_{\ell \in I \cup J} f_\ell(B)$, where $f_\ell(B) := \max\{\sum_{\ell' \in S} \phi^2_{\ell,\ell'} :\ |S| \leq K_\ell,\ S \subseteq B\}$ (by independence of like decisions across users). Also, we can limit variables $\lambda_{\ell,B}$ to subsets $B \subseteq \mathcal{X}_\ell$ because $x^\star_{\ell',\ell} = 0$ implies that $\lambda_{\ell,B} = 0$ for all B such that $\ell' \in B$. Hence, we can write $G(\mathbf{x}^\star, \mathbf{w}^\star) = \sum_{\ell \in I \cup J} G_\ell(\mathbf{x}^\star, \mathbf{w}^\star)$, where

$$G_\ell(\mathbf{x}^\star, \mathbf{w}^\star) := \max \left\{ \sum_{B \subseteq \mathcal{X}_\ell} f_\ell(B) \cdot \lambda_B :\ \sum_{B \subseteq \mathcal{X}_\ell} \lambda_B = 1,\ \sum_{\substack{B \subseteq \mathcal{X}_\ell \\ \ell' \in B}} \lambda_B = \phi^1_{\ell',\ell},\ \forall\, \ell' \in \mathcal{X}_\ell \right\}. \tag{25}$$

To show our main result, we study the dual formulation of (24) and (25), which are given by (resp.):

$$\begin{aligned} F^D_\ell(\mathbf{x}^\star, \mathbf{w}^\star) := \min \quad & K_\ell \cdot \theta + \sum_{\ell' \in \mathcal{X}_\ell} \gamma_{\ell'} \cdot \phi^1_{\ell',\ell} && && (26)\\ s.t. \quad & \theta + \gamma_{\ell'} \geq \phi^2_{\ell,\ell'} && \forall \ell' \in \mathcal{X}_\ell, \\ & \theta, \gamma_{\ell'} \geq 0, && \forall \ell' \in \mathcal{X}_\ell \end{aligned}$$

and

$$
\begin{aligned}
G^D_\ell(\mathbf{x}^\star,\mathbf{w}^\star) := \min \quad & \overline{\theta} + \sum_{\ell'\in\mathcal{X}_\ell} \phi^1_{\ell',\ell}\cdot\overline{\gamma}_{\ell'} && \\
s.t. \quad & \overline{\theta} + \sum_{\ell'\in B} \overline{\gamma}_{\ell'} \geq f_\ell(B) && \forall B\subseteq \mathcal{X}_\ell \\
& \overline{\theta},\, \overline{\gamma}_{\ell'} \in \mathbb{R}, && \forall \ell'\in\mathcal{X}_\ell.
\end{aligned} \tag{27}
$$

In Lemma 9, we characterize some useful properties of the dual problem $G^D_\ell(\mathbf{x}^\star,\mathbf{w}^\star)$. We defer its proof to Appendix B.3.1.

Lemma 9. *Let $G^D_\ell(\mathbf{x}^\star,\mathbf{w}^\star)$ be as defined in (27). Then, the following properties hold:*

1. *There is an optimal solution $(\overline{\theta}^\star,\overline{\gamma}^\star)$ such that $\overline{\theta}^\star$, $\overline{\gamma}^\star_{\ell'}\geq 0$, for all $\ell'\in\mathcal{X}_\ell$.*
2. *The feasible region can be equivalently restricted to constraints for $B\subseteq\mathcal{X}_\ell$ such that $|B|\leq K_\ell$.*
3. *There is an optimal solution $(\overline{\theta}^\star,\overline{\gamma}^\star)$ and a subset $B^\star\subseteq\mathcal{X}_\ell$ with $|B^\star|=K_\ell$ such that, for some $\alpha^\star\geq 0$, we have: (i) $\phi^2_{\ell,\ell'}-\overline{\gamma}^\star_{\ell'}=\alpha^\star$ for all $\ell'\in B^\star$, (ii) $\phi^2_{\ell,\ell'}-\overline{\gamma}^\star_{\ell'}\leq\alpha^\star$ for all $\ell'\notin B^\star$, and (iii) $\overline{\theta}^\star=K_\ell\alpha^\star$.*

One of the key elements in the proof of Theorem 1 is the following lemma, which shows that the optimal solution of the distribution problem provides an upper bound for the problem solved by DH (11).

Lemma 10. $F(\mathbf{x}^\star,\mathbf{w}^\star)\leq G(\mathbf{x}^\star,\mathbf{w}^\star)$.

Proof of Lemma 10. Note that it is enough to show that $F_\ell(\mathbf{x}^\star,\mathbf{w}^\star)\leq G_\ell(\mathbf{x}^\star,\mathbf{w}^\star)$ for all $\ell\in I\cup J$. Given a solution $(\mathbf{x}^\star,\mathbf{w}^\star)$, there are two possible cases.

1. If $|\mathcal{X}_\ell|\leq K_\ell$, we show that $F_\ell(\mathbf{x}^\star,\mathbf{w}^\star)=G_\ell(\mathbf{x}^\star,\mathbf{w}^\star)$. On the one hand, observe that the optimal solution in (24) is such that $y_{\ell'}=\phi^1_{\ell',\ell}$ for all $\ell'\in\mathcal{X}_\ell$ because $\sum_{\ell'\in\mathcal{X}_\ell} y_{\ell,\ell'}=\sum_{\ell'\in\mathcal{X}_\ell}\phi^1_{\ell',\ell}\leq|\mathcal{X}_\ell|\leq K_\ell$. Therefore, the optimal objective value in (24) is $F_\ell(\mathbf{x}^\star,\mathbf{w}^\star)=\sum_{\ell'\in\mathcal{X}_\ell}\phi^2_{\ell,\ell'}\phi^1_{\ell',\ell}$.
On the other hand, note that $|\mathcal{X}_\ell|\leq K_\ell$ implies that $f_\ell(B)=\sum_{\ell'\in\mathcal{X}_\ell}\phi^2_{\ell,\ell'}\mathbb{1}_{\{\ell'\in B\}}$ for all $B\subseteq\mathcal{X}_\ell$ since the backlog is smaller than the budget and, thus, the optimal decision in the second period is to show the entire backlog. Therefore, the optimal objective value in (25) is

$$
G_\ell(\mathbf{x}^\star,\mathbf{w}^\star)=\sum_{B\subseteq\mathcal{X}_\ell}\lambda_B\cdot f_\ell(B)=\sum_{B\subseteq\mathcal{X}_\ell}\lambda_B\cdot\sum_{\ell'\in\mathcal{X}_\ell}\phi^2_{\ell,\ell'}\mathbb{1}_{\{\ell'\in B\}}=\sum_{\ell'\in\mathcal{X}_\ell}\phi^2_{\ell,\ell'}\sum_{B:B\ni\ell'}\lambda_B=\sum_{\ell'\in\mathcal{X}_\ell}\phi^2_{\ell,\ell'}\phi^1_{\ell',\ell},
$$

where in the last equality we use the feasibility of λ. Thus, we conclude that $F_\ell(\mathbf{x}^\star,\mathbf{w}^\star)=G_\ell(\mathbf{x}^\star,\mathbf{w}^\star)$ in this case.

2. If $|\mathcal{X}_\ell|>K_\ell$, we show that $F_\ell(\mathbf{x}^\star,\mathbf{w}^\star)\leq G_\ell(\mathbf{x}^\star,\mathbf{w}^\star)$. To see this, first note that we can combine Lemma 9 (parts 1. and 2.) and the fact that $f_\ell(B)=\sum_{\ell'\in B}\phi^2_{\ell,\ell'}$ for any $|B|\leq K_\ell$ to rewrite the dual of $G_\ell(\mathbf{x}^\star,\mathbf{w}^\star)$ as

$$
\begin{aligned}
G^D_\ell(\mathbf{x}^\star,\mathbf{w}^\star) := \min \quad & \overline{\theta} + \sum_{\ell'\in\mathcal{X}_\ell} \phi^1_{\ell',\ell}\cdot\overline{\gamma}_{\ell'} && \\
s.t. \quad & \overline{\theta} + \sum_{\ell'\in B} \overline{\gamma}_{\ell'} \geq \sum_{\ell'\in B}\phi^2_{\ell,\ell'} && \forall B\subseteq \mathcal{X}_\ell,\ |B|\leq K_\ell \\
& \overline{\theta},\, \overline{\gamma}_{\ell'} \geq 0, && \forall \ell'\in\mathcal{X}_\ell.
\end{aligned} \tag{28}
$$

Let $\left(\bar{\theta}^\star, \bar{\gamma}^\star\right)$ be an optimal solution of (28) and $\alpha^\star$ be the corresponding constant as defined in Lemma 9 (part 3.). We can construct a feasible solution (θ, γ) in (26) by taking $\theta = \alpha^\star$ and $\gamma_{\ell'} = \overline{\gamma}^\star_{\ell'}$ for all $\ell' \in \mathcal{X}_\ell$. Clearly, these values are non-negative. Moreover, for each $\ell' \in \mathcal{X}_\ell$,

$$\theta + \gamma_{\ell'} = \alpha^\star + \overline{\gamma}^\star_{\ell'} \geq \phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} + \overline{\gamma}^\star_{\ell'} = \phi^2_{\ell,\ell'}.$$

Finally, the objective value of (θ, γ) is $K_\ell \alpha^\star + \sum_{\ell'} \phi^1_{\ell',\ell} \overline{\gamma}^\star_{\ell'}$, which is equal to objective value of $(\overline{\theta}^\star, \overline{\gamma}^\star)$ in (28). Hence, we know that

$$G_\ell(\mathbf{x}^\star, \mathbf{w}^\star) = G^D_\ell(\mathbf{x}^\star, \mathbf{w}^\star) = K_\ell \theta + \sum_{\ell' \in \mathcal{X}_\ell} \phi^1_{\ell',\ell} \gamma_{\ell'} \geq F^D_\ell(\mathbf{x}^\star, \mathbf{w}^\star) = F_\ell(\mathbf{x}^\star, \mathbf{w}^\star),$$

where in the first and last equalities we use strong duality.[19] Hence, we conclude that $F_\ell(\mathbf{x}^\star, \mathbf{w}^\star) \leq G_\ell(\mathbf{x}^\star, \mathbf{w}^\star)$ in this case.

Proof of Theorem 1. Let us recall for a moment that the original objective function of our problem can be re-written as

$$\mathcal{M}(\mathbf{x}, \mathbf{w}) = \mathcal{M}^1(\mathbf{x}, \mathbf{w}) + \mathcal{M}^2(\mathbf{x}, \mathbf{w}) = \sum_{e \in E} \beta^1_e \cdot w_e + \sum_{\ell \in I \cup J} \sum_{B \subseteq \mathcal{P}^1_\ell} f_\ell(B) \cdot \mathbb{P}_{\mathbf{x}^1}(B),$$

where $\mathbb{P}_{\mathbf{x}^1}(B) = \prod_{\ell' \in B} \phi^1_{\ell',\ell} \cdot x_{\ell',\ell} \prod_{\ell' \notin B} (1 - \phi^1_{\ell',\ell} \cdot x_{\ell',\ell})$ and $f_\ell(B) = \max_{S \subseteq B} \left\{ \sum_{\ell' \in S} \phi^2_{\ell,\ell'} : |S| \leq K_\ell \right\}$. In particular, given an optimal solution of DH (11) $(\mathbf{x}^\star, \mathbf{w}^\star, \mathbf{y}^\star)$, we can define $\lambda^{\mathsf{ind}}_{\ell,B} = \mathbb{P}_{\mathbf{x}^\star_\ell}(B)$, which is a feasible solution in (16). The correlation gap, introduced in (Agrawal et al. 2010), lower bounds the ratio between the objective value of the independent distribution $\lambda^{\mathsf{ind}}_{\ell,B}$ and the optimal value in (16). Formally, let $\lambda^\star_{\ell,B}$ be an optimal solution in (16). Since $f_\ell(\cdot)$ is a monotone submodular function for each ℓ, we know that

$$\frac{\sum_{\ell \in I \cup J} \sum_{B \subseteq \mathcal{P}^1_\ell} f_\ell(B) \cdot \lambda^{\mathsf{ind}}_{\ell,B}}{\sum_{\ell \in I \cup J} \sum_{B \subseteq \mathcal{P}^1_\ell} f_\ell(B) \cdot \lambda^\star_{\ell,B}} \geq 1 - 1/e.$$

Note that the numerator is $\mathcal{M}^2(\mathbf{x}^\star, \mathbf{w}^\star)$ (as defined in (6)) and the denominator is $G(\mathbf{x}^\star, \mathbf{w}^\star)$. Then, thanks to Lemma 10, we obtain

$$\mathcal{M}^2(\mathbf{x}^\star, \mathbf{w}^\star) \geq (1 - 1/e) \cdot G(\mathbf{x}^\star, \mathbf{w}^\star) \geq (1 - 1/e) \cdot F(\mathbf{x}^\star, \mathbf{w}^\star)$$

Finally, we conclude the proof by noting that

$$\mathcal{M}^1(\mathbf{x}^\star, \mathbf{w}^\star) + \mathcal{M}^2(\mathbf{x}^\star, \mathbf{w}^\star) \geq \mathcal{M}^1(\mathbf{x}^\star, \mathbf{w}^\star) + (1 - 1/e) \cdot F(\mathbf{x}^\star, \mathbf{w}^\star) \geq (1 - 1/e) \cdot \mathrm{OPT}' \geq (1 - 1/e) \cdot \mathrm{OPT}$$

where OPT' is the optimal value of (11) and, in the last inequality (i.e., $\mathrm{OPT}' \geq \mathrm{OPT}$), we use Lemma 11, which we prove in Appendix B.3.1.

Lemma 11. *Problem* (11) *is an upper bound of Problem 2.*

[19] Both duals are always feasible, since we can consider $\theta = \overline{\theta} = 0$ and $\gamma_{\ell'} = \overline{\gamma}_{\ell'} = \phi^2_{\ell,\ell'}$ for all $\ell' \in \mathcal{X}_\ell$.

B.3.1. Remaining Proofs of Technical Lemmas.

Proof of Lemma 9. We prove separately each point.

1. Let $(\overline{\theta}^\star, \overline{\gamma}^\star)$ be an optimal solution. The non-negativity of $\overline{\theta}_\ell^\star$'s results from taking $B=\emptyset$ in the constraint ($f_\ell(\emptyset)=0$). To prove the non-negativity of $\overline{\gamma}_{\ell'}^\star$ we note that in the optimal solution, there is a set $B^\star \subseteq \mathcal{X}_\ell$ such that $\overline{\theta}_\ell^\star = f_\ell(B^\star) - \sum_{\ell'\in B^\star}\overline{\gamma}_{\ell'}^\star$; which is the set that maximizes $f_\ell(B) - \sum_{\ell'\in B}\overline{\gamma}_{\ell'}^\star$. Denote by $C_\ell = \{\ell' \in \mathcal{X}_\ell : \overline{\gamma}_{\ell'}^\star < 0\}$. Let us redefine $\overline{\theta}_\ell' = \overline{\theta}_\ell^\star + \sum_{\ell'\in C_\ell}\phi_{\ell',\ell}\cdot\overline{\gamma}_{\ell'}^\star$. Clearly $\overline{\theta}_\ell' \geq 0$ since

$$\begin{aligned}\overline{\theta}_\ell' &= \overline{\theta}_\ell^\star + \sum_{\ell'\in C_\ell}\phi^1_{\ell',\ell}\cdot\overline{\gamma}_{\ell'}^\star \\ &= f_\ell(B^\star) - \sum_{\ell'\in B^\star}\overline{\gamma}_{\ell'}^\star + \sum_{\ell'\in C_\ell}\phi^1_{\ell',\ell}\cdot\overline{\gamma}_{\ell'}^\star \\ &\geq f_\ell(C_\ell) - \sum_{\ell'\in C_\ell}\overline{\gamma}_{\ell'}^\star + \sum_{\ell'\in C_\ell}\overline{\gamma}_{\ell'}^\star \\ &= f_\ell(C_\ell) \geq 0\end{aligned}$$

where we use that $B^\star$ is the maximizing set and $\phi^1_{\ell',\ell}\cdot\overline{\gamma}_{\ell'}^\star \geq \overline{\gamma}_{\ell'}^\star$ as $\overline{\gamma}_{\ell'}^\star < 0$ for $\ell'\in C_\ell$ and $\phi^1_{\ell',\ell} \leq 1$. Redefine $\overline{\gamma}_{\ell'}' = \overline{\gamma}_{\ell'}^\star$ for all $\ell' \notin C_\ell$ and zero otherwise. Note that the objective of $\overline{\theta}', \overline{\gamma}'$ is the same than the one of $\overline{\theta}^\star, \overline{\gamma}^\star$. Finally, the constraints are satisfied because, for any $B\subseteq\mathcal{X}_\ell$,

$$\begin{aligned}\overline{\theta}_\ell' + \sum_{\ell'\in B}\overline{\gamma}_{\ell'}' &= f_\ell(B^\star) - \sum_{\ell'\in B^\star}\overline{\gamma}_{\ell'}^\star + \sum_{\ell'\in C_\ell}\phi^1_{\ell',\ell}\cdot\overline{\gamma}_{\ell'}^\star + \sum_{\ell'\in B\setminus C_\ell}\overline{\gamma}_{\ell'}^\star \\ &\geq f_\ell(B\cup C_\ell) - \sum_{\ell'\in B\cup C_\ell}\overline{\gamma}_{\ell'}^\star + \sum_{\ell'\in C_\ell}\overline{\gamma}_{\ell'}^\star + \sum_{\ell'\in B\setminus C_\ell}\overline{\gamma}_{\ell'}^\star \\ &= f_\ell(B\cup C_\ell) \\ &\geq f_\ell(B),\end{aligned}$$

where in the first equality we use that $\sum_{\ell'\in B}\overline{\gamma}_{\ell'}' = \sum_{\ell'\in B\setminus C_\ell}\overline{\gamma}_{\ell'}^\star$ since $\overline{\gamma}_{\ell'}' = 0$ for $\ell'\in C_\ell$. The first inequality follows by the optimality of $B^\star$ i.e. $f_\ell(B^\star) - \sum_{\ell'\in B^\star}\overline{\gamma}_{\ell'}^\star \geq f_\ell(B\cup C_\ell) - \sum_{\ell'\in B\cup C_\ell}\overline{\gamma}_{\ell'}^\star$ and that $\sum_{\ell'\in C_\ell}\phi^1_{\ell',\ell}\cdot\overline{\gamma}_{\ell'}^\star \geq \sum_{\ell'\in C_\ell}\overline{\gamma}_{\ell'}^\star$. The last inequality is due to monotonicity of f_ℓ.

2. Let $(\overline{\theta}^\star, \overline{\gamma}^\star)$ be an optimal solution. Then, $\overline{\theta}^\star = \max_{B\subseteq\mathcal{X}_\ell}\left\{f_\ell(B) - \sum_{\ell'\in B}\overline{\gamma}_{\ell'}^\star\right\}$. Let $B^\star$ be the corresponding maximizer, i.e., $\overline{\theta} = f_\ell(B^\star) - \sum_{\ell'\in B^\star}\overline{\gamma}_{\ell'}^\star$, and let $S^\star \subseteq B^\star$ be an optimal solution in $f_\ell(B^\star)$, i.e., $|S^\star| \leq K_\ell$ and $f_\ell(B^\star) = \sum_{\ell'\in S^\star}\phi^2_{\ell,\ell'}$. Therefore, $\overline{\theta}^\star = \sum_{\ell'\in S^\star}\phi^2_{\ell,\ell'} - \sum_{\ell'\in B^\star}\overline{\gamma}_{\ell'}^\star$.
To find a contradiction, suppose that $|B^\star| > K_\ell$. Then, we know that $S^\star \subset B^\star$ and, since $\overline{\gamma}_{\ell'}^\star \geq 0$, we can remove terms in the second sum and potentially increase this difference, i.e.,

$$\sum_{\ell'\in S^\star}\phi^2_{\ell,\ell'} - \sum_{\ell'\in B^\star}\overline{\gamma}_{\ell'}^\star \leq \sum_{\ell'\in S^\star}\phi^2_{\ell,\ell'} - \sum_{\ell'\in S^\star}\overline{\gamma}_{\ell'}^\star,$$

contradicting the optimality of $B^\star$.

3. Let $(\overline{\theta}^\star, \overline{\gamma}^\star)$ be an optimal solution. By part 2. of this lemma, we know that there exists a subset $B^\star$ such that $|B^\star| \leq K_\ell$ and

$$\overline{\theta}^\star = \sum_{\ell'\in B^\star}(\phi^2_{\ell,\ell'} - \overline{\gamma}_{\ell'}^\star) \geq \max_{B\subseteq\mathcal{X}_\ell}\left\{f_\ell(B) - \sum_{\ell'\in B}\overline{\gamma}_{\ell'}^\star\right\}.$$

Observe that $\phi^2_{\ell,\ell'} - \overline{\gamma}_{\ell'}^\star \geq 0$, $\forall \ell'\in\mathcal{X}_\ell$. To see this, we argue by contradiction; suppose there exists $\ell'\in\mathcal{X}_\ell$ for which the difference is strictly negative (i.e., $\overline{\gamma}_{\ell'}^\star > \phi^2_{\ell,\ell'}$), then:

- If $\ell' \in B^\star$, we can remove it from $B^\star$ and easily show that $\overline{\theta}^\star$ does not satisfy the constraint for $B^\star \setminus \{\ell'\}$, which contradicts the feasibility of $(\overline{\theta}^\star, \overline{\gamma}^\star)$.
- If $\ell' \in \mathcal{X}_\ell \setminus B^\star$, we can create a new solution by decreasing $\overline{\gamma}^\star_{\ell'}$ until $\phi^2_{\ell,\ell'}$. This solution is still feasible since, for any B that contains ℓ', we have

$$\overline{\theta}^\star + \sum_{\ell'' \in B \setminus \ell'} \overline{\gamma}^\star_{\ell''} + \overline{\gamma}^\star_{\ell'} = \overline{\theta}^\star + \sum_{\ell'' \in B \setminus \ell'} \overline{\gamma}^\star_{\ell''} + \phi^2_{\ell,\ell'} \geq \sum_{\ell'' \in B \setminus \ell'} \phi^2_{\ell,\ell''} + \phi^2_{\ell,\ell'} = \sum_{\ell'' \in B} \phi^2_{\ell,\ell''}.$$

The feasibility for any B that does not contain ℓ' also holds. More importantly, this solution has a lower objective value, which contradicts the optimality of $(\overline{\theta}^\star, \overline{\gamma}^\star)$.

Finally, note that due to non-negativity we have $\phi^2_{\ell,\ell''} - \overline{\gamma}^\star_{\ell''} \leq \min_{\ell' \in B^\star} \{\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'}\}$, $\forall \ell' \in \mathcal{X} \setminus B^\star$, otherwise we could swap the corresponding terms and contradict the optimality of $B^\star$.

Given the properties above, without loss of generality, suppose that the indexes $\ell' \in \mathcal{X}_\ell$ are sorted in decreasing order of $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'}$, i.e.,

$$\phi^2_{\ell,1} - \overline{\gamma}^\star_1 \geq \phi^2_{\ell,2} - \overline{\gamma}^\star_2 \geq \cdots \geq \phi^2_{\ell,|\mathcal{X}_\ell|} - \overline{\gamma}_{|\mathcal{X}_\ell|} \geq 0.$$

By its optimality, we know that $B^\star$ consists of the first K_ℓ elements in this ordering, i.e., $B^\star = \{1, \ldots, K_\ell\}$. We now show that $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} = \phi^2_{\ell,\ell''} - \overline{\gamma}^\star_{\ell''} = \alpha^\star$ for any $\ell', \ell'' \in B^\star$. To find a contradiction, suppose that this does not hold. Then, there exists $\ell' \in \{1, \ldots, K_\ell - 1\}$ such that $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} > \phi^2_{\ell,\ell'+1} - \overline{\gamma}^\star_{\ell'+1}$. Let ℓ' be the smallest index that this happens. Let $\left(\overline{\theta}', \overline{\gamma}'\right)$ be such that $\overline{\theta}' = \overline{\theta}^\star - \ell' \cdot \epsilon$, $\overline{\gamma}'_{\ell''} = \overline{\gamma}^\star_{\ell''} + \epsilon$ for all $\ell'' \in \{1, \ldots, \ell'\}$, and $\overline{\gamma}'_{\ell''} = \overline{\gamma}^\star_{\ell''}$ for all $\ell'' \in \{\ell'+1, \ldots, |\mathcal{X}_\ell|\}$, where $\epsilon > 0$ is such that $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} - \epsilon = \phi^2_{\ell,\ell'+1} - \overline{\gamma}^\star_{\ell'+1}$. In words, this new solution shifts all the terms in $\{1, \ldots, \ell'\}$ to make them equal to the $(\ell'+1)$-th term.

Clearly, by feasibility of $\overline{\gamma}^\star$ and the fact that $\epsilon > 0$, we know that $\overline{\gamma}'$ is non-negative, while this also holds for $\overline{\theta}'$ since

$$\begin{aligned}
\overline{\theta}' &= \overline{\theta}^\star - \ell' \cdot \epsilon \\
&\geq \sum_{\ell''=1}^{\ell'} (\phi^2_{\ell,\ell''} - \overline{\gamma}^\star_{\ell''}) - \ell' \cdot \epsilon \\
&= \ell' \cdot (\phi^2_{\ell,\ell'+1} - \overline{\gamma}^\star_{\ell'+1}) \\
&\geq 0,
\end{aligned}$$

where the first inequality is due to the feasibility of $\overline{\theta}^\star$ and the fact that $\phi^2_{\ell,\ell''} - \overline{\gamma}^\star_{\ell''} \geq 0$, $\forall \ell'' \in \mathcal{X}_\ell$, while the last equality follows from the definition of ϵ.

Also, note that the constraints for $B \subseteq \mathcal{X}_\ell$ such that $\{1, \ldots, \ell'\} \subseteq B$ still hold since

$$\begin{aligned}
\overline{\theta}' &= \overline{\theta}^\star - \ell' \cdot \epsilon \\
&= \sum_{\ell'' \in B^\star} (\phi^2_{\ell,\ell''} - \overline{\gamma}^\star_{\ell''}) - \ell' \cdot \epsilon \\
&= \sum_{\ell'' \in B^\star \setminus \{1,\ldots,\ell'\}} (\phi^2_{\ell,\ell''} - \overline{\gamma}^\star_{\ell''}) + \sum_{\ell''=1}^{\ell'} (\phi^2_{\ell,\ell''} - \overline{\gamma}^\star_{\ell''}) - \ell' \cdot \epsilon \\
&= \sum_{\ell'' \in B^\star \setminus \{1,\ldots,\ell'\}} (\phi^2_{\ell,\ell''} - \overline{\gamma}'_{\ell''}) + \sum_{\ell''=1}^{\ell'} (\phi^2_{\ell,\ell''} - \overline{\gamma}'_{\ell''}) \\
&\geq \sum_{\ell'' \in B \setminus \{1,\ldots,\ell'\}} (\phi^2_{\ell,\ell''} - \overline{\gamma}'_{\ell''}) + \sum_{\ell''=1}^{\ell'} (\phi^2_{\ell,\ell''} - \overline{\gamma}'_{\ell''}),
\end{aligned}$$

where the inequality is because $B^\star$ is composed by top elements. The argument for $B \subseteq \mathcal{X}_\ell$ such that $\{1,\ldots,\ell'\} \cap (\mathcal{X}_\ell \setminus B) \neq \emptyset$ is analogous. Hence, $\left(\overline{\theta}', \overline{\gamma}'\right)$ is a feasible solution of Problem (28). Finally, note that this new feasible solution leads to the following objective

$$\begin{aligned}\overline{\theta}' + \sum_{\ell''=1}^{\ell'} \phi^1_{\ell'',\ell}\overline{\gamma}'_{\ell''} + \sum_{\ell'' \in \mathcal{X}_\ell \setminus \{1,\ldots,\ell'\}} \phi^1_{\ell'',\ell}\overline{\gamma}'_{\ell''} &= \overline{\theta}^\star - \ell' \cdot \epsilon + \sum_{\ell''=1}^{\ell'} \phi^1_{\ell'',\ell}(\overline{\gamma}^\star_{\ell''} + \epsilon) + \sum_{\ell'' \in \mathcal{X}_\ell \setminus \{1,\ldots,\ell'\}} \phi^1_{\ell'',\ell}\overline{\gamma}^\star_{\ell''} \\ &= \overline{\theta}^\star + \sum_{\ell'' \in \mathcal{X}_\ell} \phi^1_{\ell'',\ell}\overline{\gamma}^\star_{\ell''} - \epsilon \sum_{\ell''=1}^{\ell'} (1 - \phi^1_{\ell',\ell}) \\ &< \overline{\theta}^\star + \sum_{\ell'' \in \mathcal{X}_\ell} \phi^1_{\ell'',\ell}\overline{\gamma}^\star_{\ell''},\end{aligned}$$

which contradicts the optimality of $(\overline{\theta}^\star, \overline{\gamma}^\star)$. Therefore, we must have $\phi^2_{\ell,\ell'} - \overline{\gamma}_{\ell'} = \phi^2_{\ell,\ell'+1} - \overline{\gamma}_{\ell'+1}$ for all $\ell' \in \{1,\ldots,K_\ell\}$ and, thus, there exists $\alpha^\star \geq 0$ such that $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} = \alpha^\star$ for all $\ell' \in B^\star$ and $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} \leq \alpha^\star$ for all $\ell' \notin B^\star$ (by optimality of $B^\star$). Finally,

$$\overline{\theta}^\star = \sum_{\ell' \in B^\star} \phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} = |B^\star| \cdot \alpha^\star = K_\ell \cdot \alpha^\star.$$

Remark 5. The same argument applies if $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} = 0$, in which case $\alpha^\star = 0$.

Remark 6. If $\alpha^* > 0$, the argument does not necessarily apply beyond the first K_ℓ elements of the sequence (although we do not care about those elements). For example, suppose that the first $K_\ell + 1$ terms are such that $\phi^2_{\ell,\ell'} - \overline{\gamma}^\star_{\ell'} = \alpha$ and strictly greater than the $(K_\ell + 2)$-th term. Then, to make them all equal, we need to reduce $\overline{\theta}^\star = K_\ell \cdot \alpha$ by $(K_\ell + 1) \cdot \epsilon$ and, thus, the resulting $\overline{\theta}' = \overline{\theta}^\star - (K_\ell + 1) \cdot \epsilon$ may be negative.

Proof of Lemma 11. Let $(\mathbf{x}, \mathbf{w})$ be a feasible solution in Problem 2. Define $\hat{\mathbf{x}} = \mathbf{x}$, $\hat{\mathbf{w}} = \mathbf{w}$ and $\mathbf{y}$ as follows: for all $\ell \in I \cup J$, $\ell' \in \mathcal{P}^1_\ell$

$$\hat{y}_{\ell',\ell} = \sum_{\mathbf{B} \subseteq \vec{E}} \mathbb{1}_{\left\{\ell \in S_{\ell'}(B_{\ell'})\right\}} \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B})$$

where $S_{\ell'}(B_{\ell'})$ is an optimal solution for user ℓ' in the problem defined in $f(\mathbf{B})$. Clearly, $\hat{\mathbf{x}}, \hat{\mathbf{w}}$ satisfy their constraints in Problem (11), so it remains to show the constraints that involve $\hat{\mathbf{y}}$ are also satisfied. First, note that

$$\begin{aligned} y_{\ell',\ell} &= \sum_{\mathbf{B} \subseteq \vec{E}} \mathbb{1}_{\left\{\ell \in S_{\ell'}(\mathbf{B})\right\}} \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}) \\ &= \sum_{\mathbf{B}' \subseteq \vec{E} \setminus (\ell,\ell')} \mathbb{1}_{\left\{\ell \in S_{\ell'}(\mathbf{B}' \cup \{(\ell,\ell')\})\right\}} \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}' \cup (\ell,\ell')) \\ &= \sum_{\mathbf{B}' \subseteq \vec{E} \setminus (\ell,\ell')} \mathbb{1}_{\left\{\ell \in S_{\ell'}(\mathbf{B}' \cup (\ell,\ell'))\right\}} \cdot \mathbb{P}_{\hat{\mathbf{x}}}((\ell,\ell')) \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}') \\ &= \hat{x}_{\ell,\ell'} \cdot \phi_{\ell,\ell'} \cdot \sum_{\mathbf{B}' \subseteq \vec{E} \setminus (\ell,\ell')} \mathbb{1}_{\left\{\ell \in S_{\ell'}(\mathbf{B}' \cup (\ell,\ell'))\right\}} \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}') \\ &\leq \hat{x}_{\ell,\ell'} \cdot \phi^1_{\ell,\ell'} \end{aligned}$$

where the third equality follows by the independent choices made by the users, the fourth uses that $\mathbb{P}_{\hat{\mathbf{x}}}(\{(\ell,\ell')\}) = \hat{x}_{\ell,\ell'} \cdot \phi^1_{\ell,\ell'}$ and the last inequality follows because the sum is at most 1. We now focus on the cardinality constraints of $\hat{\mathbf{y}}$:

$$\begin{aligned}
\sum_{\ell \in \mathcal{P}_{\ell'}} y_{\ell',\ell} &= \sum_{\ell \in \mathcal{P}_{\ell'}} \sum_{\mathbf{B} \subseteq \vec{E}} \mathbb{1}_{\left\{\ell \in S_{\ell'}(\mathbf{B})\right\}} \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}) \\
&= \sum_{\mathbf{B} \subseteq \vec{E}} \sum_{\ell \in \mathcal{P}_{\ell'}} \mathbb{1}_{\left\{\ell \in S_{\ell'}(\mathbf{B})\right\}} \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}) \\
&= \sum_{\mathbf{B} \subseteq \vec{E}} \left[\sum_{k=0}^{K_{\ell'}-1} k \cdot \mathbb{P}_{\hat{\mathbf{x}}}\left(|B_{\ell'}| = k\right) + K_{\ell'} \cdot \mathbb{P}_{\hat{\mathbf{x}}}\left(|B_{\ell'}| \geq K_{\ell'}\right) \right] \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}) \\
&\leq \sum_{\mathbf{B} \subseteq \vec{E}} K_{\ell'} \cdot \mathbb{P}_{\hat{\mathbf{x}}}(\mathbf{B}) \\
&= K_{\ell'} \quad \square
\end{aligned}$$

Appendix C: Appendix to Section 5

C.1. Estimation of the like probabilities and sample for experiments

To estimate the probability that each user i likes a profile $j \in \mathcal{P}_i^t$, we use a logit model with user-fixed effects:

$$y_{ijt} = \alpha_i + \lambda_t + X'_{i,j}\beta + \epsilon_{ijt}. \tag{29}$$

The dependent variable, y_{ijt}, is a latent variable that is related to the like decision $\Phi^t_{i,j}$ according to

$$\Phi^t_{i,j} = \begin{cases} 1 & \text{if } y_{ijt} > 0, \\ 0 & \text{otherwise.} \end{cases}$$

We control for users' unobserved heterogeneity by including user fixed-effects, α_i. We also control for period-dependent unobservables by including period fixed-effects, λ_t. The third term on the right-hand side, $X'_{i,j}\beta$, controls for observable characteristics of profile j, and also for their interaction with user i's observable characteristics. Specifically, we control for the attractiveness score, age, height and education level (measured in a scale from 1 to 3) of the profile evaluated. In addition, for each of these variables we control for the square of the positive and negative difference between the value for the user evaluating and that of the profile evaluated. Finally, we also control for whether the users share the same race and religion. Finally, ϵ_{ijt} is an idiosyncratic shock that follows an extreme value distribution. In Table 2, we report the estimation results.

Using these coefficients, for each user i and profile $j \in \mathcal{P}_i^1$ we compute the probability $\phi_{i,j}$. In Figure 3 we plot the distribution of like probabilities separating by gender, estimated using the parameters from column (2) in Table 2. These are the probabilities we use on our simulation study.

In Table 3, we report summary statistics for the sample used in the numerical experiments (standard deviations in parenthesis).

Appendix D: Appendix to Section 6

D.1. Multiple Periods

D.1.1. Proof of Theorem 2. The first part of the analysis is similar to that of Theorem 1, i.e., we compare the duals of two formulations and use the concept of correlation gap. With a slight abuse of notation,

Table 2 Estimation Results

	(1)	(2)
Batch size	-0.004***	-0.004***
	(0.0004)	(0.0004)
Score	0.832***	0.832***
	(0.014)	(0.014)
Score - Positive difference	0.012***	0.012***
	(0.003)	(0.003)
Score - Negative difference	-0.011***	-0.011***
	(0.003)	(0.003)
Age	-0.026***	-0.026***
	(0.004)	(0.004)
Age - Positive difference	-0.002***	-0.002***
	(0.0004)	(0.0004)
Age - Negative difference	-0.0009**	-0.0009**
	(0.0005)	(0.0005)
Height	0.055***	0.055***
	(0.008)	(0.008)
Height - Positive difference	0.002***	0.002***
	(0.0006)	(0.0006)
Height - Negative difference	-0.004***	-0.004***
	(0.0006)	(0.0006)
Education level	0.060**	0.061**
	(0.024)	(0.024)
Education level - Positive difference	-0.001	-0.0008
	(0.014)	(0.014)
Education level - Negative difference	-0.077***	-0.077***
	(0.016)	(0.016)
Share religion	0.078***	0.078***
	(0.013)	(0.013)
Share race	0.457***	0.458***
	(0.030)	(0.030)
User	✓	✓
Date		✓
Observations	396,226	396,226
Pseudo R^2	0.386	0.386

Figure 3 Distribution of Like Probabilities

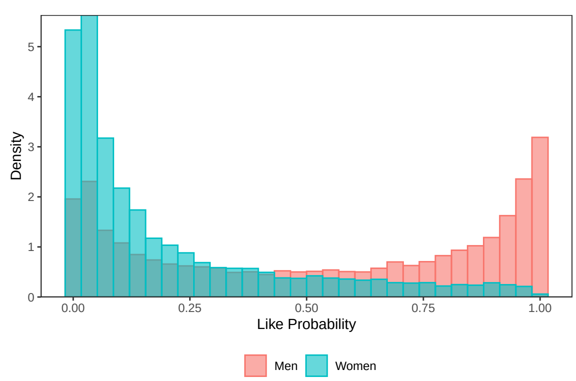


Table 3 Descriptives of Instance

	N	Score	Potentials	Like Prob.
Women	173	5.330	69.491	0.270
		(2.423)	(15.978)	(0.174)
Men	113	2.763	94.035	0.571
		(1.504)	(22.701)	(0.132)

we will use the same notation as in the proof of Theorem 1. Given an optimal solution $(\mathbf{x}^\star, \mathbf{w}^\star, \mathbf{y}^\star)$ of Problem (17), we define

$$
\begin{aligned}
F(\mathbf{x}^\star, \mathbf{w}^\star) := \max \quad & \sum_{\ell \in I \cup J} \sum_{\ell' \in \mathcal{P}_\ell^1} \phi_{\ell,\ell'} \cdot y_{\ell,\ell'} && (30)\\
s.t. \quad & y_{\ell,\ell'} \le x^\star_{\ell',\ell} \cdot \phi_{\ell',\ell}, && \forall \ell \in I \cup J,\ \ell' \in \mathcal{P}_\ell^1\\
& \sum_{\ell' \in \mathcal{P}_\ell^1} y_{\ell,\ell'} \le \tilde{K}_\ell^T, && \forall \ell \in I \cup J\\
& y_{\ell,\ell'} \ge 0, && \forall \ell \in I \cup J, \ell' \in \mathcal{P}_\ell^1.
\end{aligned}
$$

where $\tilde{K}_\ell^T := K_\ell \cdot T - \sum_{\ell' \in \mathcal{P}_\ell^1} x^\star_{\ell,\ell'} - \sum_{e \in E : \ell \in e} w^\star_e$ for each $\ell \in I \cup J$. We now define our second formulation:

$$
\begin{aligned}
G(\mathbf{x}^\star, \mathbf{w}^\star) := \max \quad & \sum_{\ell \in I \cup J} \sum_{L \subseteq \mathcal{P}_\ell^1} f_\ell^{[T]}(L) \cdot \lambda_{\ell,L} && (31)\\
s.t. \quad & \sum_{B \subseteq \mathcal{P}_\ell^1} \lambda_{\ell,L} = 1 && \forall \ell \in I \cup J\\
& \sum_{L \subseteq \mathcal{P}_\ell^1} \lambda_{\ell,L} = \phi_{\ell',\ell} \cdot x^\star_{\ell',\ell}, && \forall \ell \in I \cup J, \ell' \in \mathcal{P}_\ell^1\\
& \lambda_{\ell,L} \ge 0, && \forall \ell \in I \cup J,\ L \subseteq \mathcal{P}_\ell^1.
\end{aligned}
$$

where

$$
f_\ell^{[T]}(L) := \max \left\{ \sum_{\ell' \in L} \phi_{\ell,\ell'} \cdot z_{\ell,\ell'} : \sum_{\ell' \in L} z_{\ell,\ell'} \le \tilde{K}_\ell^T,\ z_{\ell,\ell'} \le \mathbb{1}_{\{\ell' \in L\}},\ z_{\ell,\ell'} \ge 0 \right\}. \tag{32}
$$

Analogous to the analysis of Theorem 1, we can obtain the following result:

LEMMA 12. $F(\mathbf{x}^\star, \mathbf{w}^\star) \le G(\mathbf{x}^\star, \mathbf{w}^\star)$

We omit the proof of this result to minimize redundancy.

Proof of Theorem 2. Consider an optimal solution $(\mathbf{x}^\star, \mathbf{w}^\star, \mathbf{y}^\star)$ of Problem (17). Let us denote by $\mathbf{L} = \{L_\ell\}_{\ell \in I \cup J}$ the random set of likes that result from showing profiles in $\mathbf{x}^\star$, i.e., $\mathbf{L} = \left\{\ell' : x^\star_{\ell,\ell'} = 1,\ \Phi^t_{\ell,\ell'} = 1 \text{ in some period } t \in [T],\ \forall \ell \in I \cup J\right\}$. Given the independence of users' decisions, the distribution of $\mathbf{L}$ is such that for every $\ell' \in I \cup J$

$$
\lambda^{\mathsf{ind}}_{\ell',L} = \prod_{\ell \in L} \phi_{\ell,\ell'} x^\star_{\ell,\ell'} \prod_{\ell \notin L} (1 - \phi_{\ell,\ell'} x^\star_{\ell,\ell'}).
$$

Since the function $f_\ell^{[T]}$ defined in (32) is monotone and submodular, then (Agrawal et al. 2010) shows that

$$
\frac{\sum_{\ell \in I \cup J} \sum_{L \subseteq \mathcal{P}_\ell^1} f_\ell^{[T]}(L) \cdot \lambda^{\mathsf{ind}}_{\ell,L}}{\sum_{\ell \in I \cup J} \sum_{L \subseteq \mathcal{P}_\ell^1} f_\ell^{[T]}(L) \cdot \lambda^\star_{\ell,L}} \ge 1 - \frac{1}{e}, \tag{33}
$$

where $\lambda^\star$ is an optimal solution of Problem (31). Note that

$$
\sum_{\ell \in I \cup J} \sum_{L \subseteq \mathcal{P}_\ell^1} f_\ell^{[T]}(L) \cdot \lambda^\star_{\ell,L} = G(\mathbf{x}^\star, \mathbf{w}^\star) \ge F(\mathbf{x}^\star, \mathbf{w}^\star),
$$

in which we use Lemma 12. Therefore, we can conclude that

$$
\sum_{\ell \in I \cup J} \sum_{L \subseteq \mathcal{P}_\ell^1} f_\ell^{[T]}(L) \cdot \lambda^{\mathsf{ind}}_{\ell,L} \ge (1 - 1/e) \cdot F(\mathbf{x}^\star, \mathbf{w}^\star). \tag{34}
$$

The final step in our proof is to show that the expected number of sequential matches achieved by Algorithm 2 is lower bounded by the term on the left in (34).

Recall the definitions in Algorithm 2, i.e., $X_\ell = \{\ell' \in \mathcal{P}^1_\ell : x^\star_{\ell,\ell'} = 1\}$ and $W_\ell = \{e \in E : \ell \in e,\ w^\star_e = 1\}$ for all $\ell \in I \cup J$. Note that Algorithm 2 exhausts all the profiles in $X_\ell \cup W_\ell$ for every $\ell \in I \cup J$. Given this, we can first conclude that the expected number of non-sequential matches achieved by our method is $\sum_{e\in E} \beta_e w^\star_e$, which coincides with the non-sequential part of the optimal objective in Problem (17).

We now focus on analyzing the expected number of sequential matches obtained by our method. For every $t \in [T]$, let $X^t_\ell = X_\ell \cap S^t_\ell$ be the set of profiles displayed to ℓ in period t as the initiating side of a sequential interaction. Similarly, let $W^t_\ell = W_\ell \cap S^t_\ell$ be the set of profiles shown to ℓ in period t as part of non-sequential interaction. Consider a sample path of Algorithm 2 and fix the realization of $\Phi^t_{\ell,\ell'}$ for every $\ell' \in X^t_\ell$ and $t \in [T]$. This implies that the backlog in each period $t \in [T]$ for each user $\ell \in I \cup J$ is a deterministic set B^t_ℓ. Denote by Z^t_ℓ the set of profiles shown from the backlog B^t_ℓ of user ℓ in stage t; note that Z^t_ℓ is deterministic (as B^t_ℓ is) and results from solving the problem in Step 14 in Algorithm 2. Finally, as introduced in Section 3, let L^t_ℓ be the set of users that liked ℓ in period t, i.e., $L^t_\ell = \{\ell' : \ell \in X^t_{\ell'}, \Phi^t_{\ell',\ell} = 1\}$. In a slight abuse of notation, let $L_\ell = \bigcup_{t\in[T]} L^t_\ell$ be the total set of profiles that liked ℓ during T periods. Note that, for any $t \in [T]$, the backlog satisfies $B^t_\ell = \bigcup_{\tau<t} L^\tau_\ell \setminus \left(\bigcup_{\tau<t} Z^\tau_\ell\right)$.

For every $\ell \in I \cup J$, the total expected number of sequential matches achieved by Algorithm 2 is

$$\sum_{t\in[T]} f^t_\ell(B^t_\ell) = \sum_{t\in[T]} \sum_{\ell' \in Z^t_\ell} \phi_{\ell',\ell} \tag{35}$$

where f^t_ℓ is as defined in (4) considering $K_\ell - |X^t_\ell| - |W^t_\ell|$ as the right-hand side of the capacity constraint. Since we include profiles in each S^t_ℓ from X_ℓ in decreasing order of $\phi_{\ell',\ell}$ (Step 7), then $\bigcup_{t\in[T]} Z^t_\ell$ contains the profiles with the highest values $\phi_{\ell',\ell}$ available in L_ℓ. This is crucial when we compare our method with the value of Problem (32) for the same set of likes L_ℓ.

For the total set of likes L_ℓ that ℓ received, let $\tilde{Z}_\ell$ be an optimal set in problem $f^{[T]}_\ell(L_\ell)$ as defined in (32). Construct a partition of $\tilde{Z}_\ell$ in T sets $\tilde{Z}^1_\ell, \ldots, \tilde{Z}^T_\ell$ as follows: (i) profiles in $\tilde{Z}_\ell$ are added to the sets in decreasing order of values $\phi_{\ell,\ell'}$, (ii) the set $\tilde{Z}^t_\ell$ is filled before continuing to $\tilde{Z}^{t+1}_\ell$, (iii) for each $t \in [T]$ we have $|\tilde{Z}^t_\ell| \le K_\ell - |X^t_\ell| - |W^t_\ell|$ and a profile ℓ' is added to $\tilde{Z}^t_\ell$ only if was not added before and $\ell' \in B^t_\ell$. This is possible because $L_\ell = \bigcup_{t\in[T]} B^t_\ell$ and

$$\sum_{t\in[T]} |\tilde{Z}^t_\ell| \le \sum_{t\in[T]} K_\ell - |X^t_\ell| - |W^t_\ell| = K_\ell \cdot T - \sum_{t\in[T]} |X^t_\ell| - \sum_{t\in[T]} |W^t_\ell| = K_\ell \cdot T - \sum_{\ell' \in \mathcal{P}^1_\ell} x^\star_{\ell,\ell'} - \sum_{e\in E:\ell\in e} w^\star_e = \tilde{K}^T_\ell,$$

where $\tilde{K}^T_\ell$ is the right-hand side of the cardinality constraint in $f^{[T]}_\ell(L_\ell)$. Note that the second equality follows as the profiles in $X_\ell \cup W_\ell$ are exhausted.

Claim 1. *For every $\ell \in I \cup J$, $t \in [T]$ we have $Z^t_\ell = \tilde{Z}^t_\ell$.*

Proof of Claim. First, note that $Z^t_\ell = \emptyset$ if and only if $B^t_\ell = \emptyset$ or $K_\ell = |X^t_\ell| + |W^t_\ell| = |S^t_\ell|$, which means that $\tilde{Z}^t_\ell = \emptyset$ due to the construction above. Second, suppose that $Z^t_\ell \neq \emptyset$ which means that $B^t_\ell \neq \emptyset$ and $K_\ell > |X^t_\ell| + |W^t_\ell|$. Since Z^t_ℓ is an optimal set for the problem defined by $f^t_\ell(B^t_\ell)$, then Z^t_ℓ contains the profiles with highest values $\phi_{\ell,\ell'}$ that are available in B^t_ℓ. Due to Step 7 of Algorithm 2, the profiles in Z^t_ℓ have higher values than profiles in Z^{t+1}_ℓ. Therefore, because of our construction was over the same set of likes L_ℓ then the sets $\tilde{Z}^t_\ell$ must coincide with Z^t_ℓ. □

The above claim implies that the total expected number of sequential matches is such that

$$\sum_{t\in[T]} f_\ell^t(B_\ell^t) = \sum_{t\in[T]} \sum_{\ell'\in Z_\ell^t} \phi_{\ell,\ell'} = \sum_{t\in[T]} \sum_{\ell'\in \tilde{Z}_\ell^t} \phi_{\ell,\ell'} = \sum_{\ell'\in\tilde{Z}_\ell} \phi_{\ell,\ell'} = f_\ell^{[T]}(L_\ell).$$

Taking expectation over the set of likes L_ℓ and summing over $\ell \in I \cup J$ we obtain

$$\mathbb{E}\left[\sum_{\ell\in I\cup J}\sum_{t\in[T]} f_\ell^t(B_\ell^t)\right] = \sum_{\ell\in I\cup J}\sum_{L_\ell\subseteq\mathcal{P}_\ell^1} f_\ell^{[T]}(L_\ell)\cdot\lambda_{\ell,L_\ell}^{\mathsf{ind}} \geq (1-1/e)\cdot F(\mathbf{x}^\star,\mathbf{w}^\star)$$

By summing non-sequential and sequential matches, we get

$$\begin{aligned}
\sum_{e\in E}\beta_e w_e^\star + \mathbb{E}\left[\sum_{\ell\in I\cup J}\sum_{t\in[T]} f_\ell^t(B_\ell^t)\right] &\geq \sum_{e\in E}\beta_e w_e^\star + (1-1/e)\cdot F(\mathbf{x}^\star,\mathbf{w}^\star)\\
&\geq (1-1/e)\cdot\left[\sum_{e\in E}\beta_e w_e^\star + F(\mathbf{x}^\star,\mathbf{w}^\star)\right]\\
&\geq (1-1/e)\cdot \mathrm{OPT}_{(17)}\\
&\geq (1-1/e)\cdot \mathrm{OPT}^{\Pi},
\end{aligned}$$

where $\mathrm{OPT}_{(17)}$ denotes the optimal value of Problem (17) and OPT^{Π} is the optimal value achieved by the best semi-adaptive policy. In the last inequality, we use Lemma 3.

D.1.2. Proof of Theorem 3. Fix the initiating side to be I. Since we restrict the space of policies to one-directional interactions and sequential matches, then our main relaxation is Problem (17) with $x_{j,i}=0$ for all $j\in J$, $i\in\mathcal{P}_j^1$ (i.e., users in J do not initiate interactions), $y_{i,j}=0$ for all $i\in I$ and $j\in\mathcal{P}_i^1$ (i.e., users in I are not followers in an interaction), $w_e=0$ for all $e\in E$ (i.e., there are no non-sequential interactions) and relaxing the integrality for the remaining variables, i.e., $x_{i,j}\in[0,1]$ for all $i\in I$, $j\in\mathcal{P}_i^1$. The resulting formulation is

$$\begin{aligned}
\max\quad & \sum_{i\in I}\sum_{j\in\mathcal{P}_i^1} y_{j,i}\cdot\phi_{j,i}\\
\text{s.t.}\quad & y_{j,i}\leq x_{i,j}\cdot\phi_{i,j}^1, \qquad \forall i\in I,\ j\in\mathcal{P}_i^1,\\
& \sum_{j\in\mathcal{P}_i^1} x_{i,j}\leq K_i\cdot T, \quad \forall i\in I,\\
& \sum_{i\in\mathcal{P}_j^1} y_{j,i}\leq K_j\cdot T, \quad \forall j\in J,\\
& \mathbf{x}\in[0,1]^{\vec{E}_I},\ \mathbf{y}\in[0,1]^{\vec{E}_J}
\end{aligned}\tag{36}$$

Lemma 13. *Problem* (36) *is a relaxation of Problem 1 for any adaptive policy with one-directional sequential interactions.*

Our main method is similar to Algorithm 2, but we include a rounding procedure since an optimal solution $(\mathbf{x}^\star,\mathbf{y}^\star)$ of Problem (36) may be fractional. Specifically, we use the dependent randomized rounding procedure introduced in (Gandhi et al. 2006), as we did to prove Proposition 4. This method outputs a random binary vector $\tilde{\mathbf{x}}\in\{0,1\}^{\vec{E}_I}$ such that: (i) $\sum_{j\in\mathcal{P}_i^1}\tilde{x}_{i,j}\leq K_i\cdot T$ for all $i\in I$ with probability 1; (ii) $\mathbb{E}[\tilde{x}_{i,j}]=x_{i,j}^\star$ for all $i\in I$, $j\in\mathcal{P}_j^1$. Finally, based on $\tilde{x}$, we proceed as in Algortihm 2 to decide the subsets of profiles to show in each period.

Proof of Theorem 3. Fix the random binary vector $\tilde{\mathbf{x}}$ obtained from the rounding. The proof from here is analogous to that of Theorem 2. Therefore, we obtain that the total expected number of sequential matches achieved by our method is such that

$$\mathbb{E}\left[\sum_{j\in J}\sum_{t\in[T]} f_j^t(B_j^t)\right] = \sum_{j\in J}\sum_{L_j\subseteq\mathcal{P}_j^1} f_j^{[T]}(L_j)\cdot\lambda_{j,L_j}^{\mathsf{ind}}$$

where $f_j^{[T]}(L_j)$ is defined as in (32) but with $\tilde{K}_j^T = K_j\cdot T$, the expectation is over the randomness of likes/dislikes and

$$\lambda_{j,L_j}^{\mathsf{ind}} = \prod_{i\in L_j}\phi_{i,j}\tilde{x}_{i,j}\prod_{i\notin L_j}(1-\phi_{i,j}\tilde{x}_{i,j}).$$

By taking expectation over the randomness of $\tilde{\mathbf{x}}$, we obtain

$$\begin{aligned}\mathbb{E}_{\tilde{\mathbf{x}}\sim\mathbf{x}^\star,\mathbf{B}}\left[\sum_{j\in J}\sum_{t\in[T]} f_j^t(B_j^t)\right] &= \sum_{j\in J}\sum_{L_j\subseteq\mathcal{P}_j^1} f_j^{[T]}(L_j)\cdot\mathbb{E}_{\tilde{\mathbf{x}}\sim\mathbf{x}^\star}[\lambda_{j,L_j}^{\mathsf{ind}}]\\ &= \sum_{j\in J}\sum_{L_j\subseteq\mathcal{P}_j^1} f_j^{[T]}(L_j)\cdot\prod_{i\in L_j}\phi_{i,j}x_{i,j}^\star\prod_{i\notin L_j}(1-\phi_{i,j}x_{i,j}^\star).\end{aligned}$$

where in the last equality we used the independence of like/dislike decisions of each user i and $\mathbb{E}_{\tilde{\mathbf{x}}\sim\mathbf{x}^\star}[\tilde{x}_{i,j}] = x_{i,j}^\star$ which is the property of the rounding algorithm.

To finish the proof, note that thanks to the correlation gap we have

$$\sum_{j\in J}\sum_{L_j\subseteq\mathcal{P}_j^1} f_j^{[T]}(L_j)\cdot\prod_{i\in L_j}\phi_{i,j}x_{i,j}^\star\prod_{i\notin L_j}(1-\phi_{i,j}x_{i,j}^\star)\geq(1-1/e)\cdot F(\mathbf{x}^\star),$$

where $F(\mathbf{x}^\star)$ is (30) with $\tilde{K}_j^T = K_j\cdot T$ and $y_{i,j}=0$ for all $i\in I,\ j\in\mathcal{P}_i^1$. Finally, $F(\mathbf{x}^\star)\geq\mathrm{OPT}^\Pi$ follows from Lemma 13, where OPT^Π is the optimal value achieved by an adaptive policies with one-directional interactions. $\square$

D.1.3. Remaining Proofs of Technical Lemmas.

Proof of Lemma 3. Consider an optimal semi-adaptive policy $\pi^\star$. Since $\pi^\star$ decides non-adaptively the profiles that initiate sequential interactions and the profiles of non-sequential interactions, then define $x_{\ell,\ell'}=1$ if $\pi^\star$ shows profile ℓ' to ℓ in some stage (as the initiation of a sequential interaction), and zero otherwise. Similarly, let $w_e=1$ if the users in e mutually see each other in some period, and zero otherwise. Let Ω be the set of sample paths. For every $\omega\in\Omega$, let $y_{\ell,\ell'}^\omega=1$ if the policy displays profile ℓ' to user ℓ from the backlog in some period in path ω, and zero otherwise. Finally, define $y_{\ell,\ell'}=\sum_{\omega\in\Omega}y_{\ell,\ell'}^\omega\cdot\mathbb{P}(\omega)$, where $\mathbb{P}(\omega)$ is the probability of sample path ω, which is completely determined by the like/dislike realizations.

Let us prove that these variables are feasible in Problem (17). Clearly, we have $x_{\ell,\ell'}+x_{\ell',\ell}+w_e\leq 1$ since the policy shows a profile either as a sequential interaction or as a non-sequential interaction. For every $\ell\in I\cup J$ and $\ell'\in\mathcal{P}_\ell^1$ we have

$$\begin{aligned}y_{\ell,\ell'} &= \sum_{\omega\in\Omega}y_{\ell,\ell'}^\omega\cdot\mathbb{P}(\omega)\\ &\leq\sum_{\omega\in\Omega}\mathbb{1}_{\{\ell'\text{ liked }\ell\text{ in }\omega\}}\cdot x_{\ell',\ell}\cdot\mathbb{P}(\omega)\\ &= x_{\ell',\ell}\cdot\sum_{\omega\in\Omega}\mathbb{1}_{\{\ell'\text{ liked }\ell\text{ in }\omega\}}\cdot\mathbb{P}(\omega)\\ &\leq\phi_{\ell',\ell}x_{\ell',\ell}.\end{aligned}$$

Finally, we have that for every $\ell \in I \cup J$

$$\sum_{\ell' \in \mathcal{P}^1_\ell} x_{\ell,\ell'} + \sum_{e \in E: \ell \in e} w_e + \sum_{\ell' \in \mathcal{P}^1_\ell} y_{\ell,\ell'} = \sum_{\omega \in \Omega} \Big[\sum_{\ell' \in \mathcal{P}^1_\ell} x_{\ell,\ell'} + \sum_{e \in E: \ell \in e} w_e + \sum_{\ell' \in \mathcal{P}^1_\ell} y^\omega_{\ell,\ell'} \Big] \cdot \mathbb{P}(\omega)$$
$$\leq K_\ell \cdot T \cdot \sum_{\omega \in \Omega} \mathbb{P}(\omega) = K_\ell \cdot T,$$

where in the inequality we used that, for every path, we show at most K_ℓ in every period and there are T periods. $\square$

Proof of Lemma 13. Consider an optimal adaptive policy $\pi^\star$ with one-directional interactions. Let Ω be the set of sample paths. For every $\omega \in \Omega$, let $x^\omega_{i,j} = 1$ if the policy displays profile j to user i in some period in path ω, and zero otherwise, and let $x_{i,j} = \sum_{\omega \in \Omega} x^\omega_{i,j} \cdot \mathbb{P}(\omega)$, where $\mathbb{P}(\omega)$ is the probability of sample path ω. Similarly, let $y^\omega_{i,j} = 1$ if the policy displayed profile i to user j from the backlog in some period in path ω, and zero otherwise. Finally, define $y_{j,i} = \sum_{\omega \in \Omega} y^\omega_{j,i} \cdot \mathbb{P}(\omega)$. Let us prove that these variables are feasible in Problem (36).

First, for every $i \in I$ we have

$$\sum_{j \in \mathcal{P}^1_i} x_{i,j} = \sum_{\omega \in \Omega} \sum_{j \in \mathcal{P}^1_i} x^\omega_{i,j} \cdot \mathbb{P}(\omega) \leq K_i \cdot T \cdot \sum_{\omega \in \Omega} \mathbb{P}(\omega) = K_i \cdot T,$$

where the first inequality is because in every sample path we have $\sum_{j \in \mathcal{P}^1_i} x^\omega_{i,j} \leq K_i \cdot T$ as there are T periods and at most K_i profiles are shown per period. In a similar fashion, we can prove that $\sum_{i \in \mathcal{P}^1_j} y_{j,i} \leq K_j \cdot T$.

Finally, for every $j \in J$ and $i \in \mathcal{P}^1_j$ we have

$$\begin{aligned} y_{j,i} &= \sum_{\omega \in \Omega} y^\omega_{j,i} \cdot \mathbb{P}(\omega) \\ &\leq \sum_{\omega \in \Omega} \mathbb{1}_{\{i \text{ liked } j \text{ in } \omega\}} \cdot x^\omega_{i,j} \cdot \mathbb{P}(\omega) \\ &\leq \phi_{i,j} \sum_{\omega \in \Omega_{\text{-}ij}} x^\omega_{i,j} \cdot \mathbb{P}_{\text{-}ij}(\omega) \\ &= \phi_{i,j} x_{i,j}, \end{aligned}$$

where in the first inequality we use that $y^\omega_{j,i} \leq \mathbb{1}_{\{i \text{ liked } j \text{ in } \omega\}} \cdot x^\omega_{i,j}$. In the following inequality, we remove the independent event that i liked j from the sample path space $\Omega_{\text{-}ij}$ and from the probability $\mathbb{P}_{\text{-}ij}(\omega)$. In the last equality, we use that the decision of $\pi^\star$ to show j to i is independent of the realization of $\Phi_{i,j}$ since the policy doesn't have access to realizations beforehand. $\square$

D.2. One-Directional Interactions and Non-Sequential Matches in both Periods

D.2.1. Proof of Theorem 4. Recall function $\mathcal{M}^2$ defined for $(\mathbf{x}^1, \mathbf{w}^1)$ as

$$\mathcal{M}^2(\mathbf{x}^1, \mathbf{w}^1) = \sum_{\mathbf{B} \subseteq \vec{E}} f(\mathbf{B}) \cdot \mathbb{P}_{\mathbf{x}^1}(\mathcal{B} = \mathbf{B}).$$

This sum is over all possible backlogs, i.e., all edges that belong to $\vec{E}$. Note that $\mathbb{P}_{\mathbf{x}^1}(\mathcal{B} = \mathbf{B})$ "restricts" the valid backlogs that come from the first-period profiles $\mathbf{x}^1$, i.e., this probability will be zero for any B that contains an element $e \in \vec{E}$ with $x^1_e = 0$. This means that, effectively, we are summing over subsets of $\vec{E}(\mathbf{x}^1) = \{e \in \vec{E} : \ x^1_e = 1\}$.

LEMMA 14. *The function $\mathcal{M}^2$ can be reformulated as*

$$\mathcal{M}^2(\mathbf{x},\mathbf{w}) = \sum_{\mathbf{B}\subseteq\vec{E}} f(\mathbf{B}\cap\vec{E}(\mathbf{x}))\cdot\mathbb{P}(\mathcal{B}=\mathbf{B}),$$

where $\mathbb{P}(\mathcal{B}=\mathbf{B}) = \prod_{e\in\mathbf{B}}\phi_e^1\prod_{\vec{E}\setminus\mathbf{B}}(1-\phi_e^1)$.

Note that in the result above the distribution of backlogs does not depend on the first-period decisions.

Proof of Lemma 14. First, note that for any $\mathbf{L}\subseteq\vec{E}(\mathbf{x})$ and any $\mathbf{B}\subseteq\vec{E}$ such that $\mathbf{L}=\mathbf{B}\cap\vec{E}(\mathbf{x})$ we have $f(\mathbf{B})=f(\mathbf{L})$. Then,

$$\begin{aligned}\mathcal{M}^2(\mathbf{x},\mathbf{w}) &= \sum_{\mathbf{B}\subseteq\vec{E}} f(\mathbf{B}\cap\vec{E}(\mathbf{x}))\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\\ &= \sum_{\mathbf{L}\subseteq\vec{E}(\mathbf{x})}\ \sum_{\substack{\mathbf{B}\subseteq\vec{E}:\\ \mathbf{B}\cap\vec{E}(\mathbf{x})=\mathbf{L}}} f(\mathbf{B}\cap\vec{E}(\mathbf{x}))\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\\ &= \sum_{\mathbf{L}\subseteq\vec{E}(\mathbf{x})}\ \sum_{\substack{\mathbf{B}\subseteq\vec{E}:\\ \mathbf{B}\cap\vec{E}(\mathbf{x})=\mathbf{L}}} f(\mathbf{L})\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\\ &= \sum_{\mathbf{L}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{L}) \sum_{\substack{\mathbf{B}\subseteq\vec{E}:\\ \mathbf{B}\cap\vec{E}(\mathbf{x})=L}} \cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\\ &= \sum_{\mathbf{L}\subseteq\vec{E}(\mathbf{x})} f(\mathbf{L})\prod_{e\in\mathbf{L}}\phi_e^1 x_e \prod_{e\in\vec{E}(\mathbf{x})\setminus\mathbf{L}}(1-\phi_e^1 x_e).\end{aligned}$$

Let $\tilde{f}(\mathbf{x}^1,\mathbf{w}^1,\mathbf{B})$ be the solution of (3) (note that $\mathcal{P}_\ell^2$ is dependent on $\mathbf{x}^1,\mathbf{w}^1$). Following a similar argument as in the proof of Lemma 14, we can define $\tilde{\mathcal{M}}^2(\mathbf{x}^1,\mathbf{w}^1)$ as

$$\tilde{\mathcal{M}}^2(\mathbf{x},\mathbf{w}) = \sum_{\mathbf{B}\subseteq\vec{E}} \tilde{f}\left(\mathbf{x},\mathbf{w}^1,\mathbf{B}\cap\vec{E}(\mathbf{x})\right)\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})$$

Now, we are ready to prove our main result.

Proof of Theorem 4. For simplicity, we show the result for the case when $K_\ell=1$ for all $\ell\in I\cup J$; for the general case, the analysis is similar since we can do it per pair of users. Let $(\mathbf{x}^{1,\star},\mathbf{w}^{1,\star})$ and $(\tilde{\mathbf{x}}^{1,\star},\tilde{\mathbf{w}}^{1,\star})$ be the optimal solutions of the two-period version of Problem 1 excluding and enabling non-sequential matches in the second period, leading to OPT_1 and OPT_2 expected matches, respectively. To ease the exposition, we will drop super-indices of the variables.

Our goal is to lower bound the following ratio

$$\frac{\mathrm{OPT}_1}{\mathrm{OPT}_2} = \frac{\sum_{e\in E} w_e\cdot\beta_e + \mathcal{M}^2(\mathbf{x},\mathbf{w})}{\sum_{e\in E}\tilde{w}_e\cdot\beta_e + \tilde{\mathcal{M}}^2(\tilde{\mathbf{x}},\tilde{\mathbf{w}})}.$$

By using Lemma 14, we redefine both objectives to consider backlog distributions that are independent of the decisions in the first period, i.e.,

$$\frac{\sum_{e\in E} w_e\cdot\beta_e + \mathcal{M}^2(\mathbf{x},\mathbf{w})}{\sum_{e\in E}\tilde{w}_e\cdot\beta_e + \tilde{\mathcal{M}}^2(\tilde{\mathbf{x}},\tilde{\mathbf{w}})} = \frac{\sum_{\mathbf{B}}\left(\sum_{e\in E} w_e\beta_e + f(\mathbf{B}\cap\vec{E}(\mathbf{x}))\right)\mathbb{P}(\mathcal{B}=\mathbf{B})}{\sum_{\mathbf{B}}\left(\sum_{e\in E}\tilde{w}_e\cdot\beta_e + \tilde{f}(\mathbf{B}\cap\vec{E}(\tilde{\mathbf{x}}),\tilde{\mathbf{x}},\tilde{\mathbf{w}})\right)\mathbb{P}(\mathcal{B}=\mathbf{B})}. \tag{37}$$

To lower bound this ratio, we now proceed to lower bound the expected contribution of each pair $(i,j)\in I\times J$.

Given that we are assuming one directional interactions with sequential matches going from I to J, the expected contribution of the pair (i,j) in OPT_1 depends on the first-period decisions:

- If $x_{i,j}=1$, then the contribution of the pair (i,j) in the numerator is

$$\Delta_{ij}=\phi_{ij}\phi_{ji}\sum_{\mathbf{B}:(i,j)\notin\mathbf{B}}\mathbb{1}_{\left\{i\in S_j(\mathbf{B},\mathbf{x})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})$$

where $S_j(\mathbf{B},\mathbf{x})$ is an optimal solution for user j in objective f of the second period for B and $\mathbf{x}$. Δ_{ij} is the product of: the probability that i liked j, the probability that i was shown to j in the second stage and the probability that j liked i. Observe that if no other agent $i'\in I$ such that $x_{i',j}=1$ liked j, then i would be part of the optimal solution of the second stage. In other words, the event in which no one (except i) likes j implies the event of i being part of the optimal solution. Therefore,

$$\Delta_{ij}=\phi_{ij}\phi_{ji}\sum_{\mathbf{B}:(i,j)\notin\mathbf{B}}\mathbb{1}_{\left\{i\in S_j(\mathbf{B},\mathbf{x})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\geq\phi_{ij}\phi_{ji}\prod_{\ell\neq i:x_{\ell,j}=1}(1-\phi_{\ell,j})\geq\phi_{ij}\phi_{ji}\left(1-\frac{1}{n}\right)^{n-1}$$

where the last inequality is due to our assumption.

- If $w_{i,j}=1$, then the contribution of the pair i,j in the numerator is $\phi_{ij}\phi_{ji}$.

Now, let us compare the contribution of (i,j) to OPT_1 relative to its contribution to OPT_2. As before, we have different cases depending on the solutions $\tilde{\mathbf{x}}$, $\tilde{\mathbf{w}}$, and their second-period responses:

- If $\tilde{x}_{i,j'}=1$ for some $j'\in\mathcal{P}_i$ with $j'\neq j$ (when $j'=j$ is analogous). In this case, the platform shows j' to i in the first period (initiating a sequential interaction) instead of j, hoping to get an extra non-sequential match in the second period. Therefore, the contribution of $\tilde{x}_{i,j'}$ to OPT_2 would potentially involve two terms. First, between i and j', we have

$$\tilde{\Delta}_{ij'}=\phi_{ij'}\phi_{j'i}\sum_{\mathbf{B}:(i,j')\notin\mathbf{B}}\mathbb{1}_{\left\{i\in S_{j'}(\mathbf{B},\tilde{\mathbf{x}},\tilde{\mathbf{w}})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})$$

where $S_{j'}(\mathbf{B},\tilde{\mathbf{x}},\tilde{\mathbf{w}})$ is an optimal solution for user j' in the objective $\tilde{f}$ of the second period with backlog $\mathbf{B}$, $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{w}}$. In the worst case, there is also a non-sequential match between i and j in the second period, which contributes (in expectation)

$$\tilde{\Delta}_{ij}=\phi_{ij}\phi_{ji}\sum_{\mathbf{B}}\mathbb{1}_{\left\{i\in S_j(\mathbf{B},\tilde{\mathbf{x}},\tilde{\mathbf{w}})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B}).$$

Therefore, we have the following contribution in the denominator

$$\tilde{\Delta}_{ij'}+\tilde{\Delta}_{ij}=\phi_{ij'}\phi_{j'i}\sum_{\mathbf{B}:(i,j')\notin\mathbf{B}}\mathbb{1}_{\left\{i\in S_{j'}(\mathbf{B},\tilde{\mathbf{x}},\tilde{\mathbf{w}})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})+\phi_{ij}\phi_{ji}\sum_{\mathbf{B}}\mathbb{1}_{\left\{i\in S_j(\mathbf{B},\tilde{\mathbf{x}},\tilde{\mathbf{w}})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\leq 2\phi_{ij}\phi_{ji},$$

where the inequality is due to two facts: (i) $\tilde{\Delta}_{ij'}\leq\phi_{ij}\phi_{ji}$, otherwise in solution $\mathbf{x}$ with value OPT_1 we could show j' to i instead of j and obtain a better solution, which would contradict the optimality of $\mathbf{x}$; (ii) $\sum_{\mathbf{B}}\mathbb{1}_{\left\{i\in S_j(\mathbf{B},\tilde{\mathbf{x}},\tilde{\mathbf{w}})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\leq 1$.

- If $\tilde{w}_{i,j'}=1$ for some $j'\in\mathcal{P}_i$; j' is potentially different than j. In this case, the platform decided to show j' to i simultaneously in the first period. Then, the contribution in the worst case is the following

$$\tilde{\Delta}_{ij'}+\tilde{\Delta}_{ij}=\phi_{ij'}\phi_{j'i}+\phi_{ij}\phi_{ji}\sum_{\mathbf{B}}\mathbb{1}_{\left\{i\in S_j(\mathbf{B},\tilde{\mathbf{x}},\tilde{\mathbf{w}})\right\}}\cdot\mathbb{P}(\mathcal{B}=\mathbf{B})\leq 2\phi_{ij}\phi_{ji}.$$

where the inequality follows as before, i.e., we must have $\phi_{j'i}\leq\phi_{ji}$ (recall both $\phi_{ij'}$ and ϕ_{ij} are at most $1/n$), since otherwise we can change the solution in $\mathbf{x}$ for OPT_1 and get a better objective as j' would be part of the second-period optimal solution (j' would see i in response) whenever j is.

- If $\tilde{x}_{i,j} = \tilde{w}_{i,j} = 0$ for all $j \in \mathcal{P}_i$. This case is similar to the previous one, but now the contribution would be composed only by a second-period term $\tilde{\Delta}_{ij}$, which is no worse (in terms of denominator) than the other cases.

Therefore, for any pair (i,j), the ratio between each contribution is at least

$$\frac{\Delta_{ij}}{\tilde{\Delta}_{ij'} + \tilde{\Delta}_{ij}} \geq \frac{\phi_{ij}\phi_{ji}\left(1-\frac{1}{n}\right)^n}{2\phi_{ij}\phi_{ji}} \geq \frac{1}{2e} - o(1)$$

Finally,

$$\frac{\sum_{\mathbf{B}}\left(\sum_{e\in E} w_e \cdot \beta_e + f(\mathbf{B} \cap \vec{E}(\mathbf{x}))\right) \cdot \mathbb{P}(\mathcal{B} = \mathbf{B})}{\sum_{\mathbf{B}}\left(\sum_{e\in E} \tilde{w}_e \cdot \beta_e + \tilde{f}(\mathbf{B} \cap \vec{E}(\tilde{\mathbf{x}}), \tilde{\mathbf{x}}, \tilde{\mathbf{w}})\right) \cdot \mathbb{P}(\mathcal{B} = \mathbf{B})} \geq \min_{i,j,j'} \left\{ \frac{\Delta_{ij}}{\tilde{\Delta}_{ij'} + \tilde{\Delta}_{ij}} \right\} \geq \frac{1}{2e} - o(1).$$

Note that we may be comparing more terms than we actually need, but we were looking just for a lower bound. To conclude the second part of the theorem, we observe that we can always obtain more matches (in expectation) when we allow non-sequential matches in both periods rather than in the first period only. Formally, for any feasible solution $(\mathbf{x}^1, \mathbf{w}^1)$ in Problem (2), then

$$\mathcal{M}^2(\mathbf{x}^1, \mathbf{w}^1) \leq \tilde{\mathcal{M}}^2(\mathbf{x}^1, \mathbf{w}^1).$$

Therefore, if we consider $\mathbf{x}^1, \mathbf{w}^1$ a γ-approximate solution for Problem (2). Then, we have

$$\tilde{\mathcal{M}}^2(\mathbf{x}^1, \mathbf{w}^1) + \sum_{e\in E} \beta_e w_e^1 \geq \mathcal{M}^2(\mathbf{x}^1, \mathbf{w}^1) + \sum_{e\in E} \beta_e w_e^1 \geq \gamma \cdot \mathrm{OPT}_1 \geq \gamma \cdot \left(\frac{1}{2e} - o(1)\right) \cdot \mathrm{OPT}_2. \quad \square$$