HAL open science

Two-sided matching markets with correlated random preferences have few stable pairs

Hugo Gimbert, Claire Mathieu, Simon Mauras

▶ **To cite this version:**

Hugo Gimbert, Claire Mathieu, Simon Mauras. Two-sided matching markets with correlated random preferences have few stable pairs. 2020. hal-03035039

HAL Id: hal-03035039
https://hal.science/hal-03035039

Preprint submitted on 2 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:1904.03890v2 [cs.DS] 14 Feb 2020

Two-sided matching markets with correlated random preferences have few stable pairs

HUGO GIMBERT, Université de Bordeaux, LaBRI, CNRS, F-33400 Talence, France
CLAIRE MATHIEU, Université de Paris, IRIF, CNRS, F-75013 Paris, France
SIMON MAURAS, Université de Paris, IRIF, CNRS, F-75013 Paris, France

Stable matching in a community consisting of N men and N women is a classical combinatorial problem that has been the subject of intense theoretical and empirical study since its introduction in 1962 in a seminal paper by Gale and Shapley.

In this paper, we study the number of stable pairs, that is, the man/woman pairs that appear in some stable matching. We prove that if the preference lists on one side are generated at random using the popularity model of Immorlica and Mahdian, the expected number of stable edges is bounded by $N \ln N + N$, matching the asymptotic value for uniform preference lists. If in addition that popularity model is a geometric distribution, then the number of stable edges is $O(N)$ and the incentive to manipulate is limited. If in addition the preference lists on the other side are uniform, then the number of stable edges is asymptotically N up to lower order terms: most participants have a unique stable partner, hence non-manipulability.

1 INTRODUCTION

In the classical stable matching problem, a certain community consists of N men and N women, all heterosexual and monogamous, and each person ranks those of the opposite sex in accordance with his or her preferences for a marriage partner. Our objective is to marry off all members of the community in such a way that the established matching is *stable*, *i.e.* such that there is no *blocking pair*: a man and a woman who are not married to each other but prefer each other to their actual mates.

In their seminal paper, Gale and Shapley [10] prove that there always exists a stable matching and give an algorithm for the problem. Their original motivation was the assignment of students to colleges, a setting to which the algorithm and results extend, and their approach was successfully implemented in many matching markets; see for example [1, 2, 5, 29].

However, there exist instances with more than one stable matching, and even extreme cases of instances in which every man/woman pair belongs to some stable matching. This raises the question of which matching to choose [12] and of possible strategic behavior (see paragraph on manipulability below).

Fortunately, there is empirical evidence showing that in many instances, in practice the stable matching is essentially unique (phenomenon often referred as "core-convergence"); see for example [4, 14, 23, 29]. One of the empirical justifications of core-convergence given by Roth and Peranson in [29] is that the preference lists are correlated: *"One factor that strongly influences the size of the set of stable matchings is the correlation of preferences among programs and among applicants. When preferences are highly correlated (i.e., when similar programs tend to agree which are the most desirable applicants, and applicants tend to agree which are the most desirable programs), the set of stable matchings is small."*

Following that direction of enquiry, we argue that "core convergence" is best captured by the total number of man/woman pairs that belong to some stable matching, and we model correlations by sampling preference lists using the popularity model of Immorlica and Mahdian [15].

We prove that if the preference lists on one side are sampled using popularities, the expected number of stable edges is bounded by $N \ln N + N$, matching the asymptotic value for uniform preference lists [25]. If in addition that popularity model is a geometric distribution, then the number of stable edges is $O(N)$ and the incentive to manipulate is limited. Further, if in addition the preference lists on the other side are uniform, then the number of stable edges is asymptotically N up to lower order terms: most participants have a unique stable partner, hence non-manipulability.

This work was done in part while one of the authors was visiting the Simons Institute for the Theory of Computing; and was partially funded by the grant ANR-19-CE48-0016 from the French National Research Agency (ANR).

1.1 Definitions and main theorems

Let N be a positive integer, $\mathcal{M} = \{m_1, m_2, \ldots, m_N\}$ a set of N men and $\mathcal{W} = \{w_1, w_2, \ldots, w_N\}$ a set of N women. Each man has a total order over $\mathcal{W}$ and each woman has a total order over $\mathcal{M}$, representing their preferences for a marriage partner.

Definition 1 (Stability, from [10])**.** Given a perfect matching over $\mathcal{M} \cup \mathcal{W}$, a pair (m, w) not in the matching is *blocking* if man m prefers woman w to his partner and w prefers m to her partner. A perfect matching over $\mathcal{M} \cup \mathcal{W}$ is *stable* if there is no blocking pair.

Definition 2. A pair is *stable* if it belongs to at least one stable matching. The *stable graph* is the graph of all stable pairs.

Preference lists generated from popularities are an input model introduced by Immorlica and Mahdian in [15] for preference lists of fixed lengths. In the present paper, the stochastic process is used to generate full length preference lists.

Definition 3 (Popularity preferences)**.** We say that a woman w has *popularity preferences* if there exists a distribution[1] $\mathcal{D}$ over the men $\mathcal{M}$ such that w builds a random preference list over $\mathcal{M}$ by repeatedly sampling from $\mathcal{D}$ without replacement (renormalizing $\mathcal{D}$ at each step). Similarly, one may define popularity preferences for a man.

The popularities of m_1, m_2 and m_3 are respectively 2, 5 and 3; the probability that $m_2 \succ m_1 \succ m_3$ is:

$$\frac{5}{2+5+3} \cdot \frac{2}{2+3} \cdot \frac{3}{3} = 0.2$$

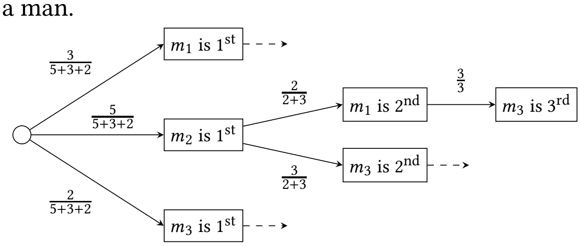


Fig. 1. Stochastic process used to generate a preference list from popularities.

Figure 1 illustrates the process of sampling a preference list over 3 men m_1, m_2 and m_3 of popularities 2, 5 and 3. Except when it is explicitly stated otherwise, we will assume that women have i.i.d. popularity preferences. When the distribution $\mathcal{D}$ is uniform, the preference lists of women are random independent uniform permutations. When the distribution is highly skewed, with high probability the women all have the same preference list, thus the distribution induces correlations between the preferences of different women.

Theorem 1. *Assume that the men have* arbitrary preferences *and the women have* i.i.d. popularity preferences. *Then the expected total number of stable pairs is at most $N + N \ln N$.*

Pittel [25] proved that in the case where men and women have uniform preferences, the expected total number of stable pairs is $\sim N \ln N$, hence Theorem 1 implies that the uniform case asymptotically maximizes the number of stable pairs, up to lower order terms. To the best of our knowledge, this is the first theoretical result proving that having uncorrelated preferences is a worst case situation for "core convergence".

Definition 4 (Geometric popularity preferences)**.** We say that a women w has *geometric popularity preferences* if $\mathcal{D}$ is a geometric distribution: there exists $0 < \lambda < 1$ such that $\mathcal{D}(m_i) = \lambda^i$ for all $1 \leq i \leq N$.

The preference lists of the women are highly correlated when λ is small and become all identical if λ tends to 0. The limit case is a well known instance where there is exactly one stable matching.

[1] actually we do not require that $\sum_{m \in \mathcal{M}} \mathcal{D}(m) = 1$, as we normalize each time a man is sampled.

Theorem 2. *Assume that the men have arbitrary preferences and the women have i.i.d. geometric popularity preferences. Then the expected total number of stable pairs is $O(N)$. Moreover, a woman using a non-truthful strategy can improve the rank of her partner in her true preference list by at most a constant, with high probability.*

Although the number of stable matchings may still be exponentially large, Theorem 2 implies that most participants only have a constant number of stable partners on average. Theorem 2 can be compared to a result from [21], where preferences are based on random utilities and Lee showed that the gain of someone trying to manipulate is almost null for almost everyone.

Theorem 3. *Assume that the men have uniform popularity preferences and the women have geometric popularity preferences. Then the expected total number of stable pairs is $N + O(\ln^3 N)$. Moreover, a woman using a non-truthful strategy will not improve the rank of her partner in her true preference list, with high probability.*

Theorem 3 can be compared to results from [3, 15] where it is shown that the fraction of persons with more than one stable partner is vanishingly small.

To summarize our results, we studied the number of stable pairs when adding correlations between the preference lists: starting from arbitrary preferences (that can be "negatively correlated"), we considered uniformly random preferences (not correlated), popularity preferences (that are "positively correlated"), geometric popularity preferences (preferences are all identical in the limit). To go further, one might want to relax the assumption that all the woman have the same popularity distribution over the men (because there might be several "types" of women). Extensions of Theorems 1, 2 and 3 are respectively discussed in subsections 3.3, 4.5 and 5.3.

1.2 Techniques

Our results follow from a probabilistic analysis of Algorithm 1 and Algorithm 2, using the *principle of deferred decisions*. Algorithm 1, introduced in [10], finds the stable matching in which women are matched to their worst stable partners. Algorithm 2, introduced in [19] as a variant of previous algorithms, finds all stable partners of a given woman.

The core of the proof of Theorem 1 is Lemma 12, that bounds the number of stable partners of a woman as a function of their popularities. In the proof of Theorem 2 we first argue that if two men have very different popularities, then the less popular one will always have lower priority than the more popular one, in every preference list. Then, men are effectively only in competition with men who have popularities similar to their own. We therefore order the men by popularity, define a corresponding ordering of women (see Definition 5), and line up men and women. We can define blocks so that a person in a block can only be married to someone in the same block (see Figure 2). In Theorem 3 we extend the analysis of Theorem 2, incorporating ideas from [3] to prove that the stable graph is nearly a perfect matching.

1.3 Previous results

Stable matchings. Gale and Shapley [10] proved that stable matchings always exist and gave an algorithm for finding one, the *men proposing deferred acceptance* procedure (see Algorithm 1), which they prove is *men-optimal*: every man is at least as well off (in the sense of the rank of his partner in his preference list) as he would be in any other stable matching. They also proved that the men-optimal stable matching is also the women-pessimal stable matching. By symmetry, there also exists a women-optimal/men-pessimal stable matching. Mc Vitie and Wilson, and then Gusfield [11, 22] showed that this women-optimal matching can also be obtained through a sequence of "rejection chains" via an extension of the men proposing deferred acceptance algorithm (see Algorithm 2, a simplified version from [19]). Those algorithms exploit the distributive lattice structure of the set of stable matchings, an observation attributed to J.H. Conway in [17]. For more information, see the books [6, 12, 18].

Manipulability. There exists instances with more than one stable matching, and even extreme cases of instances in which every man/woman pair belongs to some stable matching. This raises the question of which matching to choose [11, 12] and of possible strategic behavior, when for some participants, reporting a true preference list may not be a best response to reported preferences of others [8, 28]. Demange, Gale and Sotomayor [7] proved the following. Assume a woman lies about her preference list, falsifying her preferences strategically by changing her order of preferences and truncating it (declaring some men unacceptable). This gives rise to a new stable matching, but she will be no better off than she would be in the true woman-optimal matching. This implies that a woman can only gain from strategic manipulation up to the maximum difference between their best and worst partners in stable matchings. In particular, if a woman has the same partner in all stable matchings (unique stable partner), then she cannot benefit from manipulation. By symmetry, this also implies that the men proposing deferred acceptance procedure is strategy-proof for men (as they will get their best possible partner by telling the truth).

Stochastic preferences. In the worst case, there are instances where every man/woman pair belongs to some stable matching. Fortunately, there is empirical evidence showing that in many instances, in practice the stable matching is essentially unique. Theoretical studies have tried to understand the "core convergence" in one-to-one matching [3, 15, 19] and in the extension to many-to-one matching [13, 20]. Analyzing instances that are less far-fetched than in the worst case is the motivation underlying the model of stochastically generated preference lists. Researchers have studied uniformly random preferences [3, 19, 24–27], preference generated with a popularity-based model [15, 20], and preferences induced by random utilities [21]. To express the idea of "core convergence" in real life markets, a number of parameters are of interest in the study of stable marriage.

Counting the number of stable matchings. Since the question was first posed by Knuth [17], the maximum (over all instances) number of stable matchings has been the object of much study, including recent work by Karlin, Gharan and Weber [16], proving an exponential upper-bound. The best known lower bound is approximately 2.28^N in [30]. However, note that a single pair may belong to many stable matchings. Therefore one may ask about the number of stable pairs.

Counting the number of stable pairs. A *stable pair* is a man/woman pair which belongs to at least one stable matching. When the stable matching is unique, there are exactly N stable pairs. But in general the number of stable pairs is a very different measure from the number of stable matchings. In the case of uniform preferences, there are $\Theta(N \log N)$ stable pairs on average [19], yet the number of stable matchings is also very low, just $\Theta(N \log N)$ [24]. In the case where all preferences are almost aligned (for example when men m_{2i} and m_{2i+1} always precedes m_{2i+2} and m_{2i+3}), there are only $O(N)$ stable pairs, yet the number of stable matchings might be exponential in N: the number of stable matchings is not a good measure of the complexity of the instance in this case.

Counting the number of persons with k stable partners. Instead of the total number of stable pairs, one may refine the analysis and study the distribution of the number of stable partners of a given participant. This was done in full detail by Pittel, Shepp and Veklerov in the case of uniform preferences [27]. The number of persons with exactly one stable partner has also been studied in [15] for constant size preference lists, and in [3, 26] for unbalanced markets.

Other measures. Given a stable matching, one may also study the average over the men of the rank of the man's partner in his preference list. One measure of the complexity of the instance is the ratio/difference between that value for the men-optimal and for the men-pessimal matchings. In the case of uniform preferences, that ratio is about $\Theta(N/\log^2 N)$ [24]. In the case of N men and $N + 1$ women, this ratio falls to $1 + o(1)$ for both men and women [3, 26]. Note that this measure is also related to the number of stables pairs: the difference in rank between someone's best and worst partner is an upper bound on their number of stable partners. Finally, a last measure uses a model where each person receives some "utility" from a given matching. Each person wants to maximize the utility they receives, thus it induces preference lists. This model was studied with randomly sampled utilities in [21] where Lee showed that almost everyone receives almost the same utility from every stable matching. When those measures are small, that has implications on manipulability, not because the number of stable partners is small, but because the incentive for someone to lie is relatively small.

2 BACKGROUND: STABLE MATCHINGS WITH RANDOM PREFERENCES

We start this section by giving formal notations for the stable matching problem, then we recall classical structural results and algorithms on the lattice of stable matchings. Finally we discuss the stochastic process used to generate random preferences.

2.1 Stable matchings

In terms of notations, we view a matching as a function $\mu : \mathcal{M} \cup \mathcal{W} \to \mathcal{M} \cup \mathcal{W}$, which is an involution ($\mu^2 = \text{Id}$), where each man is paired with a woman ($\mu(\mathcal{M}) \subseteq \mathcal{W}$), and where each woman is paired with a man ($\mu(\mathcal{W}) \subseteq \mathcal{M}$). We denote by $\succ_m$ the total order of man m over $\mathcal{W}$ ($w \succ_m w'$ means that m prefers w to w'), similarly $\succ_w$ for the total order of woman w over $\mathcal{M}$.

Algorithm 1 Men Proposing Deferred Acceptance.

Input: Preferences of men $(\succ_m)_{m \in M}$ and women $(\succ_w)_{w \in \mathcal{W}}$.
Initialization : Start with an empty matching. ▹ define $\mu \leftarrow \text{id}_{\mathcal{M} \cup \mathcal{W}}$
While there is a man m who is single**, do** ▹ while $\exists m \in \mathcal{M},\ \mu(m) = m$
  m proposes to his favorite woman w he has not proposed to yet. ▹ in the total order $\succ_m$
  If m is w's favorite man among all proposals she received, ▹ if $\mu(w) = w$ or $m \succ_w \mu(w)$
    w accepts m's proposal, and rejects her previous husband if she was married.
    ▹ execute $\mu(\mu(w)) \leftarrow \mu(w)\ ;\ \mu(m) \leftarrow w\ ;\ \mu(w) \leftarrow m$
Output: Resulting matching. ▹ return $\mu_\perp \leftarrow \mu$

Algorithm 2 Extended Men Proposing Deferred Acceptance.

Input: Preferences of men $(\succ_m)_{m \in M}$ and women $(\succ_w)_{w \in \mathcal{W}}$. Fixed woman $w^* \in \mathcal{W}$.
Initialization : Start by executing Algorithm 1, then break (virtually) the pair between w^* and her husband m, who becomes the *proposer*.
▹ compute matching $\mu \leftarrow \mu_\perp$, then define $m \leftarrow \mu(w^*)$ and $S \leftarrow \{m\}$
While the *proposer* m has not proposed to every woman**, do**
  At this point, man m and woman w^* are (virtually) single, and everyone else is paired.
  m proposes to his favorite woman w he has not proposed to yet. ▹ in the total order $\succ_m$
  If m is w's favorite man among all proposals she received, ▹ if $m \succ_w \mu(w)$
    If $w = w^*$**,** then w^* accepts m's proposal, m is a new stable husband of w^*, then break (virtually) the pair between w^* and m, who stays the *proposer*.
    ▹ if $w = w^*$, then execute $\mu(m) \leftarrow w\ ;\ \mu(w) \leftarrow m, S \leftarrow \{m\} \cup S$
    Else w rejects her previous husband m', accepts m's proposal, the *proposer* becomes m'.
    ▹ otherwise execute $m' \leftarrow \mu(w), \mu(m) \leftarrow w\ ;\ \mu(w) \leftarrow m, m \leftarrow m'$
Output: Set of stable husbands of w^*. ▹ return S

See Appendix A for a detailed execution of Algorithms 1 and 2.

Lemma 4 (Adapted from [10])**.** *Algorithm 1 outputs a stable matching $\mu_\perp$ in which every man (resp. woman) has his best (resp. her worst) stable partner.*

Lemma 5 (Adapted from [19])**.** *Given a woman w^*, Algorithm 2 enumerates all the* stable husbands *of w^* (that is, her partners in all stable matchings).*

The following is a corollary of Lemma 4 obtained by reversing the roles of men and women.

Lemma 6. *If for each woman w we denote by $\mu_\top(w)$ her favorite stable husband, then $\mu_\top$ is a stable matching in which every man has his worst stable partner.*

2.2 Popularity preferences

The *popularity based* stochastic process we use to generate ordered preference lists has previously been defined and studied in [15, 20]. In this model each person has a positive popularity, and in a preference list, a person of popularity p_1 has a probability $p_1/(p_1+p_2)$ to be ranked before a person of popularity p_2.

More precisely, each man $m \in \mathcal{M}$ has a popularity $\mathcal{D}(m) \in \mathbb{R}_+^*$. A woman $w \in \mathcal{W}$ builds her preference list, a total order $>_w$, by sampling without replacement from $\mathcal{M}$ using the (normalized) distribution $\mathcal{D}$ (she samples her first choice, then her second, …). We note $\mathcal{D}^N$ the induced distribution over preference lists. This sampling method is also known as *weighted random sampling* [9].

Lemma 7. *Among a set of men $\mathcal{S} \subseteq \mathcal{M}$, the probability that a man m is w's favorite is $\frac{\mathcal{D}(m)}{\sum_{s\in\mathcal{S}}\mathcal{D}(s)}$.*

$$\forall S \subseteq \mathcal{M}, \quad \forall m \in S \setminus \{m\}, \quad \mathbb{P}[\forall s \in \mathcal{S}, m >_w s] = \frac{\mathcal{D}(m)}{\sum_{s\in\mathcal{M}}\mathcal{D}(s)}$$

Proof. During the sampling of $>_w$ consider the first time when a man in $\mathcal{S}$ is chosen. The probability that that man is m, given that it is a man of S that is chosen, equals the above quantity. □

Lemma 8. *The preferences of woman w can be computed "online": during the execution of Algorithms 1 and 2, the only questions about w's preferences are of the form: "is m the best man among all proposals received by w so far?", and w's answer to that question is independent of her answers to previous questions.*

Proof sketch. Consider a situation where we know men m_1 and m_2 proposed to a woman w and that woman w prefers m_2 to m_1. The probability that a proposal from man m_3 will be accepted (that is, m_3 is better that m_2 and m_1) given that w prefers m_2 to m_1 is $\mathbb{P}[m_3 >_w m_2, m_1 \mid m_2 >_w m_1]$ which can be computed as:

$$\frac{\mathbb{P}[m_3 >_w m_2 >_w m_1]}{\mathbb{P}[m_2 >_w m_1]} = \frac{\mathbb{P}[m_3 >_w m_2, m_1] \cdot \mathbb{P}[m_2 >_w m_1]}{\mathbb{P}[m_2 >_w m_1]} = \mathbb{P}[m_3 >_w m_2, m_1].$$

Thus $(m_3 >_w m_2, m_1)$ and $(m_2 >_w m_1)$ are independent. The same proof works in general. □

Remark 9. *Most of our proofs in sections 3, 4 and 5 will use the* principle of deferred decisions *(see [19]) and are based on a stochastic analysis of Algorithm 2. We condition on the "initial" matching $\mu_\perp$, obtained at the end of Algorithm 1. Indeed, one can notice that knowing $\mu_\perp$ and the preference lists of the men, one can deduce, for each woman, the proposals received so far (each man proposed to everyone on his preference list up to his current wife) and her favorite man among among those (namely, her current husband). Using Lemmas 7 and 8, this (partial) knowledge about the execution of Algorithm 1 suffices to continue the execution of Algorithm 2, with women answering questions online.*

3 DEFERRED ACCEPTANCE: PROOF OF THEOREM 1

In this section we prove an upper bound on the number of stable pairs through the analysis of the men proposing deferred acceptance algorithm. In previous work [15], Immorlica and Mahdian studied a setting where women have complete arbitrary preferences and men have truncated popularity preferences. In our analysis, women have complete popularity preferences and men have (truncated) arbitrary preferences. The main difference is that in [15], the proposals are sampled online, with deterministic answers. Here the randomness is moved to the answering part, with deterministic proposals.

3.1 One woman has popularity preferences

In this subsection, we are given a woman w^*. The preference lists of all other participants are given and arbitrary. Woman w^* has popularity preferences, and her preferences are sampled "online" during the execution of Algorithms 1 and 2, according to Lemmas 7 and 8.

Let us detail a possible real life execution of Algorithm 1: some day, w^* receives a proposal from man m_1 and accepts because she is single. Later she receives a proposal from man m_2. With probability $\mathcal{D}(m_1)/(\mathcal{D}(m_1) + \mathcal{D}(m_2))$, she prefers m_1 to m_2, and then she rejects the proposal from m_2 and stays with m_1. Later still, she receives a proposal from man m_3. With probability $\mathcal{D}(m_3)/(\mathcal{D}(m_1)+\mathcal{D}(m_2)+\mathcal{D}(m_3))$, she prefers m_3 to both m_1 and m_2, and then she accepts m_3 and rejects her previous husband m_1. Note that knowing which is her preferred husband between m_1 and m_2 is irrelevant to computing the probability of her answer to m_3.

Lemma 10. *Let $\mu_\perp$ denote the matching obtained at the end of Algorithm 1 and $p_\perp$ denote the sum of popularities of proposals received by w^* during Algorithm 1, including $\mu_\perp(w^*)$. Conditioning on $\mu_\perp$ and on the preferences of other participants, the sequence $x_1, x_2, \dots, x_K$ of proposals received by w^* during the rest of the execution of Algorithm 2 is completely determined and independent of w^*'s preferences, and her answers to the proposals are independent from one another. Let $p_i = \mathcal{D}(x_i)$ for all $1 \le i \le K$. Then:*

$$\forall i \in \{1, \dots, K\}, \quad \mathbb{P}[\textit{proposal } x_i \textit{ is accepted by } w^* \mid \mu_\perp] = \frac{p_i}{p_\perp + p_1 + \cdots + p_i}$$

Proof. As the preferences of everyone except w^* are deterministic, recall from Remark 9 that conditioning on the value of $\mu_\perp$, we know everything needed for the rest of the execution of Algorithm 2. The only source of randomness during the execution of Algorithm 2 comes from the answers of w^* to the proposals she receives, and the only impact of her accepting of refusing a proposal is to add or not the proposer to the set of stable husbands of w^*. □

Lemma 11. *Let $\mu_\perp$ denote the matching obtained at the end of Algorithm 1 and $\mu_\top(w^*)$ denote the last man accepted by w^* during the execution of Algorithm 2 (her favorite stable husband). Let K be the number of proposals received by w^* in Algorithm 2. Then:*

$$\mathbb{E}[\textit{Nb of stable husbands of } w^* \mid \mu_\perp] \le 1 + \mathbb{E}[\ln(K+1) \mid \mu_\perp] + \mathbb{E}\left[\ln(\mathcal{D}(\mu_\top(w^*))) \mid \mu_\perp\right] - \ln(\mathcal{D}(\mu_\perp(w^*)))$$

Proof. The main ingredient is a sum/integral comparison.

$$\sum_{i=1}^{K} \frac{p_i}{p_\perp + p_1 + \cdots + p_i} \le \sum_{i=1}^{K} \int_{p_\perp + p_1 + \cdots + p_{i-1}}^{p_\perp + p_1 + \cdots + p_i} \frac{\mathrm{d}t}{t} = \ln(p_\perp + p_1 + \cdots + p_K) - \ln p_\perp$$

See Appendix B for the full proof. □

Taking the inequality of Lemma 11 and averaging over $\mu_\perp$ immediately yields Lemma 12.

Lemma 12. *Let $\mu_\perp$ be the men optimal matching and $\mu_\top$ the women optimal matching ($\mu_\perp$ and $\mu_\top$ are random variables). Let K_{w^*} be the number of men who rank w^*. Then:*

$$\mathbb{E}[\textit{Number of stable husbands of } w^*] \leq 1 + \ln K_{w^*} + \mathbb{E}\left[\ln \frac{\mathcal{D}(\mu_\top(w^*))}{\mathcal{D}(\mu_\perp(w^*))}\right]$$

3.2 All women have *i.i.d.* popularity preferences

In this subsection, all women have popularity preferences, and sample independently their lists using the same distribution $\mathcal{D} : \mathcal{M} \to \mathbb{R}_+^*$.

Theorem 1. *Assume that the men have* arbitrary preferences *and the women have* i.i.d. popularity preferences. *Then the expected total number of stable pairs is at most $N + N \ln N$.*

Proof. All preferences lists $(>_w)_{w\in\mathcal{W}}$ being independent given the popularities, Lemma 12 is valid for each woman $w \in \mathcal{W}$. Indeed, the case where all the other women have popularity preferences is actually a linear combination of cases where those women have deterministic preferences. Thus we write $Y = \sum_{w\in\mathcal{W}} Y_w$, with Y the total number of stable pairs and Y_w the number of stable husbands of w. We can use Lemma 12 and linearity of expectation to obtain:

$$\mathbb{E}[Y] = \sum_{w\in\mathcal{W}} \mathbb{E}[Y_w] \leq N + N\ln N + \mathbb{E}\left[\sum_{w\in\mathcal{W}} \ln(\mathcal{D}(\mu_\top(w))) - \sum_{w\in\mathcal{W}} \ln(\mathcal{D}(\mu_\perp(w)))\right]$$

Since $\mu_\perp$ and $\mu_\top$ are both perfect matchings, in the expectation the two sums cancel each other, because they both are equal to $\sum_{m\in\mathcal{M}} \ln(\mathcal{D}(m))$. □

Remark 13. *When men have arbitrary preference lists of length K, Theorem 1 gives an upper bound of $N + N\ln K$ on the expected number of stable pairs: let K_w be the number of men who rank woman w, by concavity of the logarithm, and using the fact that $\sum_{w\in\mathcal{W}} K_w = K \cdot N$, we have $\sum_{w\in W} \ln K_w \leq N \ln K$.*

3.3 Women have independent but non-identical popularity preferences

When women have i.i.d. popularity preferences, each man has ex-ante the same value in the eyes of every woman. One might want to relax this property, by allowing the distribution $\mathcal{D}$ to vary from one woman to another.

Theorem 14. *Assume that the men have arbitrary preferences lists and the women have independent popularity preferences: each woman has her own distribution $\mathcal{D}_w$. For each man m, let us define $C_m = \frac{\max_{w\in W} \mathcal{D}_w(m)}{\min_{w\in W} \mathcal{D}_w(m)}$. Then the expected total number of stable pairs is at most $N + N\ln N + \sum_{m\in M} \ln C_m$.*

Proof. In the proof of Theorem 1, we bound the last two terms: for any $\mu_\perp$ and $\mu_\top$ we have

$$\sum_{w\in\mathcal{W}} \Big(\ln(\mathcal{D}(\mu_\top(w))) - \ln(\mathcal{D}(\mu_\perp(w)))\Big) \leq \sum_{m\in\mathcal{M}} \Big(\max_{w\in\mathcal{W}} \mathcal{D}_w(m) - \min_{w\in\mathcal{W}} \mathcal{D}_w(m)\Big) \leq \sum_{m\in\mathcal{M}} \ln C_m$$

□

As long as the ratios C_m are polynomial in N, the expected number of stable pairs is $O(N \ln N)$.

4 INSTANCE DECOMPOSITION: PROOF OF THEOREM 2

In this section, we view the set of stable matchings as the cartesian product of the stable matchings of connected components of the stable graph. An instance of stable matching can thus be "decomposed" into several smaller sub-instances.

In this section the men have arbitrary preferences and the women have i.i.d. geometric popularity preferences: there exists λ such that for all $1 \leq i \leq N$ we have $\mathcal{D}(m_i) = \lambda^i$ with $0 < \lambda < 1$.

We now recall Theorem 2. We prove in Theorem 22 that the number of stable pairs is linear; and in Theorem 23 that the difference of rank between a woman best and worst stable husband is constant.

Theorem 2. *Assume that the men have arbitrary preferences and the women have i.i.d. geometric popularity preferences. Then the expected total number of stable pairs is $O(N)$. Moreover, a woman using a non-truthful strategy can improve the rank of her partner in her true preference list by at most a constant, with high probability.*

4.1 Decomposition into blocks

Recall that $\mu_\perp$ denotes the matching computed by Algorithm 1. Up to *relabeling* the women, we may assume that for all $1 \le i \le N$ we have $w_i = \mu_\perp(m_i)$. We now define *separators*: one separator at $t = 0$, and one separator for every $t \in [1, N]$ such that every woman $w_1, \ldots, w_t$ prefers her husband in $\mu_\perp$ to every man $m_{t+1}, \ldots, m_N$. This defines a sequence of separators $0 = x_0, x_1, x_2, \ldots, x_b = N$. If $x_i = t'$ and $x_{i+1} = t$ denote two consecutive separators, then we define a *block* as the set of men and women $\{m_{t'+1}, m_{t'+2}, \ldots, m_t, w_{t'+1}, w_{t'+2}, \ldots, w_t\}$.

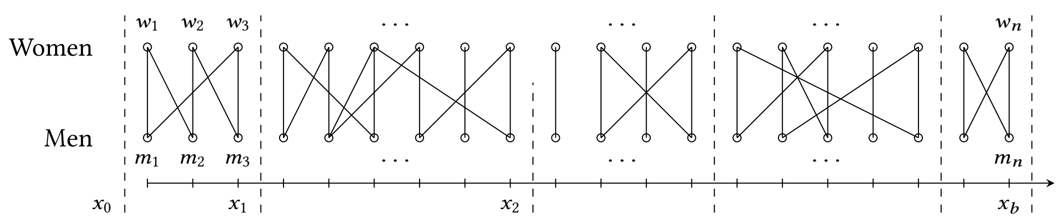


Fig. 2. Example of the decomposition of the stable graph into blocks using separators. The men are ordered by decreasing popularities, and $\mu_\perp$ is the vertical matching.

Lemma 15. *If man m_j and woman w_i are stable partners then w_i and m_j belong to the same block.*

Proof. Let t be a separator. First consider a woman w_i such that $i \le t$. Consider executing Algorithm 2 to find the stable partners of w_i. By definition of separators, w_i prefers m_i (and therefore also her subsequent stable husbands) to all men $m_{t+1}, \ldots, m_N$, so she will never accept any proposal from any man among those, so (w_i, m_j) is not stable for any $j > t$. Let μ be a stable matching. We have $\mu(m_i) \in \{m_1, m_2, \ldots m_t\}$ for all $i \le t$. Since the two sets have equal cardinality, we must therefore have $\mu(\{w_1, w_2, \ldots, w_t\}) = \{m_1, m_2, \ldots m_t\}$.

Let $t' < t$ be the previous separator. We must also have $\mu(\{w_1, w_2, \ldots, w_{t'}\}) = \{m_1, m_2, \ldots m_{t'}\}$. Therefore, $\mu(\{w_{t'+1}, w_{t'+2}, \ldots, w_t\}) = \{m_{t'+1}, m_{t'+2}, \ldots m_t\}$: in other words, μ matches inside the block. Since that is true for every stable matching μ, all stable edges are internal to blocks. □

Thanks to Lemma 15, we can state that the total number of stable pairs is at most the sum of the squares of the block sizes, $\sum_{i=1}^{b}(x_i - x_{i-1})^2$, so it only remains to analyze the distribution of the block sizes.

Remark 16. *Stable matchings with $N-1$ man and N woman having uniform preference lists have been studied by Ashlagi, Kanoria and Leshno [3], then by Pittel [26]. One can notice that unbalanced markets are easy to simulate with popularity preferences. Imagine a man m having a popularity ε so small that with high probability every woman will rank m last (for example a gap of $1/N^4$ is enough). The woman married to m in $\mu_\perp$ has only one stable partner, and she can in fact be considered as single. This is caused by a separator being placed between the first $N-1$ men and the last (least popular) one.*

4.2 Stochastic analysis of the size of the first block

In this subsection, we condition probabilities on the value of the matching $\mu_\perp$ obtained at the end of Algorithm 1. In this subsection we focus on the distribution of the size of the first block, x_1; later we will argue that the analysis is also valid for the size $(x_i - x_{i-1})$ of any block $1 \le i \le b$. To study x_1, we view the construction of the first separator as a Markov process, and use stochastic domination to derive upper bounds.

Lemma 17. *Let $t \in [1, N]$. With probability at least* $\exp\left(-\frac{\lambda^{\Delta+1}}{(1-\lambda)^2}\right)$, *for every $j \le t$ and every $i > t+\Delta$, woman w_j prefers m_j to m_i.*

Proof. Given $j \in [1, t]$, by definition of $\mu_\perp$, woman w_j prefers m_j to all the men who proposed to her during the execution of Algorithm 1. Using Lemmas 7 and 8, the probability that w_j prefers m_j to all the other contenders in $\{m_i\}_{t+\Delta<i\le n}$ is at least

$$\frac{\lambda^j}{\lambda^j + \sum_{t+\Delta<i\le n} \lambda^i} \ge \frac{1}{1 + \frac{\lambda^{\Delta+1+t-j}}{1-\lambda}} \ge \exp\left(-\frac{\lambda^{\Delta+1+t-j}}{1-\lambda}\right).$$

By independence, the probability that this holds for every $j \in [1, t]$ is at least

$$\prod_{1\le j\le t} \exp\left(-\frac{\lambda^{\Delta+1+t-j}}{1-\lambda}\right) = \exp\left(-\frac{\lambda^{\Delta+1}}{1-\lambda}\sum_{1\le j\le t} \lambda^{t-j}\right) \ge \exp\left(-\frac{\lambda^{\Delta+1}}{(1-\lambda)^2}\right).$$

□

In particular, Lemma 17 for $\Delta = 0$ means that t is a separator with probability at least $\exp(-\frac{\lambda}{(1-\lambda)^2})$. However, because of correlations, we cannot infer anything useful about the expectation of x_1 (and even less about the expectation of x_1^2), so a more refined analysis is necessary.

Given $\mu_\perp$, we view the separators as being constructed in an online manner; we proceed by order of increasing i and reveal the entire preference list of w_i by generating her preferences with respect to all the men who have not proposed to w_i during the execution of Algorithm 1.

Definition 5. Let $t \in [1, N]$, let $\Delta \ge 0$ be an integer. We denote by $(P_{t,\Delta})$ the property that for every $j \le t$ and $i > t + \Delta$, w_j prefers m_j to m_i. To simplify notations, we also define $(P_t) := (P_{t,0})$. Notice that (P_t) holds if and only if t is a separator.

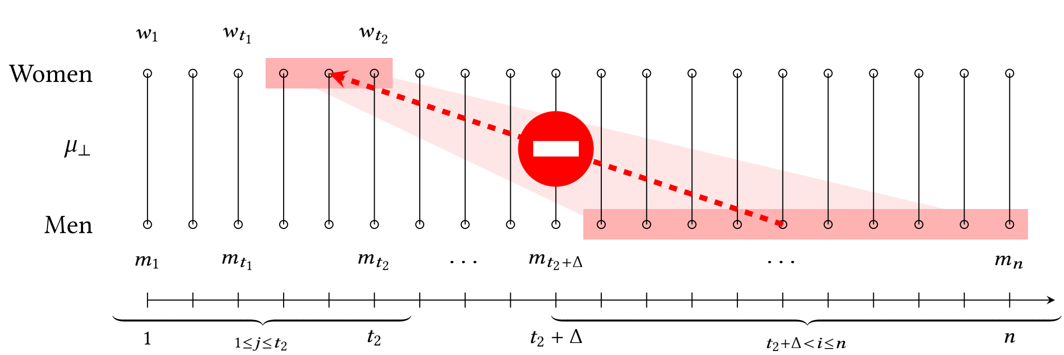


Fig. 3. Graphical representation of the analysis of event $(P_{t_2,\Delta})$ when we already know that every woman w_j for $j \le t_1$ prefers m_j to every man m_i with $i > t_2$. One then looks at the preference lists of women $w_{t_1+1}, w_{t_1+2}, \ldots, w_{t_2}$ and finds Δ such that each of those women prefers her husband to all men m_i with $i > t_2 + \Delta$.

During the first iteration, consider the preference list of w_1. The minimum value x_1 might conceivably have is $t_0 := 1$. If m_1 is w_1's favorite man, event P_1 holds, we have a separator, $x_1 = 1 = t_0$ and we are done. Else, let $t_1 = t_0 + \Delta_0$ denote the maximum index of the men that w_1 prefers to m_1. Separator x_1 can be no smaller than t_1. Note that by Lemma 17,

$$\forall k \geq 0, \quad \mathbb{P}[\Delta_0 \leq k \mid \mu_\perp] \geq \exp\left(-\frac{\lambda^{k+1}}{(1-\lambda)^2}\right).$$

We now reveal the preference lists of $w_2, w_3, \ldots, w_{t_1}$. If all of them prefer their respective husbands to $m_{t_1+1}, \ldots, m_N$, event P_{t_1} holds (since by definition of t_1, we already know that w_1 prefers m_1 to $m_{t_1+1}, \ldots, m_N$), we have a separator, $x_1 = t_1$ and we are done. Else, let $t_2 = t_1 + \Delta_1$ denote the maximum index of the men that some woman among $\{w_2, \ldots, w_{t_1}\}$ prefers to her husband. Separator x_1 can be no smaller than t_2. Note that by Lemma 17 again,

$$\forall k \geq 0, \quad \mathbb{P}[\Delta_1 \leq k \mid \mu_\perp] \geq \exp\left(-\frac{\lambda^{k+1}}{(1-\lambda)^2}\right).$$

Continuing in this manner, we obtain the following characterization of x_1:

$$x_1 = 1 + \Delta_0 + \Delta_1 + \cdots + \Delta_{s-1} \text{ with } \Delta_0, \Delta_1, \cdots, \Delta_{s-1} > 0, \Delta_s = 0,$$

where t_i and Δ_i are defined inductively: $t_0 = 1$; given t_i, Δ_i is defined as the minimum such that property (P_{t_i,Δ_i}) holds; and given t_i and Δ_i, we have $t_{i+1} = t_i + \Delta_i$.

Lemma 18. *Consider independent random variables $(\tilde{\Delta}_i)_{i\geq 0}$ with the identical distribution*

$$\forall k \geq 0, \quad \mathbb{P}[\tilde{\Delta}_i \leq k] = \exp\left(-\frac{\lambda^{k+1}}{(1-\lambda)^2}\right)$$

Let

$$\tilde{x}_1 = 1 + \tilde{\Delta}_0 + \tilde{\Delta}_1 + \cdots + \tilde{\Delta}_{\tilde{s}-1} \text{ with } \tilde{\Delta}_0, \tilde{\Delta}_1, \cdots, \tilde{\Delta}_{\tilde{s}-1} > 0, \tilde{\Delta}_{\tilde{s}} = 0$$

Then we have stochastic domination[2]: $\tilde{x}_1 \geq x_1$.

Proof sketch. Observe that Δ_i only depends on the preference lists of women $\{w_{1+t_{i-1}}, \ldots, w_{t_i}\}$, so the Δ_i's are independent. By Lemma 17, $\tilde{\Delta}_i \geq \Delta_i$. Proceeding inductively, summing independent random variables up to independent times s and $\tilde{s}$ yields that for all k, $\mathbb{P}[\tilde{x}_1 \geq k] \geq \mathbb{P}[x_1 \geq k \mid \mu_\perp]$, hence stochastic domination. □

Lemma 19. *The radius of convergence of the probability generating function of $\tilde{x}_1$ is > 1.*

Proof. See appendix C. □

Corollary 20. *The second moment of $\tilde{x}_1$ is bounded, independently of N. That is, $E[\tilde{x}_1^2] = O(1)$.*

Proof. Using Lemma 19, the probability generating function of $\tilde{x}_1$ is a power series $G_{\tilde{x}_1}$, and converges on an open set containing 1, thus $G_{\tilde{x}_1}$ can be derived twice. Recall that $\mathbb{E}[\tilde{x}_1^2] = G''_{\tilde{x}_1}(1) + G'_{\tilde{x}_1}(1)$. A proof giving explicit bounds on $\mathbb{E}[\tilde{x}_1^2]$ is also given in Appendix C. □

Corollary 21. *The random variable x_1 has an exponential tail, independently of N.*

$$\forall k \geq 0, \quad \mathbb{P}[x_1 \geq k \mid \mu_\perp] = \exp(-\Omega(k))$$

[2] Given to random variables A and B, we note $A \geq B$ and say that A stochastically dominates B when for all k we have $\mathbb{P}[A \geq k] \geq \mathbb{P}[B \geq k]$.

Proof. We use Lemma 18 for the domination of x_1, Lemma 19 for the existence of a constant $1+\varepsilon$ for which the probability generating function of $\tilde{x}_1$ is defined, and Markov's inequality.

$$\exists \varepsilon > 0, \forall k \geq 0, \quad \mathbb{P}[x_1 \geq k \mid \mu_\perp] \leq \mathbb{P}[\tilde{x}_1 \geq k] = \mathbb{P}[(1+\varepsilon)^{\tilde{x}_1} \geq (1+\varepsilon)^k] \leq (1+\varepsilon)^{-k} \cdot \mathbb{E}[(1+\varepsilon)^{\tilde{x}_1}]$$

□

4.3 The expected number of stable pairs is linear

Theorem 22. *Assume that the men have arbitrary preferences and the women have i.i.d. geometric popularity preferences. Then the expected total number of stable pairs is $O(N)$.*

Proof. As mentioned earlier, the number Y of stable pairs is at most the sum of the square of the size of each block:

$$Y \leq \sum_{i=1}^{b} (x_i - x_{i-1})^2.$$

Note that the total number b of blocks is at most N.

Observe that separator x_i is defined before the full preference lists of women w_j for $j > x_i$ are generated. Since the men m_j for $j > x_i$ are rejected by all women $\{w_1, \dots, w_{x_i}\}$, they might as well drop those women from their preference lists. Thus, given the first separator, the remaining problem is the same as the original problem, except that the number of participants is now $2(N-x_1)$ instead of $2N$. Thus the stochastic domination proved for the first block also holds for the second block, third block, etc.: with the same techniques we define N independent variables $y_1, \dots, y_N$, each distributed identically to $\tilde{x}_1$. Since $(x_i - x_{i-1}) \leq y_i$ for each i, we deduce that

$$Y \leq y_1^2 + \cdots + y_N^2$$

Using Corollary 20, the expected value of each y_i^2 is bounded by a constant independent of N.

$$\mathbb{E}[Y \mid \mu_\perp] \leq \mathbb{E}\left[\sum_{i=1}^{N} y_i^2\right] = \sum_{i=1}^{N} \mathbb{E}[y_i^2] = N\mathbb{E}[\tilde{x}_1^2] = O(N)$$

Now recall that all probabilities and expectation written above are conditioned on the value of the man optimal stable matching $\mu_\perp$. To obtain a bound on the unconditional expectation, write the law of total probability, the constant hidden in the O being independent from the choice of $\mu_\perp$. □

4.4 The difference of rank between two stable husbands is constant

Recall that a woman w^* who has more than one stable husband could falsely report her preference list to ensure that she will be matched with her best stable partner. One can imagine a situation where w^* has two stable husbands: the man she ranks first and the man she ranks last. In that case, w^* has very high incentives not to report her true preferences.

In the previous subsection, we proved that stable pairs can only exist within a block, and that the expected size of a block is constant. This does not imply that w^* cannot manipulate. In this subsection, we prove that in w^* preference list, the difference of rank between two men from her block is a constant. Thus, the incentives for w^* to lie are very limited.

Theorem 23. *Assume that the men have arbitrary preferences and the women have i.i.d. geometric popularity preferences. Then, for every women w^*, the difference of rank between her best and worst husbands is constant in expectation, and has an exponential tail.*

Proof. In subsection 4.2, the separators are constructed in an online manner: we proceed by order of increasing i and reveal the entire preference list of w_i by generating her preferences with respect to all the men who have not proposed to w_i during the execution of Algorithm 1.

Let w^* be a woman and let i such that $w^* = w_i$. We denote S the state of the algorithm after having run Algorithm 1, and after having sampled the preference lists of women $w_1, \ldots, w_{i-1}$.

- Let t' and t be the separators which define w^*'s block. Recall that $t' < i \leq t$. Conditioning on S, t' is a constant and t is a random variable. Let $x = t - t'$ be the size of w^*'s block.
- Let B be the set of men that have not proposed to w^* yet.
- Let $B_{high} = B \cap \{m_1, \ldots, m_{t'}\}$ and $B_{low} = B \cap \{m_{t'+1}, \ldots m_N\}$.
- Let z be the number of men from B_{high} that w^* ranks after a man from B_{low}.
- Let δ be the difference in rank between w^*'s best and worst partner.

Corollary 21 allows us to bound $x = t - t'$. We conclude the proof with Lemmas 24 and 25. □

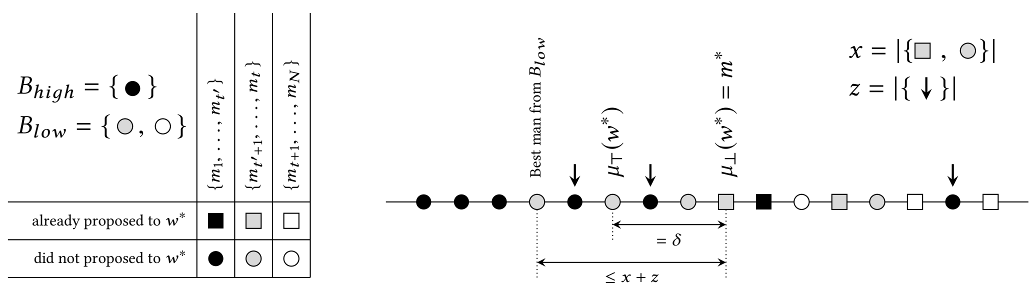


Fig. 4. Preference list of w^*. From left to right, men are represented from best to worst. Here $x = 6$ is the number of gray shapes; and $z = 3$ men from B_{high} are ranked behind a man from B_{low}.

Lemma 24. *Let w^* be a woman. We define S, x, z and δ as above. Conditioning on S, for every sampling of the remaining randomness we have: $\delta \leq x + z$.*

Proof. Let m^* be w^*'s worst stable husband ($m^* = \mu_\bot(w^*)$). Observe that w^* ranks every man who already proposed to w^* (during Algorithm 1) after m^*; those men are represented by squares in Figure 4. Moreover, by definition of blocks, w^* ranks men $\{m_{t+1}, \ldots, m_N\}$ after m^*; those men are represented by white shapes in Figure 4.

Thus, every man ranked between $\mu_\top(w^*)$ and $\mu_\bot(w^*)$ has not proposed to w^* yet, and is in or above her block. We bound the number of such men: in her block by x, and above her block by z. □

Lemma 25. *We have $\mathbb{E}[z] = O(1)$ and z has an exponential tail.*

Proof. Consider N objects $\{o_1, \ldots o_N\}$. For all $1 \leq i < N$, item o_i has a popularity of λ^i; and item o_N has a popularity of $\lambda^N/(1-\lambda)$. We use the popularities to build a preference list over the objects, and we write $1 \leq r \leq N$ the rank of object o_N. Let $\tilde{z} = N - r$. We are going to prove that:

- Conditioning on S, we have $\tilde{z} \geq z$.
- We have $\mathbb{E}[\tilde{z}] = O(1)$ and $\tilde{z}$ has an exponential tail.

To prove the stochastic domination, observe that because no man from B proposed to w^* yet, their ordering in w^*'s preference list is independent from S. To compute z, we sample men from B up until a man from B_{low} is sampled, then we count the number of men from B_{high} that have not been sampled. In this process, notice that we replace all men from B_{low} by one virtual man whose popularity is equal to the sum of their popularities. Thus z is equal to $|B_{high}| + 1$ minus the rank of this virtual man. Without loss of generality, we can multiply every popularity by $\lambda^{N-1-t'}$. Now, the popularities of men from B_{high} are a subset of $\{\lambda^1, \dots, \lambda^{N-1}\}$. The sum of popularities of men from B_{low} is at most $\lambda^N/(1-\lambda)$. Thus, by construction $\tilde{z} \geq z$.

Now let us bound $\tilde{z}$. For all $1 \leq i < N$, we have $o_N \succ o_i$ with probability:

$$\frac{\lambda^N/(1-\lambda)}{\lambda^i + \lambda^N/(1-\lambda)} = \frac{\lambda^{N-i}}{1-\lambda+\lambda^{N-i}} < \frac{\lambda^{N-i}}{1-\lambda}$$

Therefore, by linearity of the expectation:

$$\mathbb{E}[\tilde{z}] = \sum_{i=1}^{N-1} \mathbb{P}[o_N \succ o_i] < \sum_{i=1}^{N-1} \frac{\lambda^{N-i}}{1-\lambda} < \frac{\lambda}{(1-\lambda)^2} = O(1)$$

Moreover, for all $k \geq 1$, if $\tilde{z} \geq k$ then there is at least an object o_i with $i \in \{1, \dots, N-k\}$ such that $o_N \succ o_i$. Using the union-bound:

$$\forall k \geq 1, \quad \mathbb{P}[\tilde{z} \geq k] \leq \sum_{i=1}^{N-k} \mathbb{P}[o_N \succ o_i] < \frac{\lambda^k}{(1-\lambda)^2} = \exp(-\Omega(k))$$

□

4.5 Weakly geometric popularities

All the analysis of this section also goes through when women have independent (but not necessarily identical) weakly geometric popularity preferences (with two constant parameters A and B).

Definition 6 (Weakly geometric popularity preferences)**.** We say that a woman w has *weakly geometric popularity preferences* if her distribution $\mathcal{D}$ satisfies that:

- there is a constant $A \geq 1$ such that for all i, j with $j > i + A$, we have $D(m_j) \leq \frac{1}{2}\mathcal{D}(m_i)$.
- there is a constant $B \geq 1$ such that for all i, j with $j > i$, we have $D(m_j) \leq B \cdot \mathcal{D}(m_i)$.

5 UNIFORMLY RANDOM PROPOSALS: PROOF OF THEOREM 3

In section 4, we studied the case where men have geometric popularities that are used to construct the women's preference lists. We showed that if we know the men-optimal stable matching $\mu_\perp$, then we can order the women (woman w_i being married with m_i in $\mu_\perp$) and decompose the instance into blocks, so that each person has all of their stable partners within their block. Corollary 21 shows that the distribution of the size of a block has an exponential tail.

In this section, we discuss which additional results can be obtained when in addition women have all equal popularities, that are used to construct the men's preference lists.

Theorem 3. *Assume that the men have uniform popularity preferences and the women have geometric popularity preferences. Then the expected total number of stable pairs is $N + O(\ln^3 N)$. Moreover, a woman using a non-truthful strategy will not improve the rank of her partner in her true preference list, with high probability.*

To prove the theorem, we will proceed as follows: Sample the preference lists of all the women. Run Algorithm 1, with men sampling their preferences online. Let $\mu_\perp$ be the resulting men-optimal stable matching and relabel the woman according to $\mu_\perp$. Record the list of all proposals made by the men so far. Use the analysis of section 4 to define blocks.

5.1 Typical instances

We start our analysis by assuming that the instance has nice properties. Then for each woman $w^* \in \mathcal{W}$, we enumerate all stable husbands of w^* using Algorithm 2, with men sampling their preferences online.

Definition 7. Let $C = O(1)$ be a constant to be defined later. Let OK denote the event that all blocks have size at most $C \ln N$, and every woman prefers man m_i to man m_j whenever $j \geq i + C \ln N$.

Proof of Theorem 3. Let Y be the number of stable pairs. By Lemma 26, OK has probability at least $1 - 1/N^2$. When OK does not hold, we use the bound $Y \leq N^2$, so we have:

$$\mathbb{E}[Y] = \underbrace{\mathbb{P}[\text{not } OK]}_{\leq 1/N^2} \cdot \underbrace{\mathbb{E}[Y \mid \text{not } OK]}_{\leq N^2} + \underbrace{\mathbb{P}[OK]}_{\leq 1} \cdot \mathbb{E}[Y \mid OK] \leq 1 + \mathbb{E}[Y \mid OK]$$

Using Lemma 28 we write:

$$\mathbb{E}[Y \mid OK] = \sum_{\mu \text{ matching}} \mathbb{P}[\mu_\perp = \mu \mid OK] \cdot \mathbb{E}[Y \mid \mu_\perp = \mu \text{ and } OK] = N + O(\ln^3 N).$$

Hence Theorem 3. □

Lemma 26. *The probability of event OK is $\geq 1 - 1/N^2$.*

Proof. For the first case of failure, recall from Corollary 21 that the size of a block has an exponential tail. Thus we can choose C such that the probability of a given block has a size greater than $C \log N$ is at most $1/(2N^3)$. There are at most N blocks, using the union bound the probability that at least one has a size exceeding $C \log N$ is at most $1/(2N^2)$. For the second case of failure, notice that the probability for a woman to prefer a man m_j to a another man m_i is $\mathcal{D}(m_j)/(\mathcal{D}(m_j) + \mathcal{D}(m_i)) \leq \lambda^{j-i} \leq \lambda^{C \ln N}$ when $i + C \ln N \leq j$. Thus we can choose C such that the probability of this happening is smaller than $1/(2N^5)$. Using the union bound over all triples of woman/m_i/m_j, the probability of a failure is at most $1/(2N^2)$. Choosing C maximal between the two values, and using the union bound over the two possible cases of failure, the probability that OK does not holds is at most $1/N^2$. □

5.2 The sequence of proposal in Algorithm 2 has a "no-return" property

Recall that to check that event OK holds, we sampled the preference lists of all women. The only remaining source of randomness is the remainder of the preference lists of men, each time a man needs to propose he will sample a woman uniformly from the set of women he has not proposed to yet.

Lemma 27. *Fix $i \in [1, N]$. Conditioning on $\mu_\perp$ and on which proposals have already been made by the men in Algorithm 1, assuming OK holds, the probability that woman w_i has more than one stable husband is at most $3C \ln N/(N + C \ln N - i)$.*

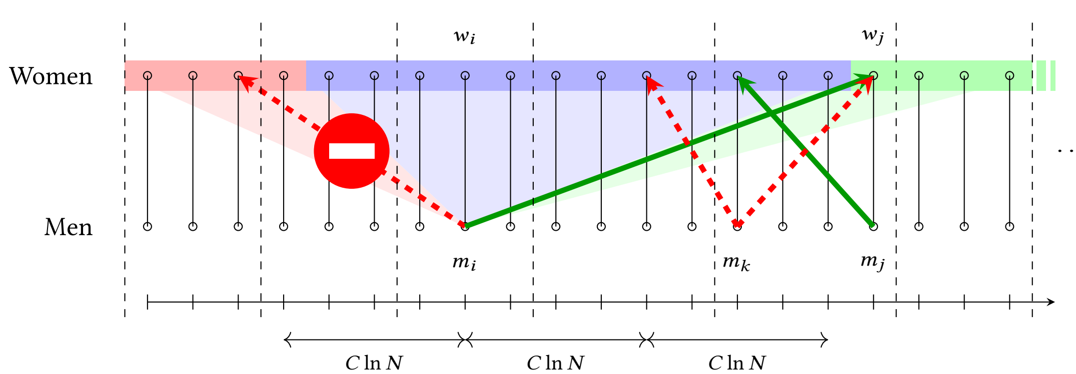


Fig. 5. Proof of Lemma 27, probability that w_i has several stable husbands ≤ ratio $\frac{\cdots}{\cdots + \cdots}$

PROOF. We run Algorithm 2 with $w^* := w_i$, so man m_i is the initial proposer. Say that a woman w_j is "red" if $j \leq i - C \ln N$, "yellow" if $i - C \ln N < j \leq i + 2C \ln N$, and "green" if $i + 2C \ln N < j$.

If m_i proposes to a red woman then he is automatically rejected (event OK), so consider the first proposal m_i does to a yellow or green woman. Let B be the number of yellow women to whom m_i has not proposed yet, and G be the number of green women. Man m_i has not proposed to any green woman (event OK).

If this proposal is made to a green woman, then she accepts (event OK). We claim that in that case m_i is the unique stable husband of w_i. To see that, let us first prove by induction on time from that point on during the rest of the execution of Algorithm 2 the following property:

(P): The proposer's block is no better (in terms of popularity) than w_j's block.

- Initially the proposer is m_j, and he is in w_j's block, so (P) holds.
- Inductively assume that (P) holds for the current proposer, a man m_k. The block of w_j has size at most $C \ln N$ (event OK), thus $k > j - C \ln N > i + C \ln N$. If m_k proposes to w_j, he is rejected because w_j is currently married to m_i (event OK). If m_k proposes to a woman in a better block than his, he is rejected by definition of blocks. Thus m_k can only be accepted by a woman currently married to a man whose block is not better than the block of w_j, and so property (P) also holds for the next proposer.

Thus (P) holds throughout the remainder of the execution, so the current proposer is always in a block worse than the block of w^*, and so will never get accepted by woman w^*; therefore w^* only has one stable husband, as claimed.

Thus the probability that woman w_i has more than one stable husband is at most the probability that m_i's first proposal to a yellow or green woman is addressed to a yellow woman. By the uniform assumption, that probability is $B/(B+G)$. Since $B \leq 3C \ln N$ and $G \geq N - i - 2C \ln N$ we have: $\Pr(w_i \text{ has more than one stable husband}) \leq \dfrac{B}{B+G} \leq \dfrac{3C \ln N}{3C \ln N + G} \leq \dfrac{3C \ln N}{N + C \ln N - i}$. □

Lemma 28. *Conditioning on $\mu_\perp$ and which proposals have already been made in Algorithm 1, assuming that event OK holds, the expected number of stable pairs is $N + O(\ln^3 N)$.*

PROOF. We proved in Lemma 27 that the probability that a woman w_i has more than one stable husband is at most $3C \ln N/(N + C \ln N - i)$. Thus, the expected number of woman with more than one stable husband is bounded by

$$\sum_{i=1}^{N} \frac{3C \ln N}{N + C \ln N - i} = \sum_{i=0}^{N-1} \frac{3C \ln N}{i + C \ln N} \leq 3C \ln N \int_{C \log N - 1}^{C \log N - 1 + N} \frac{\mathrm{d}t}{t} = 3C \ln N \ln \left(\frac{C \log N - 1 + N}{C \log N - 1} \right)$$

When N is large enough, we can simplify this bound to $3C \ln^2 N$. As event OK holds, we can say that each of those women has at most $C \ln N$ stable husbands. Conditioning on the value of $\mu_\perp$, the expected number of stable pairs when event OK holds is at most $N + 3C^2 \ln^3 N$. □

5.3 Bounded popularity preferences

All the analysis of this section also goes through when women have independent (but not necessarily identical) weakly geometric popularity preferences (see Definition 6) and men have independent (but not necessarily identical) bounded popularity preferences (with a constant parameter A).

Definition 8 (Bounded popularity preferences)**.** We say that a man m has *bounded popularity preferences* (with parameter $A \geq 1$) if in his distribution $\mathcal{D}$ over the women, the ratio between the maximum and minimum popularities is at most A.

REFERENCES

[1] Atila Abdulkadiroğlu, Parag A Pathak, and Alvin E Roth. 2005. The new york city high school match. *American Economic Review* 95, 2 (2005), 364–367.

[2] Atila Abdulkadiroğlu, Parag A Pathak, Alvin E Roth, and Tayfun Sönmez. 2005. The Boston public school match. *American Economic Review* 95, 2 (2005), 368–371.

[3] Itai Ashlagi, Yash Kanoria, and Jacob D Leshno. 2017. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy* 125, 1 (2017), 69–98.

[4] Abhijit Banerjee, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune. 2013. Marry for what? Caste and mate selection in modern India. *American Economic Journal: Microeconomics* 5, 2 (2013), 33–72.

[5] Jose Correa, Rafael Epstein, Juan Escobar, Ignacio Rios, Bastian Bahamondes, Carlos Bonet, Natalie Epstein, Nicolas Aramayo, Martin Castillo, Andres Cristi, et al. 2019. School Choice in Chile. *Proceedings of the 2019 ACM Conference on Economics and Computation* (2019), 325–343.

[6] Manlove David. 2013. *Algorithmics of matching under preferences*. World Scientific.

[7] Gabrielle Demange, David Gale, and Marilda Sotomayor. 1987. A further note on the stable matching problem. *Discrete Applied Mathematics* 16, 3 (1987), 217–222.

[8] Lester E Dubins and David A Freedman. 1981. Machiavelli and the Gale-Shapley algorithm. *The American Mathematical Monthly* 88, 7 (1981), 485–494.

[9] Pavlos S Efraimidis and Paul G Spirakis. 2006. Weighted random sampling with a reservoir. *Inform. Process. Lett.* 97, 5 (2006), 181–185.

[10] David Gale and Lloyd S Shapley. 1962. College admissions and the stability of marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15.

[11] Dan Gusfield. 1987. Three fast algorithms for four problems in stable marriage. *SIAM J. Comput.* 16, 1 (1987), 111–128.

[12] Dan Gusfield and Robert W Irving. 1989. *The stable marriage problem: structure and algorithms*. MIT press.

[13] Avinatan Hassidim, Assaf Romm, and Ran I Shorrer. 2018. Need vs. merit: The large core of college admissions markets. (2018).

[14] Gunter J Hitsch, Ali Hortaçsu, and Dan Ariely. 2010. Matching and sorting in online dating. *American Economic Review* 100, 1 (2010), 130–63.

[15] Nicole Immorlica and Mohammad Mahdian. 2015. Incentives in large random two-sided markets. *ACM Transactions on Economics and Computation* 3, 3 (2015), 14.

[16] Anna R Karlin, Shayan Oveis Gharan, and Robbie Weber. 2018. A simply exponential upper bound on the maximum number of stable matchings. *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing* (2018), 920–925.

[17] Donald Ervin Knuth. 1976. *Mariages stables et leurs relations avec d'autres problemes combinatoires: introduction a l'analysis mathematique des algorithmes-*. Les Presses de l'Universite de Montreal.

[18] Donald Ervin Knuth. 1997. *Stable marriage and its relation to other combinatorial problems: An introduction to the mathematical analysis of algorithms*. Vol. 10. American Mathematical Soc.

[19] Donald E Knuth, Rajeev Motwani, and Boris Pittel. 1990. Stable husbands. *Proceedings of the first annual ACM-SIAM symposium on Discrete algorithms* (1990), 397–404.

[20] Fuhito Kojima and Parag A Pathak. 2009. Incentives and stability in large two-sided matching markets. *American Economic Review* 99, 3 (2009), 608–27.

[21] SangMok Lee. 2016. Incentive compatibility of large centralized matching markets. *The Review of Economic Studies* 84, 1 (2016), 444–463.

[22] David G McVitie and Leslie B Wilson. 1971. The stable marriage problem. *Commun. ACM* 14, 7 (1971), 486–490.

[23] Parag A Pathak and Tayfun Sönmez. 2008. Leveling the playing field: Sincere and sophisticated players in the Boston mechanism. *American Economic Review* 98, 4 (2008), 1636–52.

[24] Boris Pittel. 1989. The average number of stable matchings. *SIAM Journal on Discrete Mathematics* 2, 4 (1989), 530–549.

[25] Boris Pittel. 1992. On likely solutions of a stable marriage problem. *The Annals of Applied Probability* (1992), 358–401.

[26] Boris Pittel. 2018. On likely solutions of the stable matching problem with unequal numbers of men and women. *Mathematics of Operations Research* (2018).

[27] Boris Pittel, Larry Shepp, and Eugene Veklerov. 2007. On the number of fixed pairs in a random instance of the stable marriage problem. *SIAM Journal on Discrete Mathematics* 21, 4 (2007), 947–958.

[28] Alvin E Roth. 1982. The economics of matching: Stability and incentives. *Mathematics of operations research* 7, 4 (1982), 617–628.

[29] Alvin E Roth and Elliott Peranson. 1999. The redesign of the matching market for American physicians: Some engineering aspects of economic design. *American economic review* 89, 4 (1999), 748–780.

[30] Edward G Thurber. 2002. Concerning the maximum number of stable matchings in the stable marriage problem. *Discrete Mathematics* 248, 1-3 (2002), 195–219.

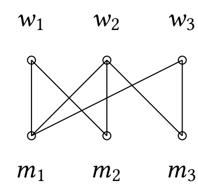
A BACKGROUND (DETAILS)

Let us give a detailed execution of Algorithms 1 and 2 on the following instance.

$$\begin{array}{ll} m_1 : w_1 \succ_{m_1} w_2 \succ_{m_1} w_3 & \quad w_1 : m_2 \succ_{w_1} m_1 \succ_{w_1} m_3 \\ m_2 : w_2 \succ_{m_2} w_3 \succ_{m_2} w_1 & \quad w_2 : m_3 \succ_{w_2} m_1 \succ_{w_2} m_2 \\ m_3 : w_3 \succ_{m_3} w_2 \succ_{m_3} w_1 & \quad w_3 : m_1 \succ_{w_3} m_3 \succ_{w_3} m_2 \end{array}$$

For this instance, there are 3 stable matchings. The stable graph has 7 edges.

$$\begin{aligned} \mu_\perp &= \{(m_1, w_1),\ (m_2, w_2),\ (m_3, w_3)\} \\ &\quad \{(m_1, w_2),\ (m_2, w_1),\ (m_3, w_3)\} \\ \mu_\top &= \{(m_1, w_3),\ (m_2, w_1),\ (m_3, w_2)\} \end{aligned}$$



When running Algorithm 1, m_1 proposes to w_1 and she accepts, m_2 proposes to w_2 and she accepts, m_3 proposes to w_3 and she accepts. Starting from the matching $\mu_\perp$, let us run Algorithm 2 to enumerate the stable husbands of w_1, w_2 and w_3.

Stable husbands of w_1

Proposition	Answer
$m_1 \to w_2$	yes
$m_2 \to w_3$	no
$m_2 \to w_1$	**yes**
$m_2 \to \emptyset$	

Stable husbands of w_2

Proposition	Answer
$m_2 \to w_3$	no
$m_2 \to w_1$	yes
$m_1 \to w_2$	**yes**
$m_1 \to w_3$	yes
$m_3 \to w_2$	**yes**
$m_3 \to w_1$	no
$m_3 \to \emptyset$	

Stable husbands of w_3

Proposition	Answer
$m_3 \to w_2$	yes
$m_2 \to w_3$	**no**
$m_2 \to w_1$	yes
$m_1 \to w_2$	no
$m_1 \to w_3$	**yes**
$m_1 \to \emptyset$	

Note that when enumerating the stable husbands of w_3 we did not enumerate the stable husbands of the other women. For example, w_2 rejected the proposition from m_1 because she already had a proposition from m_3. Algorithm 2 has been introduced in [19]. It is a simplification of an algorithm that enumerates all stable husbands at the same time [11, 22].

In literature, Algorithm 2 is often described using a "break-marriage" operation. If we want to compute all stable husbands of a woman w^*, first we compute the stable matching $\mu_\perp$ using the men-proposing deferred acceptance algorithm, finding her first husband. Then we declare this pair unacceptable, and continue the deferred acceptance algorithm until a new stable matching is found. We repeat this process, an "outer" loop regularly breaking the marriage of w^*, in order to find all her possible stable partners.

One can notice that after a new stable husband is found, he will be the next person to propose, as the next step is to break w^*'s marriage. Thus we decided to merge the "outer" and "inner" loops. This choice has two effects. First it might possible that this version is a bit harder to apprehend (if one wants to prove that Algorithm 2 effectively outputs the set of all stable husbands). However the resulting procedure is much easier to explain, and the transition from the algorithm to the proofs (see for example Lemma 10) is genuinely more natural.

B DEFERRED ACCEPTANCE (PROOFS)

Lemma 11. *Let $\mu_\perp$ denote the matching obtained at the end of Algorithm 1 and $\mu_\top(w^*)$ denote the last man accepted by w^* during the execution of Algorithm 2 (her favorite stable husband). Let K be the number of proposals received by w^* in Algorithm 2. Then:*

$$\mathbb{E}[\textit{Nb of stable husbands of } w^* \mid \mu_\perp] \leq 1+\mathbb{E}[\ln(K+1) \mid \mu_\perp]+\mathbb{E}\left[\ln(\mathcal{D}(\mu_\top(w^*))) \mid \mu_\perp\right]-\ln(\mathcal{D}(\mu_\perp(w^*)))$$

Proof. In this proof we use notations from Lemma 10. Let $x_0 = \mu_\perp(w^*)$, and let $x_1, \ldots, x_K$ be the sequence of proposals received by w_* during Algorithm 2. Let $p_\perp$ denote the sum of popularities of proposals received by w^* during Algorithm 1 (including x_0) and let $p_i = \mathcal{D}(x_i)$ for all $0 \leq i \leq K$. We note Y_{w^*} the number of stable husbands of w^*, and for all $1 \leq i \leq K$ we note X_i the 0/1 random variable equal to 1 when "proposal x_i is accepted by w^*". We know that $Y_{w^*} = 1 + X_1 + \cdots + X_K$. We use the linearity of expectation, and compute the expected value of each Bernoulli trial.

$$\mathbb{E}[Y_{w^*} \mid \mu_\perp] = 1 + \sum_{i=1}^{K} \mathbb{E}[X_i \mid \mu_\perp] = 1 + \sum_{i=1}^{K} \mathbb{P}[X_i = 1 \mid \mu_\perp] = 1 + \sum_{i=1}^{K} \frac{p_i}{p_\perp + p_1 + \cdots + p_i}$$

Then, we compare the sum to an integral.

$$\sum_{i=1}^{K} \frac{p_i}{p_\perp + p_1 + \cdots + p_i} \leq \sum_{i=1}^{K} \int_{p_\perp+p_1+\cdots+p_{i-1}}^{p_\perp+p_1+\cdots+p_i} \frac{\mathrm{d}t}{t} = \ln(p_\perp + p_1 + \cdots + p_K) - \ln p_\perp$$

Still conditioning on $\mu_\perp$, man $\mu_\top(w^*)$ is the overall best proposition received by w^*. It is x_0 with probability proportional to $p_\perp$ and it is x_i $(1 \leq i \leq K)$ with probability proportional to p_i.

$$\begin{aligned}\mathbb{E}[\ln(\mathcal{D}(\mu_\top(w^*))) \mid \mu_\perp] &= \frac{p_\perp \ln p_0 + p_1 \ln p_1 + \cdots + p_K \ln p_K}{p_\perp + p_1 + \cdots + p_K} \\ &= \underbrace{\frac{p_\perp \ln p_\perp + p_1 \ln p_1 + \cdots + p_K \ln p_K}{p_\perp + p_1 + \cdots + p_K}}_{S} + \frac{p_\perp \ln p_0 - p_\perp \ln p_\perp}{p_\perp + p_1 + \cdots + p_K}\end{aligned}$$

Using the convexity of $t \mapsto t \ln t$ and Jensen's inequality, we give a lower-bound on S.

$$\begin{aligned}S &= \frac{K+1}{p_\perp + p_1 + \cdots + p_K} \cdot \frac{p_\perp \ln p_\perp + p_1 \ln p_1 + \cdots + p_K \ln p_K}{K+1} \\ S &\geq \frac{K+1}{p_\perp + p_1 + \cdots + p_K} \cdot \left(\frac{p_\perp + p_1 + \cdots + p_K}{K+1}\right) \ln\left(\frac{p_\perp + p_1 + \cdots + p_K}{K+1}\right) \\ S &\geq \ln(p_\perp + p_1 + \cdots + p_K) - \ln(K+1)\end{aligned}$$

Gathering all four equations, we obtain

$$\begin{aligned}&\mathbb{E}[Y_{w^*} \mid \mu_\perp] \leq 1 + \ln(p_\perp + p_1 + \cdots + p_K) - \ln p_\perp \\ &\mathbb{E}[Y_{w^*} \mid \mu_\perp] \leq 1 + \ln(K+1) + S - \ln p_\perp \\ &\mathbb{E}[Y_{w^*} \mid \mu_\perp] \leq 1 + \ln(K+1) + \mathbb{E}[\ln(\mathcal{D}(\mu_\top(w^*))) \mid \mu_\perp] - \frac{p_\perp \ln p_0 - p_\perp \ln p_\perp}{p_\perp + p_1 + \cdots + p_K} - \ln p_\perp\end{aligned}$$

Finally, using the fact that $p_0 \leq p_\perp$, we can bound the last term.

$$\frac{p_\perp \ln p_0 - p_\perp \ln p_\perp}{p_\perp + p_1 + \cdots + p_K} + \ln p_\perp = \frac{(p_\perp) \ln p_0 + (p_1 + \cdots + p_K) \ln p_\perp}{(p_\perp) + (p_1 + \cdots + p_K)} \geq \ln p_0 = \ln(\mathcal{D}(\mu_\perp(w^*)))$$

Combining the last two inequalities concludes the proof. □

C INSTANCE DECOMPOSITION (PROOFS)

Before proving Lemma 19, let us start with a direct proof giving explicit bounds on $\mathbb{E}[\tilde{x}_1^2]$.

Corollary 20. *The second moment of $\tilde{x}_1$ is bounded, independently of N. That is, $E[\tilde{x}_1^2] = O(1)$.*

Proof. By definition, $\tilde{\Delta}_0, \ldots, \tilde{\Delta}_{\tilde{s}-1}$ are *i.i.d.*, thus

$$\mathbb{E}[(\tilde{x}_1 - 1)^2] = \sum_{k=0}^{+\infty} \mathbb{P}[\tilde{s} = k] \cdot \mathbb{E}\left[\left(\textstyle\sum_{i=0}^{k-1} \tilde{\Delta}_i\right)^2 \,\middle|\, \forall i \in [0, k-1],\ \tilde{\Delta}_i > 0\right]$$

$$= \sum_{k=0}^{+\infty} \mathbb{P}[\tilde{s} = k] \cdot \left(k(k-1) \cdot \mathbb{E}[\tilde{\Delta}_0 \mid \tilde{\Delta}_0 > 0]^2 + k \cdot \mathbb{E}[\tilde{\Delta}_0^2 \mid \tilde{\Delta}_0 > 0]\right)$$

We compute the expected value of $\tilde{\Delta}_0$ and $\tilde{\Delta}_0^2$ conditioning on the fact that $\tilde{\Delta}_0 > 0$.

$$\mathbb{E}[\tilde{\Delta}_0 \mid \tilde{\Delta}_0 > 0] = \frac{\mathbb{E}[\tilde{\Delta}_0]}{\mathbb{P}[\tilde{\Delta}_0 > 0]} \quad \text{and} \quad \mathbb{E}[\tilde{\Delta}_0^2 \mid \tilde{\Delta}_0 > 0] = \frac{\mathbb{E}[\tilde{\Delta}_0^2]}{\mathbb{P}[\tilde{\Delta}_0 > 0]}$$

The random variable $\tilde{s}$ follows a geometric distribution of success parameter $\mathbb{P}[\tilde{\Delta}_0 = 0]$.

$$\forall k \geq 0, \quad \mathbb{P}[\tilde{s} = k] = \mathbb{P}[\tilde{\Delta}_0 = 0] \cdot \mathbb{P}[\tilde{\Delta}_0 > 0]^k$$

Using those values in the expression of $\mathbb{E}[(\tilde{x}_1 - 1)^2]$, we obtain

$$\mathbb{E}[(\tilde{x}_1 - 1)^2] = \sum_{k=0}^{+\infty} \mathbb{P}[\tilde{\Delta}_0 = 0] \cdot \mathbb{P}[\tilde{\Delta}_0 > 0]^k \left(\frac{k(k-1) \cdot \mathbb{E}[\tilde{\Delta}_0]^2}{\mathbb{P}[\tilde{\Delta}_0 > 0]^2} + \frac{k \cdot \mathbb{E}[\tilde{\Delta}_0^2]}{\mathbb{P}[\tilde{\Delta}_0 > 0]}\right)$$

Splitting the sum in two, and starting respectively at $k = 2$ and $k = 1$

$$\mathbb{E}[(\tilde{x}_1 - 1)^2] = \mathbb{E}[\tilde{\Delta}_0]^2 \cdot \mathbb{P}[\tilde{\Delta}_0 = 0] \cdot \underbrace{\sum_{k=0}^{+\infty} (k+2)(k+1) \cdot \mathbb{P}[\tilde{\Delta}_0 > 0]^k}_{= 2/\mathbb{P}[\tilde{\Delta}_0 = 0]^3} + \mathbb{E}[\tilde{\Delta}_0^2] \cdot \mathbb{P}[\tilde{\Delta}_0 = 0] \cdot \underbrace{\sum_{k=0}^{+\infty} (k+1) \cdot \mathbb{P}[\tilde{\Delta}_0 > 0]^k}_{= 1/\mathbb{P}[\tilde{\Delta}_0 = 0]^2}$$

Finally we obtain a formula similar to Wald's equation.

$$\mathbb{E}[(\tilde{x}_1 - 1)^2] = \frac{2\mathbb{E}[\tilde{\Delta}_0]^2}{\mathbb{P}[\tilde{\Delta}_0 = 0]^2} + \frac{\mathbb{E}[\tilde{\Delta}_0^2]}{\mathbb{P}[\tilde{\Delta}_0 = 0]} \quad \text{and} \quad \mathbb{E}[\tilde{x}_1] = \frac{\mathbb{E}[\tilde{\Delta}_0]}{\mathbb{P}[\tilde{\Delta}_0 = 0]}$$

From the definition of $\tilde{\Delta}_0$, for all $k \geq 0$ we have $\mathbb{P}[\tilde{\Delta}_0 \geq k] = 1 - \exp\left(-\frac{\lambda^k}{(1-\lambda)^2}\right) \leq \frac{\lambda^k}{(1-\lambda)^2}$.

$$\mathbb{E}[\tilde{\Delta}_0] = \sum_{k=1}^{+\infty} \mathbb{P}[\tilde{\Delta}_0 \geq k] \leq \sum_{k=1}^{+\infty} \frac{\lambda^k}{(1-\lambda)^2} = \frac{\lambda}{(1-\lambda)^3}$$

$$\mathbb{E}[\tilde{\Delta}_0^2] = \sum_{k=0}^{+\infty} k^2 \cdot (\mathbb{P}[\tilde{\Delta}_0 \geq k] - \mathbb{P}[\tilde{\Delta}_0 \geq k+1]) = \sum_{k=1}^{+\infty} (2k-1) \cdot \mathbb{P}[\tilde{\Delta}_0 \geq k] \leq \sum_{k=1}^{+\infty} \frac{(2k-1)\lambda^k}{(1-\lambda)^2} = \frac{\lambda + \lambda^2}{(1-\lambda)^4}$$

Using those values in the expression of $\mathbb{E}[\tilde{x}_1^2] = \mathbb{E}[(\tilde{x}_1 - 1)^2] + 2\mathbb{E}[\tilde{x}_1] - 1$, we obtain $\mathbb{E}[\tilde{x}_1^2] = O(1)$. □

Lemma 19. *The radius of convergence of the probability generating function of $\tilde{x}_1$ is > 1.*

Proof. For any random variable X, we can define $G_X(z) = \mathbb{E}[z^X]$ for all real z such that $|z| < 1$.

$$G_{\tilde{x}_1}(z) = \mathbb{E}[z^{\tilde{x}_1}] = \sum_{k=0}^{+\infty} \mathbb{P}[\tilde{s} = k] \cdot \mathbb{E}\left[z^{1+\sum_{i=0}^{k-1} \tilde{\Delta}_i} \,\middle|\, \forall i \in [0, k-1],\ \tilde{\Delta}_i > 0\right]$$

Using the fact that all $\tilde{\Delta}_i$'s are *i.i.d.* we can simplify the expectation of the product.

$$G_{\tilde{x}_1}(z) = z \cdot \sum_{k=0}^{+\infty} \mathbb{P}[\tilde{s} = k] \cdot \mathbb{E}\left[z^{\tilde{\Delta}_0} \,\middle|\, \tilde{\Delta}_0 > 0\right]^k = z \cdot G_{\tilde{s}}\left(\mathbb{E}\left[z^{\tilde{\Delta}_0} \,\middle|\, \tilde{\Delta}_0 > 0\right]\right)$$

The conditional expectation can be expressed as follows.

$$G_{\tilde{\Delta}_0}(z) = \mathbb{E}\left[z^{\tilde{\Delta}_0}\right] = \mathbb{P}[\tilde{\Delta}_0 > 0] \cdot \mathbb{E}\left[z^{\tilde{\Delta}_0} \,\middle|\, \tilde{\Delta}_0 > 0\right] + \mathbb{P}[\tilde{\Delta}_0 = 0]$$

$$\mathbb{E}\left[z^{\tilde{\Delta}_0} \,\middle|\, \tilde{\Delta}_0 > 0\right] = \frac{G_{\tilde{\Delta}_0}(z) - \mathbb{P}[\tilde{\Delta}_0 = 0]}{\mathbb{P}[\Delta_0 > 0]}$$

Now let us compute the generating function of $\tilde{s}$

$$G_{\tilde{s}}(z) = \mathbb{E}[z^{\tilde{s}}] = \sum_{k=0}^{+\infty} z^k \cdot \mathbb{P}[\tilde{s} = k] = \sum_{k=0}^{+\infty} z^k \cdot \mathbb{P}[\tilde{\Delta}_0 > 0]^k \cdot \mathbb{P}[\tilde{\Delta}_0 = 0] = \frac{\mathbb{P}[\tilde{\Delta}_0 = 0]}{1 - z \cdot \mathbb{P}[\tilde{\Delta}_0 > 0]}$$

Combining the three previous equations we obtain

$$G_{\tilde{x}_1}(z) = \frac{z \cdot \mathbb{P}[\tilde{\Delta}_0 = 0]}{1 + \mathbb{P}[\tilde{\Delta}_0 = 0] - G_{\tilde{\Delta}_0}(z)}$$

Formally, we can compute the generating function of $\tilde{\Delta}_0$ with an Abel transform

$$G_{\tilde{\Delta}_0}(z) = \sum_{k=0}^{+\infty} z^k \cdot \mathbb{P}[\tilde{\Delta}_0 = k] = \sum_{k=0}^{+\infty} z^k \cdot (\mathbb{P}[\tilde{\Delta}_0 > k-1] - \mathbb{P}[\tilde{\Delta}_0 > k])$$

$$= 1 + \sum_{k=0}^{+\infty} \mathbb{P}[\tilde{\Delta}_0 > k] \cdot (z^{k+1} - z^k) = 1 + (z-1) \sum_{k=0}^{+\infty} z^k \cdot \mathbb{P}[\tilde{\Delta}_0 > k]$$

To obtain the convergence radius of $G_{\tilde{\Delta}_0}$, we are going to use an the equivalent of $\mathbb{P}[\tilde{\Delta}_0 > k]$.

$$\mathbb{P}[\tilde{\Delta}_0 > k] = 1 - \exp\left(\frac{-\lambda^{k+1}}{(1-\lambda)^2}\right) \sim \frac{\lambda^{k+1}}{(1-\lambda)^2}$$

Thus the radius of convergence of $G_{\tilde{\Delta}_0}$ is $1/\lambda$. In order $G_{\tilde{x}_1}(z)$ to be well defined, we need to have $G_{\tilde{\Delta}_0}(z) < 1 + \mathbb{P}[\tilde{\Delta}_0 = 0]$. Being a probability generating function, $G_{\tilde{\Delta}_0}$ is increasing when $0 < z < 1$, and tends to 1 when $z \to 1$. Let us give an upped bound for $G_{\tilde{\Delta}_0}$ when $1 \le z < 1/\lambda$. To do that, we are using the fact that the equivalence relation above is also an upper-bound.

$$\sum_{k=0}^{+\infty} z^k \cdot \mathbb{P}[\tilde{\Delta}_0 > k] \le \sum_{k=0}^{+\infty} z^k \cdot \frac{\lambda^{k+1}}{(1-\lambda)^2} = \frac{\lambda}{(1-z\lambda)(1-\lambda)^2}$$

Thus a sufficient condition for $G_{\tilde{x}_1}$ to be well defined is to have $z \ge 1$ and

$$G_{\tilde{\Delta}_0}(z) - 1 \le \frac{(z-1)\lambda}{(1-z\lambda)(1-\lambda)^2} < \mathbb{P}[\tilde{\Delta}_0 = 0]$$

The middle is a continuous function on $[1, 1/\lambda)$, thus there is an $\varepsilon > 0$ such that having $1 \le z \le 1+\varepsilon$ is sufficient for the rightmost inequality to hold, and thus for $G_{\tilde{x}_1}(z)$ to be well defined. ☐