

Welfare Distribution in Two-sided Random Matching Markets

Itai Ashlagi*

Mark Braverman[†]

Geng Zhao[‡]

February 16, 2023

Abstract

We study the welfare structure in two-sided large random matching markets. In the model, each agent has a latent personal score for every agent on the other side of the market and her preferences follow a logit model based on these scores. Under a contiguity condition, we provide a tight description of stable outcomes.

First, we identify an intrinsic fitness for each agent that represents her relative competitiveness in the market, independent of the realized stable outcome. The intrinsic fitness values correspond to scaling coefficients needed to make a mutual latent matrix bi-stochastic, where the latent scores can be interpreted as a-priori probabilities of a pair being matched.

Second, in every stable (or even approximately stable) matching, the welfare or the ranks of the agents on each side of the market, when scaled by their intrinsic fitness, have an approximately exponential empirical distribution. Moreover, the average welfare of agents on one side of the market is sufficient to determine the average on the other side.

Overall, each agent’s welfare is determined by a global parameter, her intrinsic fitness, and an extrinsic factor with exponential distribution across the population.

1 Introduction

This paper is concerned with the welfare in random two-sided matching markets. In a two-sided matching market there are two kinds of agents, where each agent has preferences over potential partners of the other kind. We assume that the outcome is stable (Gale and Shapley, 1962), meaning that there are no blocking pairs of agents who would rather match to each other over the their assigned partners.

A large literature initiated by (Gale and Shapley, 1962) has deepened our understanding of two-sided matching markets, generating a blend of rich theory and market designs.¹ Less understood, however are welfare properties in typical markets. We study the welfare structure in matching markets when agents have latent preferences generated according to observed characteristics. Specifically we are interested in the empirical welfare distribution of agents on each side of the market under stable outcomes as well as the relation between the outcomes of each side of the market.

We study this question in large randomly generated markets, which allow for both vertical and horizontal differentiation. The model assumes that every agent has an observed personal score for every other agent in the market, and her preferences follows a Logit model based on these scores. We impose that no agent is a-priori overwhelmingly more desirable than any other agent. We find

*Department of Management Science and Engineering, Stanford University

[†]Department of Computer Science, Princeton University

[‡]Department of Electrical Engineering and Computer Sciences, University of California, Berkeley

¹See, e.g., Roth and Sotomayor (1992); Roth (2018).

that the observed characteristics alone determine the empirical welfare distribution on each side of the market. Moreover, the joint surplus in the market is fixed, and the average welfare of one side of the market is a sufficient statistic to determine the empirical welfare distribution on both sides of the market.

The model we consider has an equal number of men and women. For every man m_i and every woman w_j , we are given non-negative scores a_{ij} and b_{ji} , which can be viewed as generated from observed characteristics. Each man and each woman have strict preference rankings generated independently and proportionally to these latent scores, as in the Logit model.² Equivalently, each man m_i has a latent value from matching with woman w_j that is distributed exponentially with rate a_{ij} (smaller values are considered better).³ Women’s latent values for men are generated similarly.⁴

We identify an intrinsic fitness for each agent that represents her relative competitiveness in the market, independent of the realized stable outcome. For every pair of agents on opposing sides of the market, we can obtain a mutual score of the pair’s match. If we write these scores in a matrix, the intrinsic fitness values correspond to scaling coefficients that make the mutual matrix bi-stochastic.⁵ Intuitively, this bi-stochastic mutual matrix can be thought of as consisting of *a-priori* probabilities of each pair matching. In particular, this representation captures the interactions between the sides of the market. We exploit this representation to further analyze typical realized outcomes in the market.

We find that the welfare, or the ranks of the agents, when scaled by their intrinsic fitness, have an approximately exponential empirical distribution on each side of the market. Moreover, the average welfare of agents on one side of the market is sufficient to determine the average on the other side. Overall, each agent’s welfare can be seen as determined by a global parameter, her intrinsic fitness, and an extrinsic factor with exponential distribution across the population. This characterization holds with high probability in every stable matching. In fact, this structure extends to matchings that are only approximately stable, which can tolerate a vanishing fraction of blocking pairs.

At its core, since our proof needs to apply to all stable matchings (and even to nearly-stable matchings), it is a union bound argument. We use inequalities derived from the integral formula for the probability that a given matching is stable, first introduced by Knuth (1976). The heterogeneous preferences brings great difficulty, which we overcome with a truncation technique to accommodate heavy tails of agents’ outcomes and a fixed-point argument on the eigenspace of the characterizing matrix of the market. The exponential empirical distribution part of the result holds intuitively because there are not too many stable matchings in expectation, and the exponential distribution has the highest entropy of all non-negative distributions with a given mean.

Closely related to our work is the remarkable paper Menzel (2015), which finds that the joint surplus in the market is unique. The focus in Menzel (2015) is on analyzing the matching rates between agents of different types, rather than the rankings and agents’ welfare. Menzel’s preference model is more general.⁶ Menzel establishes that, at the limit, agents choose partners according to a logit model from opportunity sets, while we consider large markets and assume agents’ preferences are logit based. There are several other key differences. First, his model requires many agents

²Numerous empirical papers that study two-sided matching market assume agents’ preferences follow a logit model (see e.g., Agarwal and Somaini (2018); Hitsch et al. (2010)).

³One can view the utility of an agent for her match to be the negative of the corresponding latent value.

⁴Special cases of this general model are markets with uniformly random preferences (Knuth et al., 1990; Pittel, 1989; Knuth, 1997; Pittel, 1992; Ashlagi et al., 2017) or when agents have common public scores (Mauras, 2021; Ashlagi et al., 2021).

⁵This representation is valid since preferences are invariant under such transformations.

⁶We note that both his and our model assume that the ratio between any two systematic scores is bounded.

of each type (with the same characteristics), while every agent in our model may have different characteristics. Second, while in our model every agent is matched, he assumes agents have a non-negligible outside option resulting in a large number of unmatched agents⁷; this assumption allows him to apply a fixed point contraction argument and establish the uniqueness and characterization result.⁸

1.1 Literature

The analysis of random two-sided markets goes back to Knuth et al. (1990); Pittel (1989, 1992), who consider markets with uniformly random complete preference lists. These papers establish the number of stable matchings as well as the average ranks on each side. A key finding is that the product of their average rank of agents on each side of the market is approximately the size of the market (Pittel, 1992), implying that stable matchings essentially lie on a parabola. Our findings generalize these findings to markets to random logit markets. We also expand these findings to describe the distributional outcomes in the market.

Several papers consider markets with uniformly drawn preferences with an unequal number of agents on each side of market (Ashlagi et al., 2017; Pittel, 2019; Cai and Thomas, 2022). A key finding is there is an essentially unique stable matching and agents on the short side have a substantial advantage. We believe that similar findings hold in random logit markets. Since our results hold for approximately stable matches, our findings extend to the imbalanced case as long as the imbalance is not too large.

Our paper further contributes to the above literature by considering also outcomes that are approximately stable outcomes.

Several papers study markets random markets when (at least on one side) agents' preferences are generated proportionally to public scores. (Immorlica and Mahdian, 2015; Kojima and Pathak, 2009; Ashlagi et al., 2014) look at the size of the core.⁹ Their analysis relies on a certain market structure (keeping preference lists short), which leaves many agents unmatched. Gimbert et al. (2019) and Ashlagi et al. (2021) assume agents have complete preference lists and their focus is on the size of the core or agents' average rank.

1.2 Notations

Denote $[n] = \{1, \dots, n\}$. Boldface letters denote vectors (lower case) and matrices (upper case), e.g., $\mathbf{x} = (x_i)_{i \in [n]}$ and $\mathbf{A} = (a_{ij})_{i \in [n], j \in [m]}$, and capital letters denote random variables.

For two identically shaped matrices (or vectors) \mathbf{M} and \mathbf{N} , $\mathbf{M} \circ \mathbf{N}$ denotes their Hadamard (entry-wise) product. For a vector $\mathbf{x} \in \mathbb{R}^n$ with non-zero entries, denote its coordinate-wise inverse by \mathbf{x}^{-1} . $\text{diag}(\mathbf{x})$ denotes the diagonal matrix whose i -th entry on the diagonal is x_i .

$\text{Exp}(\lambda)$ and $\text{Poi}(\lambda)$ denote, respectively, the exponential distribution and the Poisson distribution with rate λ . We denote the probability density function (pdf) and cumulative distribution function (CDF) of $\text{Exp}(\lambda)$ by f_λ and F_λ , respectively. $\text{Bern}(p)$ denotes the Bernoulli distribution with success probability $p \in [0, 1]$. For distributions \mathcal{D}_1 and \mathcal{D}_2 over space \mathcal{X} , $\mathcal{D}_1 \otimes \mathcal{D}_2$ denotes their product distribution over \mathcal{X}^2 . $\hat{\mathcal{F}}(\mathbf{x})$ denotes the empirical distribution function for the components of a vector \mathbf{x} , treated as a function from \mathbb{R} to $[0, 1]$. $\mathcal{F}(\mathcal{D})$ denotes the CDF of a distribution \mathcal{D} on \mathbb{R} . For real-valued random variables X and Y , $X \preceq Y$ denotes stochastic domination of X by Y .

⁷Menzel (2015) identifies how to scale the market under this assumption to capture realistic outcomes.

⁸Technically, such substantial outside options keep rejection chains short and prevent them from cycling.

⁹They further consider the related issue of strategizing under stable matching mechanisms.

We use the standard $O(\cdot)$, $o(\cdot)$, $\Omega(\cdot)$, and $\Theta(\cdot)$ notations to hide constant factors. For functions $f, g : \mathbb{N} \rightarrow \mathbb{R}_+$, we say $f = O(g)$ (resp. $\Omega(g)$) if there exists an absolute constant $K \in (0, \infty)$ such that $f \leq Kg$ (resp. $f \geq Kg$) for n sufficiently large; $f = o(g)$ if $f/g \rightarrow 0$ as $n \rightarrow \infty$; and $f = \Theta(g)$ if $f = O(g)$ and $f = \Omega(g)$. We say $f = o_\alpha(g)$ if $f/g \rightarrow 0$ as $\alpha \rightarrow 0$ (uniformly over all other parameters, such as n). For example, $\sqrt{\epsilon} = o_\epsilon(1)$.

2 Model

We study two-sided matching markets with randomly generated preferences. Next we formalize the model, how preferences are generated and key assumptions.

Setup. A matching market consists of two sets of agents, referred to as men \mathcal{M} and women \mathcal{W} . Unless specified otherwise, we assume that $|\mathcal{M}| = |\mathcal{W}| = n$, men are labeled $\mathbf{m}_1, \dots, \mathbf{m}_n$ and women are labeled $\mathbf{w}_1, \dots, \mathbf{w}_n$. Each man \mathbf{m}_i has a complete strict preference list $\succ_{\mathbf{m}_i}$ over the set of women and each woman \mathbf{w}_j has a complete strict preference list $\succ_{\mathbf{w}_j}$ over the set of men. A *matching* is a bijection $\mu : \mathcal{M} \rightarrow \mathcal{W}$. To simplify the notation, men and women will be presented using the set of integers $[n] = \{1, 2, \dots, n\}$ and we write $\mu : [n] \rightarrow [n]$ so that $\mu(i) = j$ and $\mu^{-1}(j) = i$ means that \mathbf{m}_i is matched with \mathbf{w}_j in μ . The *rank* for man \mathbf{m}_i , denoted by $R_i(\mu)$, is the position of $\mu(i)$ on \mathbf{m}_i 's preference list (e.g., if an agent is matched to the second agent on her list, her rank is two). Write $\mathbf{R}(\mu) := (R_i(\mu))_{i \in [n]}$ for the men's rank vector in matching μ .

The matching μ is *unstable* if there is a pair of man \mathbf{m}_i and woman \mathbf{w}_j such that $\mathbf{w}_j \succ_{\mathbf{m}_i} \mathbf{w}_{\mu(i)}$ and $\mathbf{m}_i \succ_{\mathbf{w}_j} \mathbf{w}_{\mu^{-1}(j)}$. A matching is said to *stable* otherwise. It is well-known that the set of stable matchings is not empty.

Logit-based random markets: the canonical form. We consider markets in which complete preferences are randomly generated as follows. For each man \mathbf{m}_i , we are given a stochastic vector $\hat{\mathbf{a}}_i = (\hat{a}_{ij})_{j \in [n]} \in \mathbb{R}_+^n$. Then, \mathbf{m}_i 's preference list is generated from a logit model based on $\hat{\mathbf{a}}_i$. In particular, let \mathcal{D}_i be the distribution on \mathcal{W} that places on \mathbf{w}_j a probability proportional to \hat{a}_{ij} ; then \mathbf{m}_i samples from \mathcal{D}_i for his favorite partner, and repeatedly sample from it without replacement for his next favorite partner until completing his list. Similarly, each woman \mathbf{w}_j preference list is generated from a logit model based on a given stochastic vector $\hat{\mathbf{b}}_j = (\hat{b}_{ji})_{i \in [n]}$. Denote by $\hat{\mathbf{A}} = (\hat{a}_{ij})_{i,j \in [n]}$ and $\hat{\mathbf{B}} = (\hat{b}_{ji})_{j,i \in [n]}$ the row-stochastic matrices. We refer to this matrix representation of the preference model as the *canonical form* and to \hat{a}_{ij} (resp. \hat{b}_{ji}) as the *canonical score* that \mathbf{m}_i (resp. \mathbf{w}_j) assigns to \mathbf{w}_j (resp. \mathbf{m}_i).

This model captures the multinomial logit (MNL) choice model, in which scores are closely related to the systematic utilities for agents over matches. The special case in which $\hat{a}_{ij} = \hat{b}_{ji} = 1/n$ for all $i, j \in [n]$ corresponds to the uniformly random preference model.

Mutual matrix and intrinsic fitness: the balanced form. While the canonical form is a useful way to describe the market, it will be helpful for the analysis to describe it using an alternative scaling scheme, which we refer to as the *balanced form*.

Observe that multiplying any row of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ by a constant does not change the behavior of the market. We look for scaling vectors $\phi, \psi \in \mathbb{R}_+^n$ for the rows of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ such that $\mathbf{M} = n^{-1} \mathbf{A} \circ \mathbf{B}$ is bistochastic¹⁰, where $\mathbf{A} = \text{diag}(\phi) \hat{\mathbf{A}}$ and $\mathbf{B} = \text{diag}(\psi) \hat{\mathbf{B}}$. As is shown by Sinkhorn (1964,

¹⁰The nonnegative matrix \mathbf{M} is bistochastic if the sum of entries in each row and each column is one.

Theorem 1), such a bistochastic matrix \mathbf{M} always uniquely exists, and the scaling vectors ϕ and ψ are unique up to constant rescaling. That is, ϕ and ψ jointly solve

$$\frac{1}{n} \text{diag}(\phi)(\hat{\mathbf{A}} \circ \hat{\mathbf{B}}^\top) \text{diag}(\psi)\mathbf{1} = \mathbf{1} \quad \text{and} \quad \frac{1}{n} \text{diag}(\psi)(\hat{\mathbf{B}} \circ \hat{\mathbf{A}}^\top) \text{diag}(\phi)\mathbf{1} = \mathbf{1}, \quad (1)$$

where $\mathbf{1}$ is the vector consisting of all 1's. The matrix \mathbf{M} will be referred to as the *mutual matrix*.

In the remainder of the paper we assume without loss of generality that the market is described in the balanced form, using \mathbf{A}, \mathbf{B} and the mutual matrix \mathbf{M} .

The bistochasticity constraint incurs the following relationship: if \hat{b}_{ji} 's increase (resp. decrease) by a factor of α simultaneously for all $j \in [n]$, the scaling factor ϕ_i , and hence all a_{ij} 's for $j \in [n]$, must decrease (resp. increase) by the same factor to maintain bistochasticity of \mathbf{M} . In other words, a uniform increase (resp. decrease) of m_i 's popularity among the women will lead to a proportional decrease (resp. increase) in ϕ_i . Thus, we can view the ϕ_i as reflecting the ‘‘average popularity’’ of man m_i among the women: Loosely speaking, the smaller $\sum_{j=1}^n a_{ij}$ is, the more popular m_i is (reflected by larger values of b_{ji} 's).

We refer to the vector ϕ and ψ as the men's and women's *intrinsic fitness* vector, respectively (and note that a smaller intrinsic fitness value means the agent is more competitive). Note that since $\hat{\mathbf{A}} = \text{diag}(\phi)^{-1}\mathbf{A}$ is row-stochastic, we conveniently have $\phi_i = \sum_{j=1}^n a_{ij}$, and similarly $\psi_j = \sum_{i=1}^n b_{ji}$ in the balanced form.

Example 2.1 (Markets with public scores). *We say a matching market has public scores when $\hat{\mathbf{a}}_i = \hat{\mathbf{a}} \in \mathbb{R}_+^n$ for all $i \in [n]$ and $\hat{\mathbf{b}}_j = \hat{\mathbf{b}} \in \mathbb{R}_+^n$ for all $j \in [n]$. In other words, agents on the same side of the market share an identical preference distribution. The fitness vectors are simply $\phi = \hat{\mathbf{b}}^{-1}$ and $\psi = \hat{\mathbf{a}}^{-1}$, where the inverse is taken component-wise. The mutual matrix $\mathbf{M} = \mathbf{J} := (n^{-1})_{i,j \in [n]}$ in this case.*

Latent values. The logit-based preference model can be generated equivalently in the following way. Let $\mathbf{X}, \mathbf{Y} \in \mathbb{R}_+^{n \times n}$ be two random matrices with independent entries X_{ij} (resp. Y_{ji}) sampled from $\text{Exp}(a_{ij})$ (resp. $\text{Exp}(b_{ji})$). The preference profile is then derived from \mathbf{X} and \mathbf{Y} as follows:

$$\begin{aligned} w_{j_1} \succeq_{m_i} w_{j_2} &\iff X_{ij_1} < X_{ij_2}, \\ m_{i_1} \succeq_{w_j} m_{i_2} &\iff Y_{ji_1} < Y_{ji_2}. \end{aligned}$$

We refer to each X_{ij} (resp. Y_{ji}) for $i, j \in [n]$ as the *latent value* (or simply *value*) of m_i (resp. w_j) if matched with w_j (resp. m_i).

Note that for every agent, a lower rank implies a lower latent value (and therefore lower values of rank and latent value are better).

Regularity assumption. We study the asymptotic behavior of two-sided matching markets as the market size grows large. Informally, we restrict attention to contiguous markets, in the sense that, ex ante, no agent finds any other agent (on the opposite side of the market) disproportionately favorable or unfavorable to other agents. The condition is formalized as follows.

A matrix $\mathbf{L} \in \mathbb{R}^{n \times n}$ with non-negative entries is called *C-bounded* for some constant $C \geq 1$ if $l_{ij} \in [1/C, C]$ for all $1 \leq i, j \leq n$. When \mathbf{L} is (bi-)stochastic, we will abuse notation and say \mathbf{L} is *C-bounded* if $n\mathbf{L}$ satisfies the definition above.

Assumption 2.2 (Contiguity). *We assume that, by choosing an appropriate scaling of ϕ and ψ in the balanced form, there exist absolute constants $C \in [1, \infty)$ and $n_0 < \infty$ such that \mathbf{A} , \mathbf{B} , and $n\mathbf{M} = \mathbf{A} \circ \mathbf{B}^\top$ are all C -bounded for all $n \geq n_0$; that is, there exists $C \in [1, \infty)$ such that*

$$\frac{1}{C} \leq \min_{i,j \in [n]} \min\{a_{ij}, b_{ji}, nm_{ij}\} \leq \max_{i,j \in [n]} \max\{a_{ij}, b_{ji}, nm_{ij}\} \leq C \quad \text{for all } n > n_0. \quad (2)$$

Remark 2.3. *It is easy to verify that Assumption 2.2 holds when no agent finds any potential partner disproportionately favorable or unfavorable based on their canonical scores: If $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are C -bounded, then there exists a choice of ϕ and ψ with all entries in $[n/C^2, nC^2]$ in the balanced form; further, \mathbf{M} is C^4 -bounded. Thus, Assumption 2.2 is equivalent to the existence of an absolute upper bound on the ratio between pairs of entries within the same row of \mathbf{A} or \mathbf{B} ; that is*

$$\limsup_{n \rightarrow \infty} \max_{i,j_1,j_2 \in [n]} \frac{a_{ij_1}}{a_{ij_2}} < \infty \quad \text{and} \quad \limsup_{n \rightarrow \infty} \max_{j,i_1,i_2 \in [n]} \frac{b_{ji_1}}{b_{ji_2}} < \infty. \quad (3)$$

This condition is agnostic to scaling of the matrices and hence easy to certify. However, the lower and upper bounds in (2) are more convenient in our later analysis, where the constant C will make an appearance (although often made implicit in the results).

Remark 2.4. *Assumption 2.2 offers a strong contiguity condition on the market, in that the attractiveness among all pairs of men and women vary at most by an (arbitrarily large) constant factor as the market grows. We expect the results to hold under a weaker assumption, which can be described through the spectral gap of the matrix \mathbf{M} . Recall that, as a bistochastic matrix, \mathbf{M} has a largest eigenvalue of 1 and all other eigenvalues of magnitude at most 1. We may think of the market as contiguous in this weaker sense if the spectral gap of \mathbf{M} , given by $1 - |\lambda_{\max}(\mathbf{M} - \mathbf{J})|$, is bounded away from zero as the market grows. The spectral gap is a common and powerful notion when studying the structure of networks and communities.¹¹ We impose Assumption 2.2 as it simplifies substantially the analysis and exposition.*

3 Main results

We denote the (random) set of stable matchings by \mathcal{S} . Recall that for a matching μ , $\mathbf{X}(\mu)$ and $\mathbf{R}(\mu)$ denote men's value and rank vectors, respectively, under μ . Denote by $\hat{\mathcal{F}}(\mathbf{v})$ the empirical distribution of the components of a vector \mathbf{v} (viewed as a function from \mathbb{R} to $[0, 1]$), and F_λ denotes the CDF of $\text{Exp}(\lambda)$.

Theorem 3.1 (Empirical distribution of values). *For any fixed $\epsilon > 0$,*

$$\mathbb{P}\left(\max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\mathbf{X}(\mu)) - F_\lambda\|_\infty \leq \epsilon\right) \rightarrow 1 \text{ as } n \rightarrow \infty. \quad (4)$$

That is, with high probability, in all stable matchings simultaneously, the empirical distribution of the men's values is arbitrarily close to some exponential distribution $\text{Exp}(\lambda)$ in Kolmogorov-Smirnov

¹¹In our model of the matching market, the spectral gap of \mathbf{M} describes the extent to which the market interconnects globally (contiguity) or decomposes into multiple sub-markets (modularity). A larger spectral gap means that the market is more cohesive, with more uniform or homogeneous preferences. For instance, the uniform market with $\mathbf{M} = \mathbf{J}$ has a unit spectral gap, the maximum possible value. On the other hand, a smaller spectral gap means that the market is more clustered, with a clearer boundary between communities and poorly mixed preferences. For instance, any block-diagonal bistochastic matrix (with more than one blocks) has a zero spectral gap, and corresponds to a market that decomposes into two or more independent sub-markets — one cannot hope to have a uniform structure result in such markets.

norm, where the parameter λ depends on the specific stable matching. In particular, the infimum over λ in (4) can be replaced with the choice of λ that can be computed from the women's value vector $\mathbf{Y}(\mu)$.

Theorem 3.2 (Empirical distribution of ranks). *For any fixed $\epsilon > 0$,*

$$\mathbb{P}\left(\max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\phi^{-1} \circ \mathbf{R}(\mu)) - F_\lambda\|_\infty \leq \epsilon\right) \rightarrow 1 \text{ as } n \rightarrow \infty, \quad (5)$$

where ϕ is the fitness vector (and can be computed as $\phi_i = \sum_{j=1}^n a_{ij}$). That is, with high probability, in all stable matchings simultaneously, the empirical distribution of rescaled ranks of the men is arbitrarily close to some exponential distribution $\text{Exp}(\lambda)$ in Kolmogorov-Smirnov norm, where the parameter λ depends on the specific stable matching (yet the scaling doesn't). Again, we may replace the infimum with the choice of λ that can be computed from the women's latent value vector $\mathbf{Y}(\mu)$.

3.1 Discussion

The results characterize outcomes of all stable matchings. A slight refinement of Theorem 3.1 will imply that the average value of one side of the market is essentially sufficient to determine the average value of the other side. Roughly, for a given stable matching μ , the value of λ in (4) and (5) is approximately the sum of the women's values in μ .¹² This suggests that the average value of men is approximately $1/\lambda \approx 1/\|\mathbf{Y}(\mu)\|_1$. Therefore multiplying the average values of the two sides of the market gives approximately $1/n$ simultaneously in all stable matchings with high probability. While we will establish such an approximation, we believe that, with a refined analysis, one should be able to show $\sup_{\mu \in \mathcal{S}} |n^{-1}\|\mathbf{X}(\mu)\|_1\|\mathbf{Y}(\mu)\|_1 - 1| \xrightarrow{P} 0$.

Moreover, the average value of men is also sufficient to predict the empirical value distribution on each side of the market. For example, if we find that 30% of the men have value h or higher, then we should expect 9% to have value $2h$ or higher.

Theorem 3.2 is similar but with respect to ranks; it implies that the product of the average scaled ranks of men and women should be asymptotically n , and the the average rank on each side determines the empirical rank distributions.

Observe that the scaling in (5) is consistent with the intuition of $\phi_i = \sum_{j=1}^n a_{ij}$ being the average fitness of m_i . Within a stable matching, a more popular man should, on average, achieve a better (smaller) rank than a less popular one. For instance, in a market with bounded public scores (Example 2.1), each man receives a number of proposals roughly inversely proportional to his fitness during the woman-proposing deferred acceptance algorithm, implying that his rank is proportional to ϕ_i in the woman optimal stable matching.

The proof of Theorem 3.1 also offers evidence that the number of stable matchings should essentially be sub-exponential. This is formally stated in Corollary 6.7.

3.2 Results for approximately stable matchings

The proof suggest that the characterization further extends to matchings that are only approximately stable in the following sense.

¹²Technically, the choice of λ can be taken as the sum of the values of women after excluding a small fraction δ of the women who are the least satisfied (those with the highest latent values) under the matching μ . This truncation, which is also done for technical reasons, avoid outliers and in fact shows that the predictions still hold under even weaker notions of stability. We believe that such trimming is unnecessary with a more careful analysis.

Definition 3.3. We say a matching μ between \mathcal{M} and \mathcal{W} is α -stable for some $0 < \alpha < 1$ if there exists a sub-market of size at least $(1 - \alpha)n$ on which μ is stable; that is, there exist subsets $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$ both with cardinality $|\mathcal{M}'| = |\mathcal{W}'| \geq (1 - \alpha)n$ such that $\mu(\mathcal{M}') = \mu(\mathcal{W}')$ and the partial matching induced by μ between \mathcal{M}' and \mathcal{W}' is stable (within this sub-market). We refer to the stable sub-matching between \mathcal{M}' and \mathcal{W}' as the stable part of μ . Denote the set of α -stable matchings by \mathcal{S}_α .

The following Theorem can be derived from the quantitative versions of Theorem 3.1 and 3.2, which will be presented in Section 7.

Theorem 3.4. Assume $\alpha < n^{-\eta}$ for some constant $\eta > 1/2$. Then, as $n \rightarrow \infty$,

$$\max_{\mu \in \mathcal{S}_\alpha} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\mathbf{X}(\mu)) - F_\lambda\|_\infty \xrightarrow{P} 0 \quad \text{and} \quad \max_{\mu \in \mathcal{S}_\alpha} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\phi^{-1} \circ \mathbf{R}(\mu)) - F_\lambda\|_\infty \xrightarrow{P} 0. \quad (6)$$

The approximately exponential empirical distribution applies to any matching that is stable except for $o(\sqrt{n})$ agents. The key observation is that, if the entire market satisfies the contiguity assumption, then each sub-market of it is also contiguous, and we can apply union bound over all sub-markets of size $(1 - \alpha)n$. As a corollary, we have the following result for slightly imbalanced markets.

Corollary 3.5. Consider a market consisting of $n - k$ men and n women, where $k < n^\beta$ for some constant $\beta < 1/2$. Assume that the contiguity condition holds as in Assumption 2.2. Then, as $n \rightarrow \infty$,

$$\max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\mathbf{X}(\mu)) - F_\lambda\|_\infty \xrightarrow{P} 0 \quad \text{and} \quad \max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\phi^{-1} \circ \mathbf{R}(\mu)) - F_\lambda\|_\infty \xrightarrow{P} 0. \quad (7)$$

Remark 3.6. The results will not hold if there is a linear imbalance in the market, i.e., $k \propto n$. This is because in such markets the men achieve a constant average rank Ashlagi et al. (2017), and therefore the convergence to the exponential distribution is impossible.

These results are not necessarily tight and one may weaken the constraints on α and k with a more careful analysis. Exploring other notions of approximate stability is left for future work.

4 Intuition and Proof Ideas

This section offers intuition and the key ideas behind the proofs.

4.1 Intuition

Let us start with providing a high-level intuition of both why the result is true, and how we should expect the proof to go. The actual proof does not follow the intuition exactly due to some technical difficulties that need to be overcome. It is possible that one can find a proof that follows the intuition below more directly.

At a very high level, the result follows from a *union bound*. There are $n!$ potential matchings. Based on *a-priori* preferences, for each matching μ , one could compute the probability P_μ that μ is stable under the realized preferences. A union bound argument just establishes that

$$\sum_{\substack{\mu \\ \mu \text{ does not satisfy the conditions of (4)}}} P_\mu = \exp(-\Omega_\epsilon(n)) \quad (8)$$

Establishing (8) directly appears to be difficult. A more approachable statement is a universal bound on the number of stable matchings overall:

$$\sum_{\mu} P_{\mu} = \exp(o(n)). \quad (9)$$

That is, in expectation, there are not so many stable matchings.

Consider the triplet of random variable (X, Y, μ) , where X and Y are preferences sampled according to the model, and μ is a uniformly random matching. Let \mathcal{S} be the event that μ is stable under preference profiles (X, Y) . If there are few stable matchings overall, as (9) implies, then we have

$$\mathbb{P}(\mathcal{S}) = \exp(o(1)) \cdot n!^{-1} \quad (10)$$

Another way of uniformly sampling from \mathcal{S} is as follows. First, sample $(X_1, Y, \mu) \in_U \mathcal{S}$. Then resample X_2 conditioned on $(X_2, Y, \mu) \in \mathcal{S}$. The triple (X_2, Y, μ) is a uniform element of \mathcal{S} . Note that for a fixed (Y, μ) the marginal distribution of X_2 conditioned on $(X_2, Y, \mu) \in \mathcal{S}$ is fairly simple to reason about: each member should prefer the pairing assigned to them by μ to all other potential blocking matches. In such a resampling, as we shall see, the empirical exponential distribution appears naturally from large deviations theory.

Suppose we prove that for all (Y, μ) ,

$$\mathbb{P}_{X_2:(X_2, Y, \mu) \in \mathcal{S}}(X_2 \text{ does not satisfy the conditions of (4)}) = \exp(-\Omega_{\epsilon}(n)) \quad (11)$$

Putting these together, we would get

$$\begin{aligned} & \mathbb{P}((X_2, Y, \mu) \in \mathcal{S} \wedge (X_2 \text{ does not satisfy the conditions of (4)})) \\ &= \mathbb{P}((X_1, Y, \mu) \in \mathcal{S}) \cdot \exp(-\Omega_{\epsilon}(n)) = n!^{-1} \cdot \exp(-\Omega_{\epsilon}(n)), \end{aligned}$$

implying (8) (together with a similar statement about Y).

A certain amount of technical work is needed to make the above blueprint go through. In particular, since our bounds need to be pretty tight, we need to worry about tail events. We end up having to perform the above resampling trick multiple times.

Why do we need the boundedness assumption 2.2? It is worth noting that while the boundedness assumption might not be the weakest assumption under which our results hold, some assumptions on the market are inevitable. In particular, if the market can be split into two independent balanced markets A and B , then there is no connection between the fortunes of men in market A and their fortunes in market B , and the empirical distribution of values on each side will be a mixture of two exponential distributions. Things will get even more complicated if markets A and B are not entirely independent, but are connected by a small number of agents. It is still possible that some version of Theorem 3.1 holds, but it will need to depend on the eigenspaces corresponding to large eigenvectors of the matrix.

It is worth noting that even (9) fails to hold when we do not have the boundedness assumption. Consider a market consisting of $n/2$ small markets with just 2 men and 2 women in each. Under the uniform preferences within each market, the expected number of stable matchings is $9/8$, thus

$$\sum_{\mu} P_{\mu} = (9/8)^{n/2} \neq \exp(o(n)).$$

4.2 Proof sketch

Case with uniformly random preferences. Let us first look at the classic case where agents' preferences are generated independently and uniformly at random, i.e., all canonical scores equal $1/n$. The proof in this case is more straightforward due to the symmetry among agents and the established high probability bound on the number of stable matchings (Pittel, 1989). We keep however the discussion informal in this section.

For a given matching μ , we study the conditional distribution of the value vectors $\mathbf{X}(\mu), \mathbf{Y}(\mu) \in \mathbb{R}^n$ conditional on stability of μ . Conditional also on women's value vector $\mathbf{Y}(\mu) = \mathbf{y}$, man m_i 's value $X_i(\mu)$ must satisfy $X_i(\mu) < X_{ij}$ for each $j \neq \mu(i)$ with $Y_{ji} < y_j$ in order not to form a blocking pair. Since X_{ij} and Y_{ji} are i.i.d. samples from $\text{Exp}(1)$ for each $j \neq \mu(i)$, one should expect $X_i(\mu)$ to be effectively less than the minimum of about $\sum_{j \neq \mu(i)} 1 - e^{-y_j} \approx \|\mathbf{y}\|_1$ number of $\text{Exp}(1)$ random variables. Such a constraint acts independently on each X_i (conditional on $\mathbf{Y}(\mu) = \mathbf{y}$), and therefore in the posterior distribution one should expect $\mathbf{X}(\mu)$ to behave like i.i.d. samples from $\text{Exp}(\|\mathbf{y}\|_1)$.

Concretely, conditional on $\mathbf{U}(\mu) = \mathbf{u} = 1 - F(\mathbf{x})$ and $\mathbf{V}(\mu) = \mathbf{v} = 1 - F(\mathbf{y})$, where we recall that $F(z) = 1 - e^{-z}$ is the CDF of $\text{Exp}(1)$ and is applied component-wise on the value vectors. The likelihood of (\mathbf{x}, \mathbf{y}) when μ is stable is $\prod_{j \neq \mu(i)} (1 - u_i v_j)$ (Pittel, 1989). With some crude first order approximation (namely, $1 - z \approx e^{-z}$), this expression can be approximated by

$$\prod_{j \neq \mu(i)} (1 - u_i v_j) \approx \exp\left(-\sum_{i,j \in [n]} x_i y_j\right) = \exp(-\|\mathbf{x}\|_1 \|\mathbf{y}\|_1), \quad (12)$$

where we also put in the terms $x_i y_{\mu(i)}$ for $i \in [n]$ despite their absence in the original product. By the Bayes rule, conditional on $\mathbf{Y}(\mu) = \mathbf{y}$ and that μ is stable, the distribution of $\mathbf{X}(\mu)$ is approximately $p(\mathbf{x} | \mu \in \mathcal{S}, \mathbf{Y} = \mathbf{y}) \propto \exp(-\|\mathbf{x}\|_1 \|\mathbf{y}\|_1) \cdot \prod_{i=1}^n e^{-x_i} = \prod_{i=1}^n \exp(-(1 + \|\mathbf{y}\|_1)x_i)$. Note that this is the joint density of the n -fold product of $\text{Exp}(1 + \|\mathbf{y}\|_1)$. Our main theorems in the case with uniformly random preferences follow directly from the convergence of empirical measures.

The general case. The entire result can be viewed abstractly from the following lens: For any matching μ , we expect the value vectors $(\mathbf{X}(\mu), \mathbf{Y}(\mu))$ to behave “nicely” with very high probability conditional on stability of μ , so that even if we apply union bound on all stable matchings (which we will show to be “rare” separately) it is still unlikely to see any “bad” value vectors. To do so requires a careful analysis on the conditional distribution of value (given stability) $\mathcal{D}_\mu \in \Delta(\mathbb{R}_+^n \times \mathbb{R}_+^n)$, which depends on both the (unconditional) preference distribution (the “prior”) and the conditional probability that μ is stable given a pair of value vector $\mathbf{X}(\mu) = \mathbf{x}$ and $\mathbf{Y}(\mu) = \mathbf{y}$, which we will denote by $p_\mu(\mathbf{x}, \mathbf{y})$. We will define the ultimate “nice” event to be the ϵ -proximity of empirical distribution of value (or rescaled rank) to some exponential distribution, but unsurprisingly it is hard to analyze this event directly from \mathcal{D}_μ , which is complicated itself, in one single step. Instead, we will follow a “layer-by-layer peeling” of the desirable events. Namely, we will find a nested sequence of subsets $\mathbb{R}_+^n \times \mathbb{R}_+^n = \Omega_0 \supseteq \Omega_1 \supseteq \dots \supseteq \Omega_K$ representing events on the joint value vectors of a stable matching, with Ω_K the desired event that the empirical distribution of value for men is ϵ -close to some exponential distribution. Step by step, we will show that a stable matching, conditional on its value vectors in Ω_i , must have value vectors in Ω_{i+1} with very high probability. Here is the roadmap to establishing these increasingly “nice” events:

- (a) As a first step, we approximate $p_\mu(\mathbf{x}, \mathbf{y})$, the likelihood of value vectors \mathbf{x} and \mathbf{y} in a stable matching μ , by the function $q(\mathbf{x}, \mathbf{y}) = \exp(-n\mathbf{x}^\top \mathbf{M}\mathbf{y})$. That is, the log likelihood of value

vectors in a stable matching is approximately bilinear in the two vectors. To establish this, we identify a weak regularity condition on the value vectors of all stable matchings in terms of the first and second moments and extremal quantiles of the value vectors, under which the approximation holds (see Section 6.1). Such a condition is met by all stable matchings with high probability (see Appendix A for details). The proof primarily consists of standard analysis of the deferred acceptance algorithm and careful use of first- and second-order approximation of $p_\mu(\mathbf{x}, \mathbf{y})$. Here we use the fact that the men-proposing and women-proposing deferred acceptance algorithms output the extremal outcomes with respect to the two sides' values among all possible stable matchings.

- (b) In the expression for $q(\mathbf{x}, \mathbf{y})$, the value vectors relate through the matching matrix \mathbf{M} . However, we show next that, in stable matchings, we can further simplify things by approximately factoring $n\mathbf{X}(\mu)^\top \mathbf{M}\mathbf{Y}(\mu)$ into a product $\|\mathbf{X}(\mu)\|_1 \|\mathbf{Y}(\mu)\|_1$ of sums of values on the two sides. More specifically, both $\mathbf{M}\mathbf{Y}(\mu)$ and $\mathbf{M}^\top \mathbf{X}(\mu)$ lie near the maximal eigenspace of \mathbf{M} , which is the span of $\mathbf{1}$ under Assumption 2.2 (see Section 6.4). The proof uses a fixed point argument to deduce that $\mathbf{M}\mathbf{Y}(\mu)$ depends almost deterministically on $\mathbf{M}^\top \mathbf{X}(\mu)$ and, symmetrically, $\mathbf{M}^\top \mathbf{X}(\mu)$ on $\mathbf{M}\mathbf{Y}(\mu)$, which forces both quantities to lie near the eigenspace. Along the way, we also deduce an upper bound for the (unconditional) probability for μ to be stable (see Section 6.3), suggesting an sub-exponential upper bound on the typical number of stable matchings (Corollary 6.7).
- (c) Under the previous event, the men's values behaves approximately like i.i.d. exponential samples with rate $\|\mathbf{Y}(\mu)\|_1$ conditional on stability of μ and $\mathbf{Y}(\mu)$ – in fact, they are conditionally independent and nearly identically distributed. The result on the empirical distribution of men's values follows immediately from a concentration inequality of Dvoretzky–Kiefer–Wolfowitz (DKW) type, generalized for nearly identically distributed independent random variables (Lemma B.4).
- (d) Finally, we translate values into ranks. Using the classic first- and second-moment method, we show that for majority of the agents in the market, the rescaled rank (based one's own scores) lies close to the value. This implies Theorem 3.2.

There is one caveat, however: In (a), a second order expansion of $p_\mu(\mathbf{x}, \mathbf{y})$ is required in order to justify the approximation with $q(\mathbf{x}, \mathbf{y})$. As a result, we need to control second order behavior of value, i.e., $\|\mathbf{X}(\mu)\|_2^2 = \sum_{i=1}^n X_i(\mu)^2$, in any stable matching μ . However, the second moment cannot be easily controlled due to the heavy tail of $\text{Exp}(1)$ (indeed, the moment generating function for X^2 does not exist for $X \sim \text{Exp}(1)$). To resolve this issue, we perform a truncation in the upper $\delta/2$ -quantile of the values on each side. By choosing δ sufficiently small, we can ensure that the truncation only affects the empirical distribution by an arbitrarily small amount in ℓ^∞ norm. As a price to pay, in (b) and (c), we will have to deal with not just all stable matchings, but all *partial* matchings on any $(1 - \delta)$ -fraction of the market that is stable. See Section 5.2 for the technical definition of truncated and partial matchings.

5 Preliminaries

5.1 Probability of stability and its approximation

For each matching μ , define the function $p_\mu : \mathbb{R}_+^n \times \mathbb{R}_+^n \rightarrow [0, 1]$ to be probability that μ is stable given values of men and women in μ . That is

$$p_\mu(\mathbf{x}, \mathbf{y}) = \mathbb{P}(\mu \in \mathcal{S} | \mathbf{X}(\mu) = \mathbf{x}, \mathbf{Y}(\mu) = \mathbf{y}). \quad (13)$$

Just like the integral formula used in Knuth (1976) and Pittel (1989, 1992) to study matching markets with uniformly random preferences, the probability of a matching μ being stable can be similarly characterized by an integral

$$\begin{aligned} \mathbb{P}(\mu \in \mathcal{S}) &= \mathbb{E}_{\mathbf{X} \sim \otimes_{i=1}^n \text{Exp}(a_{i,\mu(i)}), \mathbf{Y} \sim \prod_{i=1}^n \text{Exp}(b_{i,\mu^{-1}(i)})} [p_\mu(\mathbf{X}, \mathbf{Y})] \\ &= \int_{\mathbb{R}_+^n \times \mathbb{R}_+^n} p_\mu(\mathbf{x}, \mathbf{y}) \prod_{i=1}^n f_{a_{i,\mu(i)}}(x_i) f_{b_{i,\mu^{-1}(i)}}(y_i) d\mathbf{x} d\mathbf{y}. \end{aligned} \quad (14)$$

The function p_μ can be further expressed in closed form. Condition on the value vector $\mathbf{X}(\mu) = \mathbf{x}$ and $\mathbf{Y}(\mu) = \mathbf{y}$ and sample the rest of the values X_{ij} and Y_{ji} for all $j \neq \mu(i)$. Each pair of $i, j \in [n]$ with $j \neq \mu(i)$ may form a blocking pair when $X_{ij} < x_i$ and $Y_{ji} < y_j$, which event happens with probability $(1 - \exp(-a_{ij}x_i))(1 - \exp(-b_{ji}y_j))$. For μ to be stable, there must be no blocking pairs and thus

$$p_\mu(\mathbf{x}, \mathbf{y}) = \prod_{\substack{i,j \in [n] \\ j \neq \mu(i)}} \left(1 - (1 - e^{-a_{ij}x_i})(1 - e^{-b_{ji}y_j}) \right). \quad (15)$$

Under Assumption 2.2, i.e., $\max_{i,j_1,j_2} a_{ij_1}/a_{ij_2} \leq C^2$ and $\max_{j,i_1,i_2} b_{ji_1}/b_{ji_2} \leq C^2$, we observe a simple upper bound

$$p_\mu(\mathbf{x}, \mathbf{y}) \leq \prod_{\mu(i) \neq j} \left(1 - (1 - e^{-\hat{x}_i/C^2})(1 - e^{-\hat{y}_j/C^2}) \right), \quad (16)$$

where $\hat{x}_i = x_i a_{i,\mu(i)}$ and $\hat{y}_j = y_j b_{j,\mu^{-1}(j)}$ for $i, j \in [n]$ are the renormalized values (thus named because they have unit mean). This bound is fairly conservative and crude for our final purpose, but will prove useful for establishing preliminary results.

To further simplify the analysis, we recognize that, through first order approximation,

$$p_\mu(\mathbf{x}, \mathbf{y}) \approx \prod_{i \neq j} (1 - a_{ij} b_{ji} x_i y_j) \leq \exp\left(-\sum_{i \neq j} a_{ij} b_{ji} x_i y_j\right) \approx \exp(-n \mathbf{x}^\top \mathbf{M} \mathbf{y}). \quad (17)$$

Define the function

$$q(\mathbf{x}, \mathbf{y}) := \exp(-n \mathbf{x}^\top \mathbf{M} \mathbf{y}) = \exp\left(-n \sum_{i,j=1}^n m_{ij} x_i y_j\right).$$

In the next section we discuss conditions under which the function $q(\mathbf{x}, \mathbf{y})$ offers a good approximation for $p_\mu(\mathbf{x}, \mathbf{y})$.

5.2 Partial matchings and truncation

In order to study also approximately stable matchings, as well as for technical reasons, we need to consider matchings that are stable on a significant subset of the market. We first formalize a general partial matching then describe a particular way to form stable partial matchings.

Let $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$ be subsets of the men and women with cardinality $|\mathcal{M}'| = |\mathcal{W}'| = n'$. A *partial matching* $\mu' : \mathcal{M}' \rightarrow \mathcal{W}'$ is a bijection between \mathcal{M}' and \mathcal{W}' . Denote the values of men among \mathcal{M}' and women among \mathcal{W}' in the partial matching μ' by $\mathbf{X}_{\mathcal{M}'}(\mu')$ and $\mathbf{Y}_{\mathcal{W}'}(\mu')$, respectively. While it may be natural to view $\mathbf{X}_{\mathcal{M}'}(\mu')$ and $\mathbf{Y}_{\mathcal{W}'}(\mu')$ as n' -dimensional vector, we choose to view them as n -dimensional vectors where components corresponding to men in $\mathcal{M} \setminus \mathcal{M}'$ and women in $\mathcal{W} \setminus \mathcal{W}'$ are zero (recall that since small is better, zero is the best possible latent value). Therefore, conditional on $\mathbf{X}_{\mathcal{M}'}(\mu') = \mathbf{x}'$ and $\mathbf{Y}_{\mathcal{W}'}(\mu') = \mathbf{y}'$ for $\mathbf{x}', \mathbf{y}' \in \mathbb{R}^n$ supported on \mathcal{M}' and \mathcal{W}' , respectively, the probability that μ' is stable (as a matching between \mathcal{M}' and \mathcal{W}') is simply $p_{\mu'}(\mathbf{x}', \mathbf{y}')$.

Given a full stable matching μ and any $\delta > 0$, we define the following routine to construct a stable partial matching of size $n - \lfloor \delta n \rfloor$: Let $\bar{\mathcal{M}}_{\mu, \delta/2} \subseteq \mathcal{M}$ be the subset of $\lfloor \delta n/2 \rfloor$ men with the largest value (i.e., the least happy men) in μ , and similarly let $\bar{\mathcal{W}}_{\mu, \delta/2} \subseteq \mathcal{W}$ be the set of $\lfloor \delta n/2 \rfloor$ least happy women. Construct $\mathcal{M}'_{\mu, \delta} \subseteq \mathcal{M} \setminus (\bar{\mathcal{M}}_{\mu, \delta/2} \cup \mu(\bar{\mathcal{W}}_{\mu, \delta/2}))$ of cardinality $n - \lfloor \delta n \rfloor$. This is always possible because $|\bar{\mathcal{M}}_{\mu, \delta/2} \cup \mu(\bar{\mathcal{W}}_{\mu, \delta/2})| \leq 2\lfloor \delta n/2 \rfloor \leq \lfloor \delta n \rfloor$ and in fact here can be multiple ways to choose $\mathcal{M}'_{\mu, \delta}$. The specific way $\mathcal{M}'_{\mu, \delta}$ is chosen (when $|\mathcal{M} \setminus (\bar{\mathcal{M}}_{\mu, \delta/2} \cup \mu(\bar{\mathcal{W}}_{\mu, \delta/2}))| > n - \lfloor \delta n/2 \rfloor$) is irrelevant to our discussion, but it may be helpful to assume that the choice is made based on some canonical ordering of the men so there is no extra randomness. Let $\mu_\delta : \mathcal{M}'_{\mu, \delta} \rightarrow \mu(\mathcal{M}'_{\mu, \delta})$ be the partial matching induced by μ on $\mathcal{M}'_{\mu, \delta}$ and their partners. Define the δ -truncate value vector for μ to be $\mathbf{X}_\delta(\mu) := \mathbf{X}_{\mathcal{M}'_{\mu, \delta}}(\mu_\delta)$ and $\mathbf{Y}_\delta(\mu) := \mathbf{Y}_{\mu(\mathcal{M}'_{\mu, \delta})}(\mu_\delta)$.

6 Regularity of values in stable matchings

In this section, we establish several (high probability) properties of stable matchings.

6.1 Moment behavior and approximation of the conditional stable probability

We first consider a set of events in the value space, which can be thought of as regularity conditions for the approximation (17) of $p_\mu(\mathbf{x}, \mathbf{y})$ by $q(\mathbf{x}, \mathbf{y})$. Define

$$\mathcal{R}_1 = \{\mathbf{u} \in \mathbb{R}_+^n : \|\mathbf{u}\|_1 \geq \underline{c}_1 \log n\}, \quad (18)$$

$$\bar{\mathcal{R}}_1 = \{\mathbf{u} \in \mathbb{R}_+^n : \|\mathbf{u}\|_1 \leq \bar{c}_1 n (\log n)^{-7/8}\}, \quad (19)$$

$$\mathcal{R}_2 = \{(\mathbf{u}, \mathbf{v}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \mathbf{u}^\top \mathbf{M} \mathbf{v} \leq c_2 (\log n)^{1/8}\}, \quad (20)$$

where $\underline{c}_1, \bar{c}_1, c_2 \in \mathbb{R}_+$ are constants to be specified later. Let

$$\mathcal{R}_1 = \underline{\mathcal{R}}_1 \cap \bar{\mathcal{R}}_1 \quad \text{and} \quad \mathcal{R} = (\mathcal{R}_1 \times \mathbb{R}_+^n) \cap (\mathbb{R}_+^n \times \mathcal{R}_1) \cap \mathcal{R}_2 = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{R}_2 : \mathbf{x}, \mathbf{y} \in \mathcal{R}_1\}.$$

The region \mathcal{R} should capture the typical behavior of $(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu))$ for any stable matching μ .

Proposition 6.1. *For any fixed $\delta > 0$ and $c \in (0, 1/2)$, the constants $\underline{c}_1, \bar{c}_1$, and c_2 in (18)-(20) can be appropriately chosen such that*

$$\mathbb{P}(\exists \mu \in \mathcal{S}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \mathcal{R}) \lesssim e^{-n^c} \quad (21)$$

asymptotically as $n \rightarrow \infty$.

Remark 6.2. *It is helpful to compare the bounds (18)-(20) to classic results in the setting with uniform preferences (cf. (Pittel, 1989, 1992)): Namely, the optimal average rank is $\Theta(\log n)$, the pessimal average rank is $\Theta(n/\log n)$, and the product of the average ranks on the two side is asymptotic to n in all stable matchings. Here, due to heterogeneity of preferences, we pay a small price of an extra constant or $(\log n)^{1/8}$ factor.*

We will defer the proof to Appendix A. In fact, we will establish even finer control over the truncated value vectors in stable matchings. For a matching μ , we define $\mathbf{U}(\mu) = F(\hat{\mathbf{X}}(\mu))$ and $\mathbf{V}(\mu) = F(\hat{\mathbf{Y}}(\mu))$, where the (standard exponential CDF) function $F(z) = 1 - e^{-z}$ is applied coordinate-wise to the renormalized value vectors. Through relating \mathbf{X} and \mathbf{Y} to \mathbf{U} and \mathbf{V} , we will specify a subregion $\mathcal{R}^* \subseteq \mathcal{R}$ in which $p_\mu(\mathbf{x}, \mathbf{y})$ can be well approximated by $q(\mathbf{x}, \mathbf{y})$.¹³ We will see in Corollary A.13 that with high probability no stable matchings μ have $(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu))$ outside \mathcal{R}^* , from which Proposition 6.1 follows.

The conditions for \mathcal{R}^* are sufficiently strong to bound the functions p_μ and q within an $\exp(o(n))$ factor of each other. This is formalized as follows.

Proposition 6.3. *For any $\delta > 0$ and $c \in (0, 1/2)$, there exists an absolute constant $\theta \in (0, \infty)$ such that the probability that a matching μ is stable with value vectors not satisfying*

$$\frac{p_\mu(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu))}{q(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu))} \leq \exp\left(\frac{\theta n}{(\log n)^{1/2}}\right) \quad (\star)$$

is $\frac{\exp(-n^c)}{n!}$. In other words, with high probability, there exist no stable matchings μ whose post-truncation value vectors $\mathbf{X}_\delta(\mu)$ and $\mathbf{Y}_\delta(\mu)$ satisfy (\star) .

Again, the proof of Proposition 6.3 is deferred to Appendix A.

The reason for using δ -truncated value vectors is that, when approximating $p(\mathbf{x}, \mathbf{y})$ to second order, there will be terms of $\|\mathbf{x}\|_2^2$ and $\|\mathbf{y}\|_2^2$, which are hard to control due to the heavy tail of the exponential distribution.¹⁴ On the other hand, changing the values of a δ fraction should affect the empirical CDF by at most δ in ℓ^∞ distance. Therefore, it suffices to show that for small enough δ all stable partial matchings of size $n - \lfloor \delta n \rfloor$ have values and ranks empirically distributed close to some exponential distribution.

The function $p_\mu(\mathbf{x}, \mathbf{y})$ cannot be approximated globally by $q(\mathbf{x}, \mathbf{y}) = \exp(-n\mathbf{x}^\top \mathbf{M}\mathbf{y})$. However, we can find a region in $\mathbb{R}_+^n \times \mathbb{R}_+^n$ where $p_\mu(\mathbf{x}, \mathbf{y})$ and $q(\mathbf{x}, \mathbf{y})$ are close (uniformly for all stable matchings) in the sense that (\star) holds; Meanwhile, Proposition 6.3 states that with high probability, no stable matchings will ever have δ -truncated value vectors outside this region.

6.2 A key reduction lemma

Proposition 6.3 allows us to study high probability behaviors in stable partial matchings obtained from truncating stable (full) matchings pretending that the conditional probability of stability were given by $q(\mathbf{x}, \mathbf{y})$. Concretely, consider a fixed constant $\delta > 0$ and a region $\Omega \subseteq \mathbb{R}_+^n \times \mathbb{R}_+^n$ that defines an event on the (truncated) value vectors. If there exists a stable matching μ whose truncated value vectors $(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \in \Omega$, the induced partial matching μ_δ of size $n - \lfloor \delta n \rfloor$ between $\mathcal{M}'_{\mu, \delta}$ and $\mu(\mathcal{M}'_{\mu, \delta})$ must also be stable with value vectors $\mathbf{X}_{\mathcal{M}'_{\mu, \delta}}(\mu_\delta) = \mathbf{X}_\delta(\mu)$ and $\mathbf{Y}_{\mu(\mathcal{M}'_{\mu, \delta})}(\mu_\delta) = \mathbf{Y}_\delta(\mu)$. Thus, we will end up either having a stable matching whose truncated value violates (\star) , or a

¹³Technically, \mathcal{R}^* has to be defined in the context of a matching μ , as \mathbf{U} and \mathbf{V} . Here we drop the dependency for convenience. See Corollary A.13 for the formal definition of $\mathcal{R}^*(\mu)$.

¹⁴Note that moment generating function does not exist for X^2 where $X \sim \text{Exp}(1)$ so the classic Hoeffding- or Bernstein-type bounds fail to apply.

stable partial matching of size $n - \lfloor \delta n \rfloor$ whose value vectors (already truncated) lies in $\Omega \cap \mathcal{R}^*$. By Proposition 6.3, the former event happens with probability $o(1)$. Therefore, we may focus on the second event, where a stable partial matching of size $n - \lfloor \delta n \rfloor$ exists with value vectors in $\Omega \cap \mathcal{R}^*$. This is summarized by the following Lemma, which will be a major tool in the remainder of the proof.

Lemma 6.4. *Let $\delta > 0$, $c \in (0, 1/2)$, and $\Omega \subseteq \mathbb{R}_+^n \times \mathbb{R}_+^n$. Then,*

$$\mathbb{P}(\exists \mu \text{ stable}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \in \Omega) \leq e^{-n^c} + \exp\left(\Theta\left(\frac{n}{(\log n)^{1/2}}\right)\right) \cdot \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M}, \mathcal{W}' \subseteq \mathcal{W} \\ |\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor}} \sum_{\substack{\mu': \mathcal{M}' \rightarrow \mathcal{W}' \\ \text{bijection}}} \mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega}(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')))]. \quad (22)$$

Proof. Note that

$$\mathbb{P}(\exists \mu \text{ stable}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \in \Omega) \leq \mathbb{P}(\exists \mu \text{ stable}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \mathcal{R}^*) + \mathbb{P}(\exists \mu \text{ stable}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \in \mathcal{R}^* \cap \Omega), \quad (23)$$

where the first term is e^{-n^c} by Corollary A.13. Let \mathcal{E} denote the event that there exists a *stable* partial matching μ' between $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$ with $|\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor$ where $(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \in \mathcal{R}^* \cap \Omega$. Clearly, the existence of a stable matching μ with $(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \in \mathcal{R}^* \cap \Omega$ implies \mathcal{E} . Thus, the second term in (23) is bounded by

$$\mathbb{P}(\mathcal{E}) \leq \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M}, \mathcal{W}' \subseteq \mathcal{W} \\ |\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor}} \sum_{\substack{\mu': \mathcal{M}' \rightarrow \mathcal{W}' \\ \text{bijection}}} \mathbb{P}(\mu' \text{ stable}, (\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \in \mathcal{R}^* \cap \Omega). \quad (24)$$

using union bound. For each \mathcal{M}' , \mathcal{W}' , and μ' in the summation, we compute the above probability through conditioning on $\mathbf{X}_{\mathcal{M}'}(\mu')$ and $\mathbf{Y}_{\mathcal{W}'}(\mu')$ as

$$\begin{aligned} & \mathbb{P}(\mu' \text{ stable}, (\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \in \mathcal{R}^* \cap \Omega) \\ &= \mathbb{E}[\mathbb{P}(\mu' \text{ stable} | \mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')); (\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \in \mathcal{R}^* \cap \Omega] \\ &= \mathbb{E}[p_{\mu'}(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \cdot \mathbb{1}_{\mathcal{R}^* \cap \Omega}(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')))] \\ &\leq \mathbb{E}\left[\exp\left(\Theta\left(\frac{n}{(\log n)^{1/2}}\right)\right) q(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \cdot \mathbb{1}_{\mathcal{R}^* \cap \Omega}(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu'))\right)] \\ &\leq \exp\left(\Theta\left(\frac{n}{(\log n)^{1/2}}\right)\right) \mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega}(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')))], \quad (25) \end{aligned}$$

which completes the proof. \square

Remark 6.5. *The choice of the constant $c \in (0, 1/2)$ affects the implicit constants in defining \mathcal{R} and \mathcal{R}^* . Once we have a target convergence of e^{-n^c} for some $c \in (0, 1/2)$, we will assume that c is fixed in the rest of our discussion unless otherwise mentioned.*

Lemma 6.4 will be a key tool in the proof to establish further likely behaviors of value (and rank) vectors. It will be a recurring theme where we first identify a likely region Ω_{likely} for truncated value vectors of stable (full) matchings to fall in (\mathcal{R} to start with), then rule out a bad event Ω_{bad} within Ω_{likely} by showing $\Omega_{\text{likely}} \cap \Omega_{\text{bad}}$ is unlikely for value vectors of any stable *partial* matching,

and apply Lemma 6.4 to conclude that Ω_{bad} is unlikely for truncated value vectors of any stable matching and that $\Omega_{\text{likely}} \cap \Omega_{\text{bad}}$ can be used as the likely region moving forward.

Based on Lemma 6.4, it now suffices to consider partial matchings μ' of size $n - \lfloor \delta n \rfloor$ between \mathcal{M}' and \mathcal{W}' and upper bound $\mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{X}_{\mathcal{W}'}(\mu')) \cdot \mathbb{1}_{\Omega}(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{X}_{\mathcal{W}'}(\mu'))]$. From now on, we fix \mathcal{M}' , \mathcal{W}' , and μ' , and make the dependency of $\mathbf{X}_{\mathcal{M}'}$ and $\mathbf{Y}_{\mathcal{W}'}$ on μ' implicit when the context is clear.

6.3 Estimating the (unconditional) stability probability

Using concentration inequalities given in Lemma B.1 and B.2, we derive the following upper bound, which essentially characterizes the (approximate) probability that a partial matching of size $n - \lfloor \delta n \rfloor$ is stable with a probable value vector for the women (i.e., $\mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1$).

Proposition 6.6. *For a fixed a partial matching μ' on \mathcal{M}' and \mathcal{W}' of size $n - \lfloor \delta n \rfloor$,*

$$\mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_1}(\mathbf{Y}_{\mathcal{W}'})] \leq e^{o(n)+o_\delta(n)} \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}. \quad (26)$$

The proof of Proposition 6.6 will be deferred to Appendix C, where we will develop intermediate results that characterize the typical behavior of $\mathbf{X}_{\mathcal{M}'}$ and $\mathbf{Y}_{\mathcal{W}'}$ relative to each other (see Appendix C.2).

Proposition 6.6 provides evidence that, heuristically, the expected number of stable partial matchings should be sub-exponential.

Corollary 6.7 (Number of stable partial matchings). *Fix any $\delta > 0$ and $c \in (0, 1/2)$. Let N_δ denote the number of stable partial matchings of size $n - \lfloor \delta n \rfloor$ satisfying the condition in Corollary A.13 (i.e., \mathcal{R}^*) in a random instance of the matching market. Then, $\mathbb{E}[N_\delta] \leq \exp(o_\delta(n))$ granted that n is sufficiently large. Further, with probability at least $1 - e^{-n^c}$, the condition \mathcal{R}^* is satisfied by all δ -truncated stable matchings.*

Remark 6.8. *Corollary 6.7 falls short of establishing a sub-exponential bound for the expected number of stable matchings in two aspects.*

- *While stable matchings that violate \mathcal{R}^* (when truncated) will not exist with high probability, we have not yet proved a bound for the expected number of such stable matchings. We believe that this can be overcome with a refined analysis of deferred acceptance, which should lead to stronger results than Lemma A.1. Note that all high probability results in Appendix A after this lemma come with an upper bound on the expected number of stable matchings under various conditions.*
- *In general, it is possible to have multiple, in the worst case $\lfloor \delta n \rfloor!$, stable matchings that produces the same δ -truncated stable partial matching.*

We believe that a sub-exponential bound for the number of stable matchings is possible with a more refined analysis.

6.4 Opportunity sets and an eigenspace property for the value vectors

Our next result states that value vectors in stable matchings are not only controlled in terms of their first and second moments, but also in a sense “close” to some constant vector, i.e., $t\mathbf{1}$ for some $t \in \mathbb{R}_+$, which are eigenvectors of \mathbf{M} corresponding to its maximal eigenvalue $\lambda_1(\mathbf{M}) = 1$.

Let us fix the women's values to be $\mathbf{Y}_{\mathcal{W}'} = \mathbf{y}$ and consider the implication for the men's outcome in any (partial) matching μ' . For a man m_i with value x_i , the expected number of blocking pairs between him and the women, conditional on x_i and \mathbf{y} , is

$$\sum_{j \neq \mu'(i)} (1 - e^{-a_{ij}x_i})(1 - e^{-b_{ji}y_j}) \approx \sum_{j=1}^n a_{ij}b_{ji}x_iy_j = n(\mathbf{M}\mathbf{y})_ix_i.$$

The next result suggests that, in a typical market, the burden of avoiding blocking pairs falls roughly equally on the men in the sense that the entries of $\mathbf{M}\mathbf{Y}_{\mathcal{W}'}$ are largely the same.

Lemma 6.9. *Let μ' be a partial matching of size $n - \lfloor \delta n \rfloor$ on $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$. Fix any $\zeta > 0$, and let*

$$\Omega_{\text{eig}}(\zeta) := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n : \exists t \in \mathbb{R}_+, \sum_{i=1}^n \mathbb{1} \left\{ |(\mathbf{M}\mathbf{y})_i - t| \geq \sqrt{\zeta}t \right\} \leq \sqrt{\zeta}n \right\}. \quad (27)$$

Then

$$\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \setminus \Omega_{\text{eig}}(\Theta(\delta) + \zeta)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \right] \leq \exp(o_\delta(n) - \Theta(\zeta^2 n)) \cdot \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}, \quad (28)$$

with the implicit constants uniform over all \mathcal{M}' , \mathcal{W}' , and μ' .

The proof of Lemma 6.9 is deferred to Appendix C.2.

Let us observe the immediate corollary of this Lemma, the proof of which is similar to that of Lemma 6.4 and deferred to Appendix C.3.

Corollary 6.10. *For $\delta > 0$ sufficiently small, there exists a choice of $\zeta = \zeta(\delta) > 0$ such that $\zeta \rightarrow 0$ as $\delta \rightarrow 0$ and that*

$$\mathbb{P}(\exists \mu \text{ stable}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \Omega_{\text{eig}}(\zeta)) \lesssim e^{-n^c} \quad (29)$$

asymptotically as $n \rightarrow \infty$.

Corollary 6.10 roughly states that, in a contiguous market, conditioning on the women's outcomes in a stable matching has an almost even impact on the men's values.

7 Empirical distribution of values and ranks

7.1 Empirical distribution of values

Knowing the eigenspace property of the value vectors allows us to characterize the empirical distribution of values.

Lemma 7.1. *Fix any $\delta, \zeta > 0$. Let μ' be a partial matching of size $n - \lfloor \delta n \rfloor$ on $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$. For any $\epsilon > 0$, consider*

$$\Omega_{\text{emp}}(\epsilon) := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \exists \lambda \in \mathbb{R}_+, \|\hat{\mathcal{F}}(\mathbf{x}) - F_\lambda\|_\infty \leq \epsilon + \Theta(\delta + \sqrt{\zeta}) \right\}. \quad (30)$$

Then

$$\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega_{\text{eig}}(\zeta) \setminus \Omega_{\text{emp}}(\epsilon)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \right] \leq \exp(o_\delta(n) - \Theta(\epsilon^2 n)) \cdot \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}, \quad (31)$$

where again the implicit constants are uniform over all \mathcal{M}' , \mathcal{W}' , and μ' .

The proof formalizes the intuition that, conditional on stability of μ' and $\mathbf{Y}_{\mathcal{W}'} = \mathbf{y}$, the value X_i for $i \in \mathcal{M}'$ should behave approximately as $\text{Exp}(\lambda_i)$ for some $\lambda_i = (1 + \Theta(\delta + \sqrt{\zeta}))\|\mathbf{y}\|_1$. The full proof is deferred to Appendix C.5. Hence, instead of looking for the optimal λ that minimizes $\|\hat{\mathcal{F}}(\mathbf{X}_{\mathcal{M}'} - \mathcal{F}(\text{Exp}(\lambda)))\|_\infty$ in the definition (30) of $\Omega_{\text{emp}}(\epsilon)$, we may simply choose $\lambda = \|\mathbf{y}\|_1$, which only differs from the right choice by at most a tolerable $\Theta(\sqrt{\zeta} + \delta)$ factor. In other words, if we define

$$\tilde{\Omega}_{\text{emp}}(\epsilon) := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \|\hat{\mathcal{F}}(\mathbf{x}) - \mathcal{F}(\text{Exp}(\|\mathbf{y}\|_1))\|_\infty \leq \epsilon + \Theta(\delta + \sqrt{\zeta}) \right\},$$

albeit with a worse implicit constant in $\Theta(\delta + \sqrt{\zeta})$, the same conclusion holds as in Lemma 7.1 with $\Omega_{\text{emp}}(\epsilon)$ replaced by $\tilde{\Omega}_{\text{emp}}(\epsilon)$; that is,

$$\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega_{\text{eig}}(\zeta) \setminus \tilde{\Omega}_{\text{emp}}(\epsilon)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \right] \leq \exp(o_\delta(n) - \Theta(\epsilon^2 n)) \cdot \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}. \quad (32)$$

Using Lemma 7.1, we now prove our first main theorem about the uniform limit of empirical distribution of men's (or women's) value in stable matchings.

Theorem 7.2 (Empirical distribution of value). *Fix any $\epsilon > 0$. Then*

$$\mathbb{P} \left(\max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\mathbf{X}(\mu)) - F_\lambda\|_\infty > \epsilon \right) \lesssim e^{-n^c} \quad (33)$$

asymptotically as $n \rightarrow \infty$. In particular, the infimum over λ in (33) can be replaced with the choice of $\lambda = \|\mathbf{Y}_\delta(\mu)\|_1$ for δ sufficiently small.

Proof. Plugging (32) into Lemma 6.4 and repeating the same arithmetic as in (111) and (112) immediately give

$$\mathbb{P}(\exists \mu \in \mathcal{S}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \in \Omega_{\text{eig}}(\zeta) \setminus \tilde{\Omega}_{\text{emp}}(\epsilon)) \leq e^{-n^c} + \exp(o_\delta(n) - \Theta(\epsilon^2 n)) \lesssim e^{-n^c}, \quad (34)$$

granted that $\epsilon \geq \epsilon_0(\delta)$, where the function $\epsilon_0(\delta) \rightarrow 0$ as $\delta \rightarrow 0$. Corollary 6.10 implies that with probability at least $1 - \Theta(e^{-n^c})$ there exists no stable matching μ with $(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \Omega_{\text{eig}}(\zeta)$, and hence

$$\mathbb{P}(\exists \mu \in \mathcal{S}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \tilde{\Omega}_{\text{emp}}(\epsilon/3)) \lesssim e^{-n^c}, \quad (35)$$

granted that $\epsilon_0(\delta) < \epsilon/3$. By choosing δ (and hence also $\zeta = \zeta(\delta)$) sufficiently small so that the $\Theta(\delta + \sqrt{\zeta})$ term in the definition of $\tilde{\Omega}_{\text{emp}}$ is upper bounded by $\epsilon/3$, we ensure

$$\mathbb{P}(\exists \mu \in \mathcal{S}, \|\hat{\mathcal{F}}(\mathbf{X}_\delta(\mu)) - \mathcal{F}(\text{Exp}(\|\mathbf{Y}_\delta(\mu)\|_1))\|_\infty \geq 2\epsilon/3) \lesssim e^{-n^c}. \quad (36)$$

By further restricting δ to be sufficiently small, we may absorb the difference caused by the δ -truncation on $\mathbf{X}(\mu)$ into an extra term of $\Theta(\delta) \leq \epsilon/3$, since $\|\hat{\mathcal{F}}(\mathbf{X}_\delta(\mu)) - \hat{\mathcal{F}}(\mathbf{X}(\mu))\|_\infty \leq \delta$. The theorem follows immediately. \square

With essentially the same analysis as in Lemma 7.1 and Theorem 7.2, except for replacing the DKW inequality with Bernstein's inequality for empirical average, we can also deduce the following result. The proof is omitted.

Proposition 7.3. *For any fixed $\epsilon, \delta > 0$ and $0 < c < 1/2$,*

$$\mathbb{P} \left(\max_{\mu \in \mathcal{S}} |n^{-1} \|\mathbf{X}_\delta(\mu)\|_1 \|\mathbf{Y}_\delta(\mu)\|_1 - 1| > \epsilon \right) \lesssim e^{-n^c}. \quad (37)$$

The effect of the δ -truncation is nontrivial to remove because the sum of values can be sensitive to outliers, in particular given the heavy tail of the exponential distribution. We believe, however, that a refined analysis should suggest that $\sup_{\mu \in \mathcal{S}} |n^{-1} \|\mathbf{X}(\mu)\|_1 \|\mathbf{Y}(\mu)\|_1 - 1| \xrightarrow{P} 0$. This is the analogue of the ‘‘law of hyperbola’’ in Pittel (1992).

7.2 Empirical distribution of ranks

Based on the previous discussion on the empirical distribution of value, we now extend the result to ranks and prove our second main theorem.

Theorem 7.4 (Empirical distribution of ranks). *For any fixed $\epsilon > 0$,*

$$\mathbb{P}\left(\max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\phi^{-1} \circ \mathbf{R}(\mu)) - F_\lambda\|_\infty > \epsilon\right) \lesssim e^{-n^\epsilon} \quad (38)$$

asymptotically as $n \rightarrow \infty$, where ϕ is men's fitness vector. As in Theorem 7.2, the infimum over λ in (33) can be replaced with the choice of $\lambda = \|\mathbf{Y}_\delta(\mu)\|_1$ for δ sufficiently small.

Heuristically, we would expect the rank $R_i(\mu)$ for a man to be proportional to his value $X_i(\mu)$ when stability of μ and the values $X_i(\mu) = x_i$ and $\mathbf{Y}(\mu) = \mathbf{y}$ are conditioned upon. Indeed, a woman w_j with $j \neq \mu(i)$ stands ahead of $w_{\mu(i)}$ in the preference of man m_i exactly when $X_{ij} < x_i$ and $Y_{ji} > y_{\mu(i)}$. As X_{ij} and Y_{ji} jointly follow the product distribution $\text{Exp}(a_{ij}) \otimes \text{Exp}(b_{ji})$ conditional on the event that $X_{ij} < x_i$ and $Y_{ji} < y_{\mu(i)}$ do not simultaneously happen, the conditional probability that $w_j \succeq_{m_i} w_{\mu(i)}$ is $\frac{(1-e^{-a_{ij}x_i})e^{-b_{ji}y_j}}{1-(1-e^{-a_{ij}x_i})(1-e^{-b_{ji}y_j})} \approx a_{ij}x_i$. By summing over $j \neq \mu(i)$, we should expect that the rank $R_i(\mu)$ to be in expectation close to $x_i \sum_{j \neq i} a_{ij} \approx x_i \phi_i$; further, as the sum of independent Bernoulli random variables, $R_i(\mu)$ should concentrate at its expectation, therefore leading to $R_i(\mu) \approx x_i \phi_i$ simultaneously for most $i \in [n]$. This intuition is formalized in the proof, which is given in Appendix C.6.

8 Conclusion

We studied the welfare structure in stable outcomes in large two-sided matching markets with logit-based preferences. Under a contiguity condition that prevents agents from disproportionately favoring or unfavoring other agents, we characterize stable and almost stable matchings outcomes in terms of the empirical distribution of latent values and ranks.

In particular, our results suggest that the welfare of an agent in a stable matching can be decomposed into three parts: a global parameter that determines the trade-off between the two sides, a personal intrinsic fitness computed from the systematic scores, and an exogenous factor behaving as a standard exponential random variable. In other words, given the market structure (i.e., the systematic scores), the average rank (or value) of the men (or women) is essentially a sufficient statistic for the outcome distribution.

References

- Nikhil Agarwal and Paulo Somaini. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica*, 86(2):391–444, 2018.
- Itai Ashlagi, Mark Braverman, and Avinatan Hassidim. Stability in large matching markets with complementarities. *Operations Research*, 62(4):713–732, 2014.
- Itai Ashlagi, Yash Kanoria, and Jacob D Leshno. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy*, 125(1):69–98, 2017.
- Itai Ashlagi, Mark Braverman, Amin Saberi, Clayton Thomas, and Geng Zhao. Tiered random matching markets: Rank is proportional to popularity. In *12th Innovations in Theoretical Computer Science Conference (ITCS)*, 2021.

- Linda Cai and Clayton Thomas. The short-side advantage in random matching markets. In *SIAM Symposium on Simplicity in Algorithms (SOSA)*, 2022.
- Benjamin Doerr. Probabilistic tools for the analysis of randomized optimization heuristics. In *Theory of evolutionary computation*, pages 1–87. Springer, 2020.
- Aryeh Dvoretzky, Jack Kiefer, and Jacob Wolfowitz. Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator. *The Annals of Mathematical Statistics*, pages 642–669, 1956.
- William Feller. *An introduction to probability theory and its applications, Volume 2*, volume 81. John Wiley & Sons, 1971.
- David Gale and Lloyd S Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15, 1962.
- Hugo Gimbert, Claire Mathieu, and Simon Mauras. Two-sided matching markets with correlated random preferences have few stable pairs. *arXiv preprint arXiv:1904.03890*, 2019.
- Günter J Hitsch, Ali Hortaçsu, and Dan Ariely. Matching and sorting in online dating. *American Economic Review*, 100(1):130–163, 2010.
- Nicole Immorlica and Mohammad Mahdian. Incentives in large random two-sided markets. *ACM Transactions on Economics and Computation (TEAC)*, 3(3):1–25, 2015.
- Donald E Knuth, Rajeev Motwani, and Boris Pittel. Stable husbands. *Random Structures & Algorithms*, 1(1):1–14, 1990.
- Donald Ervin Knuth. *Mariages stables et leurs relations avec d’autres problèmes combinatoires: introduction à l’analyse mathématique des algorithmes*. Les Presses de l’Université de Montréal, 1976.
- Donald Ervin Knuth. *Stable marriage and its relation to other combinatorial problems: An introduction to the mathematical analysis of algorithms*, volume 10. American Mathematical Soc., 1997.
- Fuhito Kojima and Parag A Pathak. Incentives and stability in large two-sided matching markets. *American Economic Review*, 99(3):608–27, 2009.
- Simon Mauras. Two-sided random matching markets: Ex-ante equivalence of the deferred acceptance procedures. *ACM Transactions on Economics and Computation*, 9(4):1–14, 2021.
- Peter McCullagh. An asymptotic approximation for the permanent of a doubly stochastic matrix. *Journal of Statistical Computation and Simulation*, 84(2):404–414, 2014.
- Konrad Menzel. Large matching markets as two-sided demand systems. *Econometrica*, 83(3):897–941, 2015.
- VB Nevzorov. Representations of order statistics, based on exponential variables with different scaling parameters. *Journal of Soviet Mathematics*, 33(1):797–798, 1986.
- Boris Pittel. The average number of stable matchings. *SIAM Journal on Discrete Mathematics*, 2(4):530–549, 1989.

- Boris Pittel. On likely solutions of a stable marriage problem. *The Annals of Applied Probability*, 2(2):358–401, 1992.
- Boris Pittel. On likely solutions of the stable matching problem with unequal numbers of men and women. *Mathematics of Operations Research*, 44(1):122–146, 2019.
- Alvin E Roth. Marketplaces, markets, and market design. *American Economic Review*, 108(7):1609–58, 2018.
- Alvin E Roth and Marilda Sotomayor. Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541, 1992.
- Pál Schweitzer. *Egy egyenlőtlenség az aritmetikai középértékről (Inequality containing the arithmetic mean)*. 1914.
- Richard Sinkhorn. A relationship between arbitrary positive matrices and doubly stochastic matrices. *The annals of mathematical statistics*, 35(2):876–879, 1964.

A Proofs of typical behaviors of scores in stable matching

Recall that for a matching μ with (latent) value vectors $\mathbf{X}(\mu)$ and $\mathbf{Y}(\mu)$, we define $\mathbf{U}(\mu) = F(\hat{\mathbf{X}}(\mu))$ and $\mathbf{V}(\mu) = f(\hat{\mathbf{X}}(\mu))$ with the standard exponential CDF $F(z) = 1 - e^{-z}$ applied component-wise to the renormalized values $\hat{X}_i(\mu) = X_i(\mu)a_{i,\mu(i)}$ and $\hat{Y}_i(\mu) = Y_i(\mu)b_{i,\mu^{-1}(i)}$ for $i \in [n]$, so that $U_i, V_i \sim \text{Unif}([0, 1])$ and are mutually independent due to the way the score matrices are generated.

Lemma A.1. *For any $c \in (0, 1/2)$, there exists a constant $\theta_1 > 0$ (depending on c and C) such that in a random instance of the market, at least $\theta_1 n \ln n$ proposals are made during man-proposing deferred acceptance with probability $1 - \exp(-n^c)$.*

Proof. This result follows from a standard analysis of the deferred acceptance algorithm executed as in (Ashlagi et al., 2021, Section 3). In Ashlagi et al. (2021), the preference model involves tiers, where fitness values among different tiers differ by at most a constant factor. It turns out that this bounded ratio of fitness is the only thing used in the proofs, and is also satisfied by our matching market under Assumption 2.2. The main steps of analysis are as follows.

1. Consider (man-proposing) deferred acceptance *with re-proposals*, where each time when man i proposes, his proposal will go to woman i with probability proportional to a_{ij} , independent of all previous proposals (and their acceptance/rejection). The total number T of proposals in this process is equal in distribution to the number of draws in the coupon collector problem, and from standard concentration bounds we can show $\mathbb{P}(T \geq kn^{1+c}) \leq e^{-n^{c-2}}$ for k sufficiently large and $\mathbb{P}(T < \alpha n \ln n) \leq \exp(-n^c - 2)$ for $\alpha > 0$ sufficiently small (see also Doerr (2020)). The details mirror Appendices A and B in Ashlagi et al. (2021).
2. Analogous to Lemmas 3.5 and 3.6 in Ashlagi et al. (2021), we can show that, with probability $1 - \exp(-n^c - 1)$, no single man makes more than ℓn^{2c} proposals during deferred acceptance with re-proposal for ℓ sufficiently large.
3. Conditional on $T \geq \alpha n \ln n$ and the maximum number of proposals any man makes being at most ℓn^{2c} , the fraction of re-proposals should be no greater than $C \ell n^{2c-1}$ in expectation, since each proposal will be a duplicate of a previous proposal independently with probability

at most $\frac{C\ell n^{2n}}{n}$. It follows immediately from a binomial concentration that the (conditional) probability that the number of repeated proposals exceeds $T/2$ is exponentially small. Hence, the actual number of proposals during deferred acceptance (without re-proposals) is at least $\frac{\alpha}{2}n \ln n$ with probability $1 - \exp(-n^c)$.

□

Lemma A.2. *For any $c \in (0, 1/2)$, there exists a constant $\theta_1 > 0$ (depending on c and C) such that, with probability $1 - \exp(-n^c)$, there exist no stable matchings μ where $\|\mathbf{U}(\mu)\|_1 \leq \theta_2 \ln n$.*

Proof. Note that $U_i(\mu) = 1 - e^{-a_{i,\mu(i)}X_i(\mu)} \in [C^{-1}F(X_i(\mu)), CF(X_i(\mu))]$ since $a_{i,\mu(i)} \in [1/C, C]$. For any stable matching μ , $\mathbf{X}(\mu) \succeq \mathbf{X}(\mu_{\text{MOSM}})$ where μ_{MOSM} is the man-optimal stable matching obtained from the man-proposing deferred acceptance, where all the n men are matched with their optimal possible stable partner (and hence achieves best value) simultaneously, and hence

$$\|\mathbf{U}(\mu)\|_1 \geq \frac{1}{C} \|F(\mathbf{X}(\mu))\|_1 \geq \frac{1}{C} \|F(\mathbf{X}(\mu_{\text{MOSM}}))\|_1 \geq \frac{1}{C^2} \|\mathbf{U}(\mu_{\text{MOSM}})\|_1. \quad (39)$$

Thus, it suffices to consider the event where $\|\mathbf{U}(\mu_{\text{MOSM}})\|_1 \leq \theta_2 C^2 \ln n$. Without loss of generality, we may assume that μ_{MOSM} matches man i with woman i for $i \in [n]$. Denote by $R_i \in [n]$ and $X_i = X_{ii} \in \mathbb{R}_+$ the (random) rank of partner and the latent value in μ_{MOSM} for man $i \in [n]$, and let $U_i = F(a_{ii}X_i)$. By definition, $X_i \sim \text{Exp}(a_{ii})$ and $R_i = \sum_{j \neq i} \mathbb{1}\{X_{ij}/a_{ij} < X_i\}$.

We condition on a specific execution of the man-proposing deferred acceptance algorithm, i.e., on the sequence of proposals, which specifies the rank R_i and an ordering over his top R_i most preferred partners for each man. Notice that the specific value X_{ij} only affect the execution through the ordering of proposals, and hence conditional on a particular ordering, the values of the men are independent. Further, the value X_i conditional on R_i and an ordering $w_{j_1} \succeq_{m_i} w_{j_2} \succeq_{m_i} w_{j_1} \succeq \dots \succeq_{m_i} w_{j_{R_i}}$ is equal in distribution to $\tilde{X}_{(R_i)}$, i.e., the R_i -th order statistic of $(\tilde{X}_1, \dots, \tilde{X}_n)$ with $\tilde{X}_j \sim \text{Exp}(a_{ij})$, conditional on $\tilde{X}_{(k)} = \tilde{X}_{j_k}$ for all $k \in [R_i]$. By the representation of exponential order statistics given in Nevzorov (1986), under such conditions,

$$\tilde{X}_{(R_i)} \stackrel{d}{=} \sum_{t=1}^{R_i} \frac{Z_{i,t}}{\sum_{k=t}^n a_{i,j_k}}, \quad (40)$$

where $Z_{i,t} \sim \text{Exp}(1)$ are independently sampled for $t \in [n]$. Conditional on R_i and the sequence j_1, \dots, j_{R_i} , we have

$$\begin{aligned} U_i = F(a_{ii}X_i) &\stackrel{d}{=} F\left(\sum_{t=1}^{R_i} \frac{a_{ii}Z_{i,t}}{\sum_{k=t}^n a_{i,j_k}}\right) \geq F\left(\frac{1}{n} \sum_{t=1}^{R_i} \frac{Z_{i,t}}{C^2}\right) \\ &\geq \frac{1}{n} \sum_{t=1}^{R_i} F(Z_{i,t}/C^2) \geq \frac{1}{C^2 n} \sum_{t=1}^{R_i} F(Z_{i,t}) = \frac{1}{C^2 n} \sum_{t=1}^{R_i} W_{i,t}, \end{aligned} \quad (41)$$

where the second last inequality follows from Jensen's inequality (applied to the concave function F), and $W_{i,t} = F(Z_{i,t}) \sim \text{Unif}([0, 1])$ independently. Thus, conditional on \mathbf{R} , we have $C^2 n \|\mathbf{U}\|_1 \succeq \sum_{i=1}^n \sum_{t=1}^{R_i} W_{i,t}$, independent of the specific ordering (and the identity) of the proposals made. Therefore, we may marginalize over this ordering to get

$$\mathbb{P}\left(\|\mathbf{U}\|_1 \leq \theta_2 C^2 \ln n \mid \mathbf{R}\right) \leq \mathbb{P}\left(\sum_{i=1}^n \sum_{t=1}^{R_i} W_{i,t} \leq \theta_2 C^4 n \ln n \mid \mathbf{R}\right). \quad (42)$$

Whenever $\|\mathbf{R}\|_1 \geq \theta_1 n \ln n$, the probability above is at most $\exp(-\Theta(n \ln n)) \ll \exp(-n^c)$ by Hoeffding's inequality granted that we choose $\theta_2 < \frac{\theta_1}{2C^4}$, and our proof is complete as we marginalize over all possible realizations of \mathbf{R} with $\|\mathbf{R}\|_1 \geq \theta_1 n \ln n$. \square

Lemma A.3. *For any $\kappa \geq 0$ and $\theta_3 > 0$, the expected number of stable matchings μ with $\|\mathbf{U}(\mu)\|_1 \|\mathbf{V}(\mu)\|_1 \geq \theta_3 n (\ln n)^{1/8}$ is upper bounded by $\exp(-\kappa n)$. In particular, with high probability, no such stable matchings exists.*

Proof. It suffices to show that the probability that any fixed matching μ is stable and satisfies $\|\mathbf{U}(\mu)\|_1 \|\mathbf{V}(\mu)\|_1 \geq \theta_3 n (\ln n)^{1/8}$ is upper bounded by $o(e^{-\kappa n}/n!)$ for any $\theta_3 > 0$. Write $I = [0, 1]$ for the unit interval. Let

$$\Omega = \left\{ (\mathbf{u}, \mathbf{v}) \in I^n \times I^n : \|\mathbf{u}\|_1 \|\mathbf{v}\|_1 > \theta_3 n (\ln n)^{1/8} \right\}$$

and let

$$P := \mathbb{P}(\mu \in \mathcal{S}, (\mathbf{U}(\mu), \mathbf{V}(\mu)) \in \Omega) = \int_{\mathbb{R}_+^n \times \mathbb{R}_+^n} p(\mathbf{x}, \mathbf{y}) \cdot \mathbb{1}_\Omega(F(\hat{\mathbf{x}}), F(\hat{\mathbf{y}})) \cdot \prod_{i=1}^n f(\hat{x}_i) f(\hat{y}_i) d\hat{\mathbf{x}} d\hat{\mathbf{y}}, \quad (43)$$

where $f(t) = e^{-t}$ denotes the standard exponential density function. We apply the simple bound (16) on $p(\mathbf{x}, \mathbf{y})$ to obtain

$$\begin{aligned} P &\leq \int_{\mathbb{R}_+^n \times \mathbb{R}_+^n} \prod_{i \neq j} \left(1 - (1 - e^{-\hat{x}_i/C^2})(1 - e^{-\hat{y}_j/C^2}) \right) \cdot \mathbb{1}_\Omega(F(\hat{\mathbf{x}}), F(\hat{\mathbf{y}})) \cdot \prod_{i=1}^n f(\hat{x}_i) f(\hat{y}_i) d\hat{\mathbf{x}} d\hat{\mathbf{y}} \\ &= \int_{I^n \times I^n} \prod_{i \neq j} \left(1 - (1 - (1 - u_i)^{1/C^2})(1 - (1 - v_j)^{1/C^2}) \right) \cdot \mathbb{1}_\Omega(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} \\ &\leq \int_{I^n \times I^n} \prod_{i \neq j} \left(1 - \frac{1}{C^4} u_i v_j \right) \cdot \mathbb{1}_\Omega(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} \\ &\leq \int_{I^n \times I^n} \exp \left(-\frac{1}{C^4} (\|\mathbf{u}\|_1 \|\mathbf{v}\|_1 - \mathbf{u} \cdot \mathbf{v}) \right) \cdot \mathbb{1}_\Omega(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v} \\ &\leq \int_{I^n \times I^n} \exp \left(\frac{1}{C^4} (n - \|\mathbf{u}\|_1 \|\mathbf{v}\|_1) \right) \cdot \mathbb{1}_\Omega(\mathbf{u}, \mathbf{v}) d\mathbf{u} d\mathbf{v}, \end{aligned} \quad (44)$$

where we use the basic facts that $1 - z^{1/C^2} \geq (1 - z)/C^2$ for all $z \in [0, 1]$ and $C \geq 1$ and that $1 + z \leq e^z$ for all $z \in \mathbb{R}$.

Let $s = \|\mathbf{u}\|_1$ and $t = \|\mathbf{v}\|_1$. It is well known (e.g., (Feller, 1971, Ch. I, Sec. 9)) that the probability density of $S = \|\mathbf{U}\|_1$ with U_1, \dots, U_n independent samples from $\text{Unif}(I)$ is bounded above by $\frac{s^{n-1}}{(n-1)!}$. Hence,

$$P \leq \int_{s, t \in [0, n] : st \geq \theta_3 n (\ln n)^{1/8}} e^{(n-st)/C^4} \cdot \frac{n^2 (st)^{n-1}}{(n!)^2} ds dt. \quad (45)$$

Note that when $st \geq n (\ln n)^2$, we have $\exp((n-st)/C^4) \leq \exp((n-n(\ln n)^2)/C^4) = o(\exp(-\kappa n))/n!$ and therefore the region $\{s, t \in [0, n] : st \geq n (\ln n)^2\}$ contributes a negligible amount to the inte-

gral. Hence,

$$\begin{aligned}
P &\leq \frac{o(e^{-\kappa n})}{n!} + \int_{s,t \in [0,n] : \theta_3 n (\ln n)^{1/8} \leq st \leq n (\ln n)^2} e^{(n-st)/C^4} \cdot \frac{n^2 (st)^{n-1}}{(n!)^2} ds dt \\
&\stackrel{(i)}{\leq} \frac{o(e^{-\kappa n})}{n!} + n^2 \cdot e^{(n-\theta_3 n (\ln n)^{1/8})/C^4} \cdot \frac{n^2 (n (\ln n)^2)^{n-1}}{(n!)^2} \\
&\stackrel{(ii)}{\leq} \frac{o(e^{-\kappa n})}{n!} + \frac{1}{n!} \cdot \exp\left(\frac{1}{C^4}(n - \theta_3 n (\ln n)^{1/8}) + 3 \ln n + 2(n-1) \ln \ln n + n\right) = \frac{o(e^{-\kappa n})}{n!}, \quad (46)
\end{aligned}$$

where in step (i) we upper bound the integral by the product of the Lebesgue measure of its domain (bounded by n^2) and the supremum of its integrand, and in step (ii) we invoke Stirling's approximation. \square

Corollary A.4. *For any constant $c \in (0, 1/2)$, there exists a constant $\theta_4 > 0$ such that, with probability $1 - \exp(-n^c)$, there exist no stable matchings μ with $\|\mathbf{U}(\mu)\|_1 \geq \frac{\theta_4 n}{(\ln n)^{7/8}}$.*

Proof. This follows immediately from Lemma A.2 and A.3. \square

Proposition A.5. *Let $\Omega \subset I^n$ be a (sequence of) regions in the n -dimensional hypercube. For $k \in \mathbb{Z}_+$, define interval $I_k = (2^{-k}n, 2^{-k+1}n]$. If*

$$\mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k) \otimes n}(\mathbf{W} \in \Omega, \|\mathbf{W}\|_1 \in I_k) \leq \frac{g(n)}{e^{6n} C^{8n}} \quad (47)$$

for some function $g(n)$ and uniformly for all $k \in \mathbb{Z}_+$, we can guarantee that the expected number of stable matchings μ with value vector $\mathbf{X}(\mu) \in \mathbb{R}^n$ satisfying $\mathbf{U}(\mu) \in \Omega$ is upper bounded by $g(n) + e^{-\Theta(n^2)}$; in particular, with high probability, no such stable matchings exist if $g(n) = o(1)$.

Proof. We focus on a fixed matching μ and by union bound, it suffices to show that

$$\mathbb{P}(\mu \in \mathcal{S}, \mathbf{U}(\mu) \in \Omega) = \frac{g(n) + e^{-\Theta(n^2)}}{n!} \quad (48)$$

under the condition of (47). The same chain of reasoning as in (44) gives

$$\begin{aligned}
P &:= \mathbb{P}(\mu \in \mathcal{S}, \mathbf{U}(\mu) \in \Omega) \leq \int_{I^n \times I^n} \exp\left(\frac{1}{C^4}(n - \|\mathbf{u}\|_1 \|\mathbf{v}\|_1)\right) \cdot \mathbb{1}_\Omega(\mathbf{u}) d\mathbf{u} d\mathbf{v} \\
&\leq e^n \int_{I^n \times I^n} \exp(-\|\mathbf{u}\|_1 \|\mathbf{v}\|_1 / C^4) \cdot \mathbb{1}_\Omega(\mathbf{u}) d\mathbf{u} d\mathbf{v} \quad (49)
\end{aligned}$$

since $C \geq 1$. Observing that $\|\mathbf{U}(\mu)\|_1 = 0$ and $\|\mathbf{V}(\mu)\|_1 = 0$ are both probability zero events, we may split the domain of the integral above into sub-regions according to which intervals $\|\mathbf{U}(\mu)\|_1$ and $\|\mathbf{V}(\mu)\|_1$ fall into, and then bound the value of the integral within each sub-region. That is, with the help of the monotone convergence theorem to interchange summation with integral,

$$\begin{aligned}
P &\leq e^n \sum_{k,\ell=1}^{\infty} \int_{I^n \times I^n} \exp(-\|\mathbf{u}\|_1 \|\mathbf{v}\|_1 / C^4) \cdot \mathbb{1}_\Omega(\mathbf{u}) \mathbb{1}_{I_k}(\|\mathbf{u}\|_1) \mathbb{1}_{I_\ell}(\|\mathbf{v}\|_1) d\mathbf{u} d\mathbf{v} \\
&\leq e^n \sum_{k,\ell=1}^{\infty} \int_{I^n \times I^n} \exp(-2^{-k-\ell} n^2 / C^4) \cdot \mathbb{1}_\Omega(\mathbf{u}) \mathbb{1}_{I_k}(\|\mathbf{u}\|_1) \mathbb{1}_{I_\ell}(\|\mathbf{v}\|_1) d\mathbf{u} d\mathbf{v}. \quad (50)
\end{aligned}$$

For all $k \in \mathbb{Z}_+$, we have $e^{2n} \exp(-2^k \|\mathbf{u}\|_1) \geq 1$ whenever $\mathbf{u} \in I^n$ with $\|\mathbf{u}\|_1 \in I_k$. Thus,

$$\begin{aligned} P &\leq e^{5n} \sum_{k,\ell=1}^{\infty} \int_{I^n \times I^n} \exp(-2^{-k-\ell} n^2 / C^4) \cdot \mathbb{1}_{\Omega}(\mathbf{u}) \mathbb{1}_{I_k}(\|\mathbf{u}\|_1) \mathbb{1}_{I_\ell}(\|\mathbf{v}\|_1) \cdot \exp(-2^k \|\mathbf{u}\|_1 - 2^\ell \|\mathbf{v}\|_1) d\mathbf{u} d\mathbf{v} \\ &= e^{5n} \sum_{k,\ell=1}^{\infty} 2^{-(k+\ell)n} \exp(-2^{-k-\ell} n^2 / C^4) \mathbb{E}_{\mathbf{U} \sim \text{Exp}(2^k)^{\otimes n}, \mathbf{V} \sim \text{Exp}(2^\ell)^{\otimes n}} [\mathbb{1}_{\Omega}(\mathbf{U}) \mathbb{1}_{I^n}(\mathbf{V}) \mathbb{1}_{I_k}(\|\mathbf{U}\|_1) \mathbb{1}_{I_\ell}(\|\mathbf{V}\|_1)]. \end{aligned} \quad (51)$$

Observe that all the terms with $k + \ell \geq n$ combined contribute at most

$$\begin{aligned} e^{5n} \sum_{k+\ell \geq n} 2^{-(k+\ell)n} &= e^{5n} \left(\sum_{k=1}^{n-1} 2^{-kn} \sum_{\ell=n-k}^{\infty} 2^{-\ell n} + \sum_{k=n}^{\infty} 2^{-kn} \sum_{\ell=1}^{\infty} 2^{-\ell n} \right) \\ &= e^{5n} \cdot 2^{-n^2-n} \cdot \frac{n(1-2^{-n})+1}{(1-2^{-n})^2} = \frac{e^{-\Theta(n^2)}}{n!}, \end{aligned} \quad (52)$$

which is negligible. Therefore, we only need to consider $O(n^2)$ terms and

$$\begin{aligned} P &\leq \frac{e^{-\Theta(n^2)}}{n!} + n^2 e^{5n} \max_{k,\ell \in \mathbb{Z}_+} \frac{\mathbb{E}_{\mathbf{U} \sim \text{Exp}(2^k)^{\otimes n}, \mathbf{V} \sim \text{Exp}(2^\ell)^{\otimes n}} [\mathbb{1}_{\Omega}(\mathbf{U}) \mathbb{1}_{I^n}(\mathbf{V}) \mathbb{1}_{I_k}(\|\mathbf{U}\|_1) \mathbb{1}_{I_\ell}(\|\mathbf{V}\|_1)]}{2^{(k+\ell)n} \exp(2^{-k-\ell} n^2 / C^4)} \\ &\leq \frac{e^{-\Theta(n^2)}}{n!} + e^{6n} \max_{k \in \mathbb{Z}_+} \frac{\mathbb{P}_{\mathbf{U} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{U} \in \Omega, \|\mathbf{U}\|_1 \in I_k)}{2^{(k+\ell)n} \exp(2^{-k-\ell} n^2 / C^4)}. \end{aligned} \quad (53)$$

Let $\alpha = 2^{-(k+\ell)}$. When $\alpha \leq C^8/n$, the denominator of the second term can be bounded as $\alpha^{-n} e^{\alpha n^2 / C^4} \geq n^n / C^{8n} \geq n! / C^{8n}$; when $\alpha > C^8/n$, let $\tau = n\alpha > C^8$ and we have $\alpha^{-n} e^{\alpha n^2 / C^4} = \frac{n^n}{\tau^n} e^{n\tau / C^4} = n^n \exp(n(\tau / C^4 - \ln \tau)) \geq n^n \exp(n(C^4 - 4 \ln C)) \geq n!$. Thus, the denominator is bounded below by $n! / C^{8n}$ and

$$P \leq \frac{e^{-\Theta(n^2)}}{n!} + \frac{e^{6n} C^{8n}}{n!} \max_{k \in \mathbb{Z}_+} \mathbb{P}_{\mathbf{U} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{U} \in \Omega, \|\mathbf{U}\|_1 \in I_k) \leq \frac{1}{n!} (e^{-\Theta(n^2)} + g(n)). \quad (54)$$

The claim follows immediately. \square

Lemma A.6. *For any fixed $\delta > 0$ and $\kappa > 0$, there exists an absolute constant θ_5 (depending on δ, κ , and C) such that, with probability $1 - \exp(-\kappa n)$, there exist no stable matchings μ with $\|\mathbf{X}_\delta(\mu)\|_\infty \leq \theta_5$. (Recall that $\mathbf{X}_\delta(\mu)$ is the value vector of the $(1 - \delta)$ -partial matching obtained from μ that excludes the least happy $\delta/2$ fraction of men and women.)*

Proof. Since $\hat{\mathbf{X}}$ and \mathbf{X} differ by at most a factor of C component-wise, we have

$$\|\mathbf{X}_\delta(\mu)\|_\infty \leq X_{(n - \lfloor \delta n / 2 \rfloor)}(\mu) \leq C \hat{X}_{(n - \lfloor \delta n / 2 \rfloor)}(\mu) = -C \log(1 - U_{(n - \lfloor \delta n / 2 \rfloor)}(\mu)). \quad (55)$$

Thus, it suffices to bound the upper $\delta/2$ quantile $U_{(n - \lfloor \delta n / 2 \rfloor)}(\mu)$ away from 1.

Let $\Omega = \{\mathbf{u} \in I^n : \mathbf{u}_{(n - \lfloor \delta n / 2 \rfloor)} > 1 - e^{-s}\}$ for some $s \geq 1$ that we will specified later. Then $\mathbf{W} \in \Omega$ implies that $\sum_{i=1}^n \mathbb{1}_{(1 - e^{-s}, 1]}(W_i) \geq \delta n / 2$. For $W_i \sim \text{Exp}(2^k)$, we have

$$\mathbb{P}(W_i \in (1 - e^{-s}, 1]) = \int_{1 - e^{-s}}^1 2^k e^{-2^k t} dt \leq e^{-s} \cdot 2^k e^{-2^{k-1}} \leq e^{-s}.$$

Thus, $\sum_{i=1}^n \mathbb{1}_{(1-e^{-s}, 1]}(W_i)$ is stochastically dominated by a $\text{Bin}(n, e^{-s})$ random variable, and as a result

$$\begin{aligned} \mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{W} \in \Omega, \|\mathbf{W}\|_1 \in I_k) &\leq \mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{W} \in \Omega) \\ &\leq \mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}\left(\sum_{i=1}^n \mathbb{1}_{(1-e^{-s}, 1]}(W_i) \geq \frac{\delta n}{2}\right) \leq \mathbb{P}_{Z \sim \text{Bin}(n, e^{-s})}\left(Z \geq \frac{\delta n}{2}\right) \leq \exp(-nD(\delta/2 \| e^{-s})). \end{aligned}$$

Since $D(\delta/2 \| z) \rightarrow \infty$ as $z \rightarrow 0$, it suffices to take s sufficiently large to guarantee $D(\delta/2 \| e^{-s}) > \kappa + 6 + 8 \log C$. Proposition A.5 then guarantees with probability $1 - \exp(-\kappa n)$ that no stable matchings μ have $\mathbf{U}(\mu) \in \Omega$, and as a result of (55), with at least the desired probability, no stable matchings μ should have $\|\mathbf{X}_\delta(\mu)\|_\infty > \theta_5 := Cs$. \square

Lemma A.7. *For any $\delta > 0$ and $\kappa > 0$, there exists an absolute constant θ_6 (again depending on δ, κ , and C) such that, with probability $1 - \exp(-\kappa n)$, there exist no stable matchings μ with $\|\mathbf{X}_\delta(\mu)\|_1 \geq \theta_6 \|\mathbf{U}(\mu)\|_1$.*

Proof. Take $\theta_6 = C \sup_{0 < x \leq C\theta_5} \frac{x}{F(x)} = \frac{C^2\theta_5}{1 - e^{-C\theta_5}}$, where θ_5 is the constant in Lemma A.6. Assume that $\|\mathbf{X}_\delta(\mu)\|_\infty \leq \theta_5$ for all stable matchings μ since the probability otherwise is at most $\exp(-\kappa n)$ as desired. Note that for each i in the support of $\mathbf{X}_\delta(\mu)$ (i.e., $(X_\delta)_i(\mu) > 0$), we have

$$\hat{X}_i(\mu) \leq CX_i(\mu) = C(X_\delta)_i(\mu) \leq C\theta_5, \quad (56)$$

and subsequently

$$U_i(\mu) = F(\hat{X}_i(\mu)) \geq \frac{C\hat{X}_i(\mu)}{\theta_6} \geq \frac{X_i(\mu)}{\theta_6} = \frac{(X_\delta)_i(\mu)}{\theta_6}, \quad (57)$$

and this final inequality is trivial for any i not in the support of $\mathbf{X}_\delta(\mu)$. The claim then follows immediately. \square

Lemma A.8. *For any fixed $\delta > 0$ and $\kappa > 0$, there exists an absolute constant θ_7 (depending on δ, κ , and C) such that, with probability $1 - \exp(-\kappa n)$, there exist no stable matchings μ with $\|\mathbf{X}_\delta(\mu)\|_2^2 \geq \theta_7 \|\mathbf{U}(\mu)\|_1^2/n$.*

Proof. Taking advantage of Lemma A.6, let us assume that $X_{(n - \lfloor \delta n/2 \rfloor)}(\mu) \leq \theta_5$ is satisfied simultaneously by all stable matchings μ (see the proof, in particular (55), for more details); the event otherwise has probability bounded by $\exp(-\kappa n)$.

Notice that

$$\|\mathbf{X}_\delta(\mu)\|_2^2 \leq \sum_{i=1}^{n - \lfloor \delta n/2 \rfloor} X_{(i)}(\mu)^2 \leq C^2 \sum_{i=1}^{n - \lfloor \delta n/2 \rfloor} \hat{X}_{(i)}(\mu)^2 \leq \theta'_6 \sum_{i=1}^{n - \lfloor \delta n/2 \rfloor} U_{(i)}(\mu)^2,$$

where the (i) subscript denotes the i -th (lower) order statistics (and in particular, $\hat{X}_{(i)}(\mu)$ is the i -th smallest entry of $\hat{\mathbf{X}}(\mu)$) with $\theta'_6 = C^2 \left(\frac{\theta_5}{F(\theta_5)}\right)^2$. Now it suffices to compare $\sum_{i=1}^{n - \lfloor \delta n/2 \rfloor} U_{(i)}(\mu)^2$ with $\|\mathbf{U}(\mu)\|_1^2/n$.

Consider $\Omega := \{\mathbf{w} \in I^n : \sum_{i=1}^{n - \lfloor \delta n/2 \rfloor} w_{(i)}^2 \geq \gamma \|\mathbf{w}\|_1^2/n\}$ for some $\gamma \in \mathbb{R}_+$ to be specified. By Proposition A.5, it suffices to show that for some appropriate value of γ we have

$$\mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{W} \in \Omega, \|\mathbf{W}\|_1 \in I_k) \leq e^{-(\kappa+6)n} C^{-8n}$$

for all $k \in \mathbb{Z}_+$. Observe that

$$\begin{aligned}
\mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{W} \in \Omega, \|\mathbf{W}\|_1 \in I_k) &= \mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}\left(\sum_{i=1}^{n-\lfloor \delta n/2 \rfloor} W_{(i)}^2 \geq \frac{\gamma \|\mathbf{W}\|_1^2}{n}, 2^{-k}n < \|\mathbf{W}\|_1 \leq 2^{-k+1}n\right) \\
&\leq \mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}\left(\sum_{i=1}^{n-\lfloor \delta n/2 \rfloor} W_{(i)}^2 \geq \gamma 2^{-2k}n\right) \\
&\leq \mathbb{P}_{\mathbf{W} \sim \text{Exp}(1)^{\otimes n}}\left(\sum_{i=1}^{n-\lfloor \delta n/2 \rfloor} W_{(i)}^2 \geq \gamma n\right) \\
&\leq \mathbb{P}_{\mathbf{W} \sim \text{Exp}(1)^{\otimes n}}(W_{(n-\lfloor \delta n/2 \rfloor)} \geq \sqrt{\gamma}) \\
&\leq \mathbb{P}_{Z \sim \text{Bin}(n, e^{-\sqrt{\gamma}})}\left(Z \geq \frac{\delta n}{2}\right).
\end{aligned}$$

By the large deviation bound for binomial distribution, choosing γ sufficiently large such that $D(\delta/2 \| e^{-\sqrt{\gamma}}) > \kappa + 6 + 8 \log C$ ensures that this probability is $o(e^{-(\kappa+6)n} C^{-8n})$. This finishes the proof with the choice of $\theta_7 = \theta'_6 \gamma$. \square

Remark A.9. *This is the only part of our analysis that relies on the δ -truncation of values. Without the truncation, $\frac{1}{n} \|\mathbf{W}\|_2^2$ would concentrate poorly – in fact not even having a finite mean – for $\mathbf{W} \sim \text{Exp}(1)^{\otimes n}$.*

Corollary A.10. *For any constant $c \in (0, 1/2)$, there exists a constant $\theta_8 > 0$ such that, with probability $1 - \exp(-n^c)$, there exist no stable matchings μ with $\sum_{i=1}^n (X_\delta)_i \mathbb{1}_{[2/C, \infty)}((X_\delta)_i) \geq \theta_8 \|\mathbf{U}\|_1 / (\ln n)^{7/8}$.*

Proof. Notice that

$$\|\mathbf{X}_\delta\|_2^2 \geq \sum_{i=1}^n (X_\delta)_i^2 \mathbb{1}_{[2/C, \infty)}((X_\delta)_i) \geq \frac{2}{C} \sum_{i=1}^n (X_\delta)_i \mathbb{1}_{[2/C, \infty)}((X_\delta)_i).$$

The statement then follows from Lemma A.8 and Corollary A.4. \square

Lemma A.11. *For any fixed $\delta > 0$ and $\kappa > 0$, there exists an absolute constant θ_9 (depending on δ, κ , and C) such that, with probability $1 - \exp(-\kappa n)$, there exist no stable matchings μ with $\|\mathbf{X}_\delta(\mu)\|_1 \leq \theta_9 \|\mathbf{U}(\mu)\|_1$.*

Proof. Since $\mathbf{X} \succeq \mathbf{U}$ component-wise, we have $\|\mathbf{X}_\delta(\mu)\|_1 \geq \|\mathbf{U}_\delta(\mu)\|_1$. Thus, it suffices to consider the condition $\|\mathbf{U}_\delta(\mu)\|_1 \leq \theta_9 \|\mathbf{U}(\mu)\|_1$.

Consider $\Omega := \{\mathbf{w} \in I^n : \exists S \subseteq [n], |S| = n - \lfloor \delta n \rfloor, \sum_{i \in S} w_i \leq \alpha \|\mathbf{w}\|_1\}$ for some $\alpha \in \mathbb{R}_+$ to be specified. By union bound, for any $k \in \mathbb{Z}_+$,

$$\begin{aligned}
\mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{W} \in \Omega, \|\mathbf{W}\|_1 \in I_k) &\leq \binom{n}{\lfloor \delta n \rfloor} \mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}\left(\sum_{i=1}^{n-\lfloor \delta n \rfloor} w_i \leq \alpha \|\mathbf{W}\|_1 \leq 2^{-k+1} \alpha n\right) \\
&= \binom{n}{\lfloor \delta n \rfloor} \mathbb{P}_{\mathbf{W} \sim \text{Exp}(1)^{\otimes n}}\left(\sum_{i=1}^{n-\lfloor \delta n \rfloor} w_i \leq 2\alpha n\right) \\
&= \exp(-h(\delta)n + o(n)) \cdot \left(\frac{2\alpha e}{1-\delta}\right)^{n-\lfloor \delta n \rfloor},
\end{aligned}$$

where in the last step we use Stirling's approximation to bound the first factor and standard (lower) concentration of $\text{Exp}(1)$ to bound the probability term (e.g., see Lemma B.1). For α sufficiently small, e.g., $\alpha < \exp\left(\frac{1}{1-\delta}(h(\delta) - \kappa - 6 - 8 \ln C) - h(\delta)\right)$, we have

$$\mathbb{P}_{\mathbf{W} \sim \text{Exp}(2^k)^{\otimes n}}(\mathbf{W} \in \Omega, \|\mathbf{W}\|_1 \in I_k) \leq e^{-(\kappa+6)n} C^{-8n}$$

for all $k \in \mathbb{Z}_+$. Invoking Lemma A.5 concludes the proof with $\theta_9 = \alpha$. \square

Corollary A.12. *For any constant $c \in (0, 1/2)$, there exists a constant $\theta_{10} > 0$ such that, with probability $1 - \exp(-n^c)$, there exist no stable matchings μ with $\|\mathbf{X}_\delta(\mu)\|_1 \leq \theta_{10} \ln n$.*

Proof. This follows from Lemmas A.2 and A.11, with $\theta_{10} = \theta_2 \theta_9$. \square

The following Corollary combines all the previous into the typical behavior of value vectors in stable matchings.

Corollary A.13. *Define $\mathcal{R}^*(\mu) \subseteq \mathbb{R}_+^n \times \mathbb{R}_+^n$, in the context of a matching μ , to be the set of all pairs of vectors $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n$ that satisfy all of the following conditions:*

$$\theta_2 \ln n \leq \|\mathbf{u}\|_1, \|\mathbf{v}\|_1 \leq \frac{\theta_4 n}{(\ln n)^{7/8}}, \quad (58)$$

$$\|\mathbf{u}\|_1 \|\mathbf{v}\|_1 \leq \theta_3 n (\ln n)^{1/8}, \quad (59)$$

$$\|\mathbf{x}_\delta\|_1 \leq \theta_6 \|\mathbf{u}\|_1 \text{ and } \|\mathbf{y}_\delta\|_1 \leq \theta_6 \|\mathbf{v}\|_1, \quad (60)$$

$$\|\mathbf{x}_\delta\|_2^2 \leq \frac{\theta_7 \|\mathbf{u}\|_1^2}{n} \text{ and } \|\mathbf{y}_\delta\|_2^2 \leq \frac{\theta_7 \|\mathbf{v}\|_1^2}{n}, \quad (61)$$

$$\sum_{i=1}^n (x_\delta)_i \mathbb{1}_{[2/C, \infty)}((x_\delta)_i) \leq \frac{\theta_8 \|\mathbf{u}\|_1}{(\ln n)^{7/8}} \text{ and } \sum_{i=1}^n (y_\delta)_i \mathbb{1}_{[2/C, \infty)}((y_\delta)_i) \leq \frac{\theta_8 \|\mathbf{v}\|_1}{(\ln n)^{7/8}}, \quad (62)$$

$$\|\mathbf{x}_\delta\|_1, \|\mathbf{y}_\delta\|_1 \geq \theta_{10} \ln n, \quad (63)$$

where $u_i = F(a_{i, \mu(i)} x_i)$ and $v_j = F(b_{j, \mu^{-1}(j)} y_j)$ for $i, j \in [n]$; \mathbf{x}_δ and \mathbf{y}_δ denote the truncated version of \mathbf{x} and \mathbf{y} ; $\theta_2, \theta_3, \theta_6, \theta_7, \theta_4, \theta_8, \theta_{10} \in \mathbb{R}_+$ are absolute constants (independent of μ) chosen appropriately as in Lemmas A.2, A.3, A.7, A.8, and Corollaries A.4, A.10, and A.12. Then, for any $c \in (0, 1/2)$, with probability $1 - \exp(-n^c)$, $(\mathbf{X}(\mu), \mathbf{Y}(\mu)) \in \mathcal{R}^*(\mu)$ for all stable matchings μ .

The proof simply summarizes the aforementioned Lemmas and Corollaries and shall be omitted.

Proposition 6.3. *For any $\delta > 0$ and $c \in (0, 1/2)$, there exists an absolute constant $\theta \in (0, \infty)$ such that the probability that a matching μ is stable with value vectors not satisfying*

$$\frac{p_\mu(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu))}{q(\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu))} \leq \exp\left(\frac{\theta n}{(\log n)^{1/2}}\right) \quad (\star)$$

is $\frac{\exp(-n^c)}{n!}$. In other words, with high probability, there exist no stable matchings μ whose post-truncation value vectors $\mathbf{X}_\delta(\mu)$ and $\mathbf{Y}_\delta(\mu)$ satisfy (\star) .

Proof. Note that $1 - e^{-tx} \geq \left(tx - \frac{t^2x}{2}\right)$ for all $x, t \geq 0$. In particular, $1 - e^{-tx} \geq \left(tx - \frac{t^2x}{2}\right) \mathbb{1}_{[0, 2/t]}(x) \geq 0$. Using this to approximate $p(\mathbf{x}, \mathbf{y})$ gives

$$\begin{aligned} p(\mathbf{x}, \mathbf{y}) &= \prod_{i \neq j} \left(1 - (1 - e^{-a_{ij}x_i})(1 - e^{-b_{ji}y_j})\right) \\ &\leq \prod_{i \neq j} \left(1 - \mathbb{1}_{[0, 2/a_{ij}]}(x_i) \mathbb{1}_{[0, 2/b_{ji}]}(y_j) \left(a_{ij}x_i - \frac{a_{ij}^2}{2}x_i^2\right) \left(b_{ji}y_j - \frac{b_{ji}^2}{2}y_j^2\right)\right) \\ &\leq \exp \left(- \sum_{i \neq j} \mathbb{1}_{[0, 2/C]}(x_i) \mathbb{1}_{[0, 2/C]}(y_j) \left(a_{ij}x_i - \frac{a_{ij}^2}{2}x_i^2\right) \left(b_{ji}y_j - \frac{b_{ji}^2}{2}y_j^2\right)\right). \end{aligned} \quad (64)$$

Taking logarithm for simplicity and expanding the expression above gives

$$\begin{aligned} \ln p(\mathbf{x}, \mathbf{y}) &\leq - \sum_{i \neq j} \left(a_{ij}b_{ji}x_iy_j - (\mathbb{1}_{(2/C, \infty)}(x_i) + \mathbb{1}_{(2/C, \infty)}(y_j))a_{ij}b_{ji}x_iy_j \right. \\ &\quad \left. - \mathbb{1}_{[0, 2/C]}(x_i) \mathbb{1}_{[0, 2/C]}(y_j) \left(a_{ij}^2b_{ji}x_i^2y_j + a_{ij}b_{ji}^2x_iy_j^2 \right) \right) \\ &\leq - \sum_{i,j=1}^n a_{ij}b_{ji}x_iy_j + \sum_{i=1}^n C^2x_iy_i \\ &\quad + \sum_{i,j=1}^n \left(C^2(\mathbb{1}_{(2/C, \infty)}(x_i) + \mathbb{1}_{(2/C, \infty)}(y_j))x_iy_j + C^3(x_i^2y_j + x_iy_j^2) \right). \end{aligned} \quad (65)$$

Notice that $-\ln q(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^n a_{ij}b_{ji} \frac{x_i}{a_{ii}} \frac{y_j}{b_{jj}}$. Thus,

$$\begin{aligned} \ln \frac{p(\mathbf{x}, \mathbf{y})}{q(\mathbf{x}, \mathbf{y})} &\leq C^2 \mathbf{x}^\top \mathbf{y} + C^2 \left(\|\mathbf{x}\|_1 \sum_{i=1}^n \mathbb{1}_{(2/C, \infty)}(y_j)y_j + \|\mathbf{y}\|_1 \sum_{i=1}^n \mathbb{1}_{(2/C, \infty)}(x_i)x_i \right) \\ &\quad + C^3 (\|\mathbf{x}\|_2^2 \|\mathbf{y}\|_1 + \|\mathbf{x}\|_1 \|\mathbf{y}\|_2^2). \end{aligned} \quad (66)$$

In light of Corollary A.13, it suffices to upper bound $\ln \frac{p(\mathbf{x}_\delta, \mathbf{y}_\delta)}{q(\mathbf{x}_\delta, \mathbf{y}_\delta)}$ by $cn/(\ln n)^{1/2}$ for all $(\mathbf{x}, \mathbf{y}) \in \mathcal{R}^*(\mu)$ and for all μ . To simplify notation, we will make the dependency on μ implicit in the rest of the proof. By Cauchy-Schwarz inequality, the first term in (66), up to a factor of C^2 , is at most

$$\|\mathbf{x}_\delta\|_2 \|\mathbf{y}_\delta\|_2 \leq \frac{\theta_7 \|\mathbf{u}\|_1 \|\mathbf{v}\|_1}{n} \leq \theta_3 \theta_7 (\ln n)^{1/8} = o\left(\frac{n}{(\ln n)^{1/2}}\right)$$

by (61) and (59). The middle term in (66), up to a factor of $2C^2$, is at most

$$\|\mathbf{x}_\delta\|_1 \sum_{i=1}^n \mathbb{1}_{(2/C, \infty)}((y_\delta)_j)(y_\delta)_j \leq \theta_6 \|\mathbf{u}\|_1 \frac{\|\mathbf{v}\|_1}{(\ln n)^{7/8}} \leq \theta_3 \theta_6 \frac{n}{(\ln n)^{3/4}}$$

by (60), (62), and (59). Finally, the last term, up to a factor of $2C^2$, is upper bounded by

$$\begin{aligned} \|\mathbf{x}_\delta\|_2^2 \|\mathbf{y}_\delta\|_1 &= \frac{\|\mathbf{x}_\delta\|_2^2}{\|\mathbf{u}\|_1^2} \frac{\|\mathbf{y}_\delta\|_1}{\|\mathbf{v}\|_1} \frac{1}{\|\mathbf{v}\|_1} (\|\mathbf{u}\|_1 \|\mathbf{v}\|_1)^2 \\ &\leq \frac{\theta_7}{n} \cdot \theta_6 \cdot \frac{1}{\theta_2 \ln n} \cdot \theta_3^2 n^2 (\ln n)^{1/4} = \frac{\theta_7 \theta_6 \theta_3^2}{\theta_2} \frac{n}{(\ln n)^{3/4}} \end{aligned}$$

due to (61), (60), (58), and (59). Putting these together gives the proposition. \square

B Some concentration inequalities

B.1 Concentration for independent non-identically distributed exponential random variables

Lemma B.1. *Let $\mathbf{u} \in \mathbb{R}_+^n$ be a vector with $\|\mathbf{u}\|_1 = n$ and let \mathbf{Z} be a random vector with independent $\text{Exp}(1)$ components. Then for any $t \in [0, 1)$, we have*

$$\mathbb{P}(\mathbf{u} \cdot \mathbf{Z} \leq tn) \leq (te^{1-t})^n \prod_{i=1}^n u_i^{-1} \leq (te)^n \prod_{i=1}^n u_i^{-1}. \quad (67)$$

In fact, the upper bound given by the second inequality holds trivially when $t \geq 1$ and is invariant under simultaneous scaling of \mathbf{u} and t .

Further, when $1/K \leq u_i \leq K$ for some constant $K \geq 1$, we have

$$\mathbb{P}(\mathbf{u} \cdot \mathbf{Z} \leq tn) \geq e^{-O(n^{2/3})} (te^{1-Kt})^n \prod_{i=1}^n u_i^{-1}. \quad (68)$$

For $t = o(1)$, this lower bound becomes

$$e^{-O(n^{2/3}) + (1-K)tn} (te^{1-t})^n \prod_{i=1}^n u_i^{-1} = e^{o(n)} (te)^n \prod_{i=1}^n u_i^{-1},$$

indicating that the upper bound is tight up to a factor of $e^{o(n)}$. In particular, when $t = O(n^{-1/3})$, the gap is $e^{O(n^{2/3})}$.

Proof. First we establish the upper bound. Directly applying Chernoff's method on $\mathbf{u} \cdot \mathbf{Z}$, we have

$$\mathbb{P}(\mathbf{u} \cdot \mathbf{Z} \leq tn) \leq \inf_{\lambda \geq 0} \frac{\mathbb{E}[\exp(-\lambda \mathbf{u} \cdot \mathbf{Z})]}{\exp(-\lambda tn)} = \inf_{\lambda \geq 0} e^{\lambda tn} \prod_{i=1}^n \frac{1}{1 + \lambda u_i}. \quad (69)$$

Taking $\lambda = 1/t - 1$ (which is the minimizer when $\mathbf{u} = \mathbf{1}$) gives

$$\mathbb{P}(\mathbf{u} \cdot \mathbf{X} \leq tn) \leq e^{n-tn} \prod_{i=1}^n \frac{t}{t + (1-t)u_i}. \quad (70)$$

Notice that $\mathbf{u} \mapsto \sum_{i=1}^n \log u_i$ is a concave function on \mathbb{R}_+^n , and hence

$$\prod_{i=1}^n (t + (1-t)u_i) \geq \left(\prod_{i=1}^n u_i \right)^{1-t} \geq \prod_{i=1}^n u_i$$

since $\prod_{i=1}^n u_i \leq (n^{-1} \sum_{i=1}^n u_i)^n = 1$. Plugging the above inequality into (70) gives the desired upper bound.

Now we establish the tightness of the bound under the additional assumption that $1/K \leq u_i \leq K$ for all $i \in [n]$. Consider independent random variables $W_i \sim \text{Exp}(u_i R/t)$ for $i = 1, \dots, n$ with $R = 1 + n^{-1/3}$, so that by Chebyshev's inequality

$$g_n := \mathbb{P}(\mathbf{u} \cdot \mathbf{W} \leq tn) = \mathbb{P}_{T \sim \Gamma(n,1)}(T \leq nR) \geq 1 - n^{-1/3}.$$

For convenience, we similarly write

$$p_n := \mathbb{P}(\mathbf{u} \cdot \mathbf{Z} \leq tn)$$

and write the (joint) distributions of \mathbf{Z} and \mathbf{W} as $P_n = \text{Exp}(1)^{\otimes n}$ and $Q_n = \bigotimes_{i=1}^n \text{Exp}(u_i R/t)$, respectively. Applying the data processing inequality to the channel \mathcal{C} that maps $\zeta \in \mathbb{R}^n$ to $\mathbb{1}\{\mathbf{u} \cdot \zeta \leq tn\}$ gives

$$\begin{aligned} D(Q_n \| P_n) &\geq D(\mathcal{C}(Q_n) \| \mathcal{C}(P_n)) = q_n \log \frac{q_n}{p_n} + (1 - q_n) \log \frac{1 - q_n}{1 - p_n} \\ &= -q_n \log p_n + (1 - q_n) \log(1 - p_n) + (q_n \log q_n + (1 - q_n) \log(1 - q_n)), \end{aligned} \quad (71)$$

where $D(\cdot \| \cdot)$ denotes the Kullback-Leibler (KL) divergence between two probability distributions. A direct computation gives

$$\begin{aligned} D(Q_n \| P_n) &= \sum_{i=1}^n \left(\frac{t}{R u_i} - 1 - \log \frac{t}{R u_i} \right) \\ &\leq \sum_{i=1}^n \left(\frac{K t}{R} - 1 - \log \frac{t}{R u_i} \right) \\ &= -n + R^{-1} K t n - n \log t + n \log R + \sum_{i=1}^n \log u_i \end{aligned} \quad (72)$$

$$\leq -n + K t n - n \log t + n^{2/3} + \sum_{i=1}^n \log u_i. \quad (73)$$

Combining this with (71) and letting $n \rightarrow \infty$ gives

$$-n + K t n - n \log t + n^{2/3} + \sum_{i=1}^n \log u_i \geq -(1 - O(n^{-1/3})) \log p_n + o(1), \quad (74)$$

where we used the fact that $\log(1 - p_n) \rightarrow 0$ (due to our upper bound). Exponentiating both sides gives the desired lower bound for p_n . \square

As a consequence, we have the following lemma.

Lemma B.2. *Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}_+^n$ be two vectors with bounded components, namely $\|\mathbf{u}\|_1 = \|\mathbf{v}\|_1 = n$ and $1/K \leq u_i, v_i \leq K$ for some fixed $K \geq 1$. For independent $Z_1, \dots, Z_n \sim \text{Exp}(1)$, we have*

$$\mathbb{P} \left(\left| \frac{\mathbf{u} \cdot \mathbf{Z}}{t \mathbf{u} \cdot \mathbf{v}^{-1}} - 1 \right| > \zeta \mid \mathbf{v} \cdot \mathbf{Z} < tn \right) \leq \exp(-\Theta(n \zeta^2)) \quad (75)$$

for $t = o(1)$ and any fixed constant $\zeta > 0$, where \mathbf{v}^{-1} denotes the component-wise inverse of vector \mathbf{v} .

Notice that this result is invariant under simultaneous scaling of vector \mathbf{u} , \mathbf{v} , and t . Essentially, we only need $tn/\|\mathbf{v}\|_1 = o(1)$ and bounded ratios between among the entries of \mathbf{u} and \mathbf{v} . Further, the result remains unchanged if $Z_i \sim \text{Exp}(c_i)$ independently with c_i 's bounded on some $[1/K', K']$; the c_i 's can simply be absorbed into \mathbf{u} and \mathbf{v} .

Proof. We first prove the concentration bound for the lower tail.

Writing

$$\begin{aligned} \mathbb{P}(u \cdot x < (1 - \zeta)tu \cdot v^{-1} | v \cdot x < tn) &= \frac{\mathbb{P}(u \cdot x < (1 - \zeta)tu \cdot v^{-1}, v \cdot x < tn)}{\mathbb{P}(v \cdot x < tn)} \\ &\leq \frac{\mathbb{P}((\lambda u + (1 - \lambda)v) \cdot x < (1 - \zeta)\lambda tu \cdot v^{-1} + (1 - \lambda)tn)}{\mathbb{P}(v \cdot x < tn)} \end{aligned} \quad (76)$$

for some $\lambda > 0$ to be determined later, the previous Lemma bounds the numerator by

$$t^n \left(1 - \lambda + \frac{(1 - \zeta)\lambda u \cdot v^{-1}}{n}\right)^n e^{n - (1 - \zeta)\lambda tu \cdot v^{-1} - (1 - \lambda)tn} \prod_{i=1}^n \frac{1}{\lambda u_i + (1 - \lambda)v_i}.$$

For lower bounding the denominator $\mathbb{P}(v \cdot x < tn)$, Lemma B.1 indicates that for $t = o(n)$, the denominator is well approximated by $t^n e^{n - tn} \prod_{i=1}^n v_i^{-1}$, up to an error of $e^{o(n)}$. Taking the ratio between the two quantities gives

$$\left(1 - \lambda + \frac{(1 - \zeta)\lambda u \cdot v^{-1}}{n}\right)^n e^{\lambda tn - (1 - \zeta)\lambda tu \cdot v^{-1}} \prod_{i=1}^n \frac{1}{\lambda u_i v_i^{-1} + 1 - \lambda} \quad (77)$$

Focus on the quantity $\prod_{i=1}^n (\lambda u_i v_i^{-1} + 1 - \lambda)$. We use the following claim for a bound on the gap between the arithmetic and geometric means.

Claim B.3. *For $z \in \mathbb{R}_+^n$ with $\bar{z} = n^{-1} \sum_{i=1}^n z_i$, the function $f : [0, 1] \rightarrow \mathbb{R}$ given by $f(\alpha) = \sum_{i=1}^n \log(\bar{z} + \alpha(z_i - \bar{z}))$ is concave with a maximum at $\alpha = 0$. (Indeed, the function $z \mapsto \sum_{i=1}^n \log z_i$ is concave on \mathbb{R}_+^n .) Hence,*

$$0 \leq f(0) - f(1) \leq -f'(1) = -\sum_{i=1}^n \frac{z_i - \bar{z}}{z_i} = -n + \bar{z} \sum_{i=1}^n \frac{1}{z_i}. \quad (78)$$

Exponentiating both sides gives

$$\bar{z}^n \prod_{i=1}^n z_i^{-1} \leq \exp\left(-n + \bar{z} \sum_{i=1}^n \frac{1}{z_i}\right). \quad (79)$$

Applying the above claim to the product in (77) with $z_i = \lambda u_i v_i^{-1} + 1 - \lambda$, we obtain

$$\prod_{i=1}^n \frac{1}{\lambda u_i v_i^{-1} + 1 - \lambda} \leq \left(1 - \lambda + \frac{\lambda u \cdot v^{-1}}{n}\right)^{-n} \exp\left(-n + \sum_{i=1}^n \frac{1 - \lambda + \lambda u \cdot v^{-1}/n}{1 - \lambda + \lambda u_i v_i^{-1}}\right) \quad (80)$$

Hence, the conditional probability of interest is upper bounded, up to $e^{o(n)}$, by the following expression

$$\left(1 - \lambda + \frac{(1 - \zeta)\lambda u \cdot v^{-1}}{n}\right)^n e^{\lambda tn - (1 - \zeta)\lambda tu \cdot v^{-1}} \left(1 - \lambda + \frac{\lambda u \cdot v^{-1}}{n}\right)^{-n} \exp\left(-n + \sum_{i=1}^n \frac{1 - \lambda + \lambda u \cdot v^{-1}/n}{1 - \lambda + \lambda u_i v_i^{-1}}\right). \quad (81)$$

Denote the negative logarithm of the n -th root of the quantity above by $\underline{\chi}(\lambda)$. That is,

$$\mathbb{P}(u \cdot x < (1 - \zeta)tu \cdot v^{-1} | v \cdot x < tn) \leq \inf_{\lambda > 0} e^{o(n) - n\underline{\chi}(\lambda)} = \exp\left(o(n) - n \sup_{\lambda > 0} \underline{\chi}(\lambda)\right) \quad (82)$$

for any $\lambda > 0$ with

$$\begin{aligned} \underline{\chi}(\lambda) := & -\log\left(1 - \lambda + \frac{(1 - \zeta)\lambda u \cdot v^{-1}}{n}\right) - \lambda t + \frac{(1 - \zeta)\lambda t u \cdot v^{-1}}{n} + \log\left(1 - \lambda + \frac{\lambda u \cdot v^{-1}}{n}\right) \\ & + 1 - \frac{1}{n} \sum_{i=1}^n \frac{1 - \lambda + \lambda u \cdot v^{-1}/n}{1 - \lambda + \lambda u_i v_i^{-1}}. \end{aligned} \quad (83)$$

The $o(n)$ factor is of lower order, and it suffices to show that there exists some λ such that $\underline{\chi}(\lambda) = \Theta(\zeta^2)$. For λ sufficiently small (e.g., $\lambda \leq K^{-2}/2$, recalling that $u_i, v_i \in [1/K, K]$), we may approximate the logarithm function near its zero and obtain

$$\log\left(1 - \lambda + \frac{\lambda u \cdot v^{-1}}{n}\right) \geq \frac{\lambda u \cdot v^{-1}}{n} - \lambda - \left(\frac{\lambda u \cdot v^{-1}}{n} - \lambda\right)^2. \quad (84)$$

Then the two log terms in (83) combined can be bounded below by

$$\lambda - \frac{(1 - \zeta)\lambda u \cdot v^{-1}}{n} + \frac{\lambda u \cdot v^{-1}}{n} - \lambda - \left(\frac{\lambda u \cdot v^{-1}}{n} - \lambda\right)^2 = \zeta \lambda \frac{u \cdot v^{-1}}{n} - \lambda^2 \left(\frac{u \cdot v^{-1}}{n} - 1\right)^2. \quad (85)$$

With $u_i, v_i \in [1/K, K]$, a naive lower bound is the following

$$\underline{\chi}(\lambda) \geq \zeta \lambda K^{-2} - \lambda^2 K^4 - \lambda t + (1 - \zeta)\lambda t K^{-2} + 1 - \frac{(2 - 2\lambda + K^2\lambda + K^{-2}\lambda)^2}{4(1 - \lambda + K^2\lambda)(1 - \lambda + K^{-2}\lambda)}, \quad (86)$$

where the summation at the end of (83) is bounded using Schweitzer's inequality Schweitzer (1914) for the ratio between arithmetic and harmonic means, stating

$$\frac{1}{n} \sum_{i=1}^n \frac{\bar{z}}{z_i} \leq \frac{(a + b)^2}{4ab}$$

for $z \in \mathbb{R}^n$ with bounded components $0 < a \leq z_i \leq b$. Further, we observe

$$1 - \frac{(2 - 2\lambda + K^2\lambda + K^{-2}\lambda)^2}{4(1 - \lambda + K^2\lambda)(1 - \lambda + K^{-2}\lambda)} = -\frac{(4K^2(K^{-2} - 1)^2 + (K - K^{-1})^4)}{4(1 - \lambda + K^2\lambda)(1 - \lambda + K^{-2}\lambda)} \lambda^2 \geq -3K^4 \lambda^2. \quad (87)$$

Taking $\lambda = \zeta K^{-6}/8$ in (86) yields

$$\underline{\chi}\left(\frac{1}{8}\zeta K^{-6}\right) \geq \frac{1}{16}\zeta^2 K^{-8} - \frac{1}{8}t\zeta K^{-6} \geq \Theta(\zeta^2), \quad (88)$$

hence finishing our proof for the lower tail.

The proof for the upper tail follows a similar structure. Writing

$$\begin{aligned} \mathbb{P}(u \cdot x > (1 + \zeta)tu \cdot v^{-1} | v \cdot x < tn) &= \frac{\mathbb{P}(u \cdot x > (1 + \zeta)tu \cdot v^{-1}, v \cdot x < tn)}{\mathbb{P}(v \cdot x < tn)} \\ &\leq \frac{\mathbb{P}((-\lambda u + (1 + \lambda)v) \cdot x < -(1 + \zeta)\lambda t u \cdot v^{-1} + (1 + \lambda)tn)}{\mathbb{P}(v \cdot x < tn)} \end{aligned}$$

for some $0 < \lambda < K^{-2}$ (so that $-\lambda u + (1 + \lambda)v \in \mathbb{R}_+^n$) to be determined later, Lemma B.1 and B.1 together imply that the ratio is, up to $e^{o(n)}$,

$$\left(1 + \lambda - \frac{(1 + \zeta)\lambda u \cdot v^{-1}}{n}\right)^n e^{(1 + \zeta)\lambda t u \cdot v^{-1} - \lambda t n} \prod_{i=1}^n \frac{1}{1 + \lambda - \lambda u_i v_i^{-1}} \quad (89)$$

As in the proof of lower tail bound, the product term can be bounded by

$$\prod_{i=1}^n (1 + \lambda - \lambda u_i v_i^{-1}) \leq \left(1 + \lambda - \frac{\lambda u \cdot v^{-1}}{n}\right)^{-n} \exp\left(-n + \sum_{i=1}^n \frac{1 + \lambda - \lambda u \cdot v^{-1}/n}{1 + \lambda - \lambda u_i v_i^{-1}}\right), \quad (90)$$

giving an upper bound, again up to $e^{o(n)}$, of

$$\left(1 + \lambda - \frac{(1 + \zeta)\lambda u \cdot v^{-1}}{n}\right)^n e^{(1+\zeta)\lambda t u \cdot v^{-1} - \lambda t n} \left(1 + \lambda - \frac{\lambda u \cdot v^{-1}}{n}\right)^{-n} \exp\left(-n + \sum_{i=1}^n \frac{1 + \lambda - \lambda u \cdot v^{-1}/n}{1 + \lambda - \lambda u_i v_i^{-1}}\right) \quad (91)$$

for the conditional probability of interest.

Denote the negative logarithm of the n -th root of the quantity above by $\bar{\chi}(\lambda)$. That is,

$$\mathbb{P}(u \cdot x > (1 + \zeta)t u \cdot v^{-1} | v \cdot x < t n) \leq \inf_{0 < \lambda < K^{-2}} e^{o(n) - n\bar{\chi}(\lambda)} \quad (92)$$

for any $\lambda \in (0, K^{-2})$ with

$$\begin{aligned} \bar{\chi}(\lambda) := & -\log\left(1 + \lambda - \frac{(1 + \zeta)\lambda u \cdot v^{-1}}{n}\right) + \lambda t - \frac{(1 + \zeta)\lambda t u \cdot v^{-1}}{n} + \log\left(1 + \lambda - \frac{\lambda u \cdot v^{-1}}{n}\right) \\ & + 1 - \frac{1}{n} \sum_{i=1}^n \frac{1 + \lambda - \lambda u \cdot v^{-1}/n}{1 + \lambda - \lambda u_i v_i^{-1}}. \end{aligned} \quad (93)$$

Again, it suffices to prove that for some choice of λ we have $\bar{\chi}(\lambda) = \Theta(\zeta^2)$. With similar arithmetic as in the proof for the lower tail, we observe that for λ sufficiently small (e.g., $\lambda \leq K^{-2}/2$)

$$\bar{\chi}(\lambda) \geq \zeta \lambda K^{-2} + \lambda t - (1 + \zeta)\lambda t K^2 - 4\lambda^2 K^4. \quad (94)$$

Again, taking $\lambda = \zeta K^{-6}/8$ in (94) gives the desired lower bound of $\Theta(\zeta^2)$ for $\sup_{0 < \lambda < K^{-2}} \bar{\chi}(\lambda)$ and thus finishes our proof. \square

B.2 A generalized DKW inequality for independent and nearly identically distributed random variables

Lemma B.4. *Let X_i , $i = 1, \dots, n$ be independent random variables each with (non-identical) distribution function G_i , and assume that there exists a constant $\delta > 0$ and a distribution F such that $\|G_i - F\|_\infty \leq \delta$ uniformly across all $i = 1, \dots, n$. Let \hat{G} be the empirical distribution function of $\{X_i\}_{i=1}^n$. Then*

$$\mathbb{P}(\|\hat{G} - F\|_\infty > 2\delta + \epsilon) < 4 \exp(-2n\epsilon^2/9). \quad (95)$$

Proof. Let $U_i = G_i(X_i)$ so that U_1, \dots, U_n are i.i.d. uniform on $[0, 1]$, and denote their empirical distribution function by \hat{J} . Let $Y_i = F^{-1}(U_i) = F^{-1}(G_i(X_i))$ so that Y_1, \dots, Y_n are i.i.d. each with distribution function F , and denote their empirical distribution function by \hat{F} . Notice that

$$\begin{aligned} \|\hat{G} - F\|_\infty &= \sup_{x \in \mathbb{R}} \left| n^{-1} \sum_{i=1}^n I_{(-\infty, x)}(X_i) - F(x) \right| \\ &= \sup_{x \in \mathbb{R}} \left| n^{-1} \sum_{i=1}^n I_{(-\infty, x)}(Y_i) - F(x) + n^{-1} \sum_{i=1}^n (I_{(-\infty, x)}(Y_i) - I_{(-\infty, x)}(X_i)) \right| \\ &\leq \sup_{x \in \mathbb{R}} \left| n^{-1} \sum_{i=1}^n I_{(-\infty, x)}(Y_i) - F(x) \right| + \sup_{x \in \mathbb{R}} \left(n^{-1} \sum_{i=1}^n |I_{(-\infty, x)}(Y_i) - I_{(-\infty, x)}(X_i)| \right) \\ &= \|\hat{F} - F\|_\infty + \sup_{x \in \mathbb{R}} A(x). \end{aligned}$$

From the classic result of DKW inequality Dvoretzky et al. (1956) applied to \hat{F} and F , we know that

$$\mathbb{P}(\|\hat{F} - F\|_\infty > \epsilon/3) < 2 \exp(-2n\epsilon^2/9). \quad (96)$$

For the second supremum of in the sum above, we now consider $U_i, i = 1, \dots, n$ as the underlying random variables. Each term in the summation in A contributes 1 to the sum if and only if

$$F^{-1}(U_i) = Y_i < x \leq X_i = G_i^{-1}(U_i) \quad \text{or} \quad G_i^{-1}(U_i) = X_i < x \leq Y_i = F^{-1}(U_i),$$

or alternatively

$$F(x) \wedge G_i(x) \leq U_i \leq F(x) \vee G_i(x),^{15}$$

where \wedge and \vee denote the operators of taking the minimum and maximum, respectively. Hence,

$$\begin{aligned} A(x) &= n^{-1} \sum_{i=1}^n |I_{(-\infty, x)}(Y_i) - I_{(-\infty, x)}(X_i)| \\ &= n^{-1} \sum_{i=1}^n I_{(F(x) \wedge G_i(x), F(x) \vee G_i(x))}(U_i) \\ &\leq n^{-1} \sum_{i=1}^n I_{(\bigwedge_j G_j(x) \wedge F(x), \bigvee_j G_j(x) \vee F(x))}(U_i) \\ &= \hat{J}(M(x)) - \hat{J}(m(x)), \end{aligned}$$

where M and m denote the maximum and minimum across F and $G_i, i = 1, \dots, n$, respectively. By our assumption that $\|G_i - F\|_\infty \leq \delta$ across all i , we have that

$$0 \leq M(x) - m(x) \leq 2\delta$$

for all $x \in \mathbb{R}$. Noticing that the true distribution function J of $U_i, i = 1, \dots, n$ is the identity function on $[0, 1]$, we have

$$\begin{aligned} A(x) &= \hat{J}(M(x)) - \hat{J}(m(x)) \\ &\leq \left| \hat{J}(M(x)) - J(M(x)) \right| + \left| \hat{J}(m(x)) - J(m(x)) \right| + |J(M(x)) - J(m(x))| \\ &\leq 2\|\hat{J} - J\|_\infty + 2\delta \end{aligned}$$

on \mathbb{R} uniformly. Therefore, applying DKW inequality again, we see that

$$\mathbb{P}(\sup A > 2\delta + 2\epsilon/3) \leq \mathbb{P}(\|\hat{J} - J\|_\infty > \epsilon/3) < 2 \exp(-2n\epsilon^2/9). \quad (97)$$

Combining (96) and (97) yields the desired bound in (95). \square

C Additional proofs

C.1 Proof of Corollary 6.7

In this section, we prove Corollary 6.7, which is restated below for convenience. We will assume Proposition 6.6, whose proof is deferred to Appendix C.4.

¹⁵We may safely ignore the case where the two sides are equal, as it happens with probability zero.

Corollary 6.7 (Number of stable partial matchings). *Fix any $\delta > 0$ and $c \in (0, 1/2)$. Let N_δ denote the number of stable partial matchings of size $n - \lfloor \delta n \rfloor$ satisfying the condition in Corollary A.13 (i.e., \mathcal{R}^*) in a random instance of the matching market. Then, $\mathbb{E}[N_\delta] \leq \exp(o_\delta(n))$ granted that n is sufficiently large. Further, with probability at least $1 - e^{-n^c}$, the condition \mathcal{R}^* is satisfied by all δ -truncated stable matchings.*

Proof. The last part is simply Corollary A.13.

For the first part, observe that for each $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$ with $|\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor$ and partial matching μ' between \mathcal{M}' and \mathcal{W}' ,

$$\begin{aligned} \mathbb{P}(\mu' \text{ is stable and satisfies } \mathcal{R}^*) &\leq e^{o(n)} \mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}^*}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'})] \leq \\ &e^{o(n)} \mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_1}(\mathbf{Y}_{\mathcal{W}'})] \leq e^{o_\delta(n)} \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i} \quad (98) \end{aligned}$$

by Proposition 6.6. Summing over \mathcal{M}' , \mathcal{W}' , and μ' bounds the expected number of such stable partial matchings above by

$$\begin{aligned} \mathbb{E}[N_\delta] &\leq \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M}, \mathcal{W}' \subseteq \mathcal{W} \\ |\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor}} \sum_{\substack{\mu': \mathcal{M}' \rightarrow \mathcal{W}' \\ \text{bijection}}} e^{o_\delta(n)} \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i} \\ &\stackrel{(i)}{=} \frac{1}{\lfloor \delta n \rfloor!} \sum_{\substack{\mu: \mathcal{M} \rightarrow \mathcal{W} \\ \text{bijection}}} \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M} \\ |\mathcal{M}'| = n - \lfloor \delta n \rfloor}} e^{o_\delta(n)} \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu(i)} b_{\mu(i), i} \\ &\stackrel{(ii)}{\leq} \sum_{\substack{\mu: \mathcal{M} \rightarrow \mathcal{W} \\ \text{bijection}}} \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M} \\ |\mathcal{M}'| = n - \lfloor \delta n \rfloor}} e^{o_\delta(n)} \frac{1}{n!} \cdot C^{2\lfloor \delta n \rfloor} \prod_{i \in \mathcal{M}} a_{i, \mu(i)} b_{\mu(i), i} \\ &\stackrel{(iii)}{\leq} e^{o_\delta(n)} \binom{n}{\lfloor \delta n \rfloor} \cdot \frac{1}{n!} \text{Perm}(\mathbf{A} \circ \mathbf{B}^\top) \\ &\stackrel{(iv)}{\leq} e^{o_\delta(n)}, \end{aligned}$$

where in (i) we use an alternative counting of partial matchings by counting sub-matchings of size $n - \lfloor \delta n \rfloor$ in full matchings and then deduplicate by a factor of $\lfloor \delta n \rfloor!$; in (ii) we use the boundedness assumption on the components of \mathbf{A} and \mathbf{B} ; in (iii) we merge $C^{2\lfloor \delta n \rfloor}$ into $e^{o_\delta(n)}$; and finally in (iv) we merge $\binom{n}{\lfloor \delta n \rfloor} = \exp(h(\delta)n + o(n))$ into $e^{o_\delta(n)}$ and bound the permanent term by $n^n \text{Perm}(\mathbf{M}) \leq \Theta(n!)$ using the moderate deviation property of \mathbf{M} (McCullagh, 2014, Sec. 3). \square

C.2 Proof of Lemma 6.9

In this section, we prove Lemma 6.9, which is restated below for convenience.

Lemma 6.9. *Let μ' be a partial matching of size $n - \lfloor \delta n \rfloor$ on $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$. Fix any $\zeta > 0$, and let*

$$\Omega_{\text{eig}}(\zeta) := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n : \exists t \in \mathbb{R}_+, \sum_{i=1}^n \mathbb{1} \left\{ |(\mathbf{M}\mathbf{y})_i - t| \geq \sqrt{\zeta} t \right\} \leq \sqrt{\zeta} n \right\}. \quad (27)$$

Then

$$\mathbb{E}\left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \setminus \Omega_{\text{eig}}(\Theta(\delta) + \zeta)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'})\right] \leq \exp(o_\delta(n) - \Theta(\zeta^2 n)) \cdot \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}, \quad (28)$$

with the implicit constants uniform over all \mathcal{M}' , \mathcal{W}' , and μ' .

To prepare for the proof of Lemma 6.9, let us denote the expectation in (28) by E , and express it as

$$\begin{aligned} E &= \int_0^\infty \mathbb{P}(q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \setminus \Omega_{\text{eig}}(\zeta)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) > s) ds \\ &= \int_0^1 \mathbb{P}(\exp(-\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'}) > s, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \mathcal{R} \setminus \Omega_{\text{eig}}(\zeta)) ds \\ &= \int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \mathcal{R}_2 \setminus \Omega_{\text{eig}}(\zeta), \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt \\ &= \int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_{\text{eig}}(\zeta), \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt, \quad (99) \end{aligned}$$

where $\bar{t} := t \wedge (c_2(\log n)^{1/8})$. If we can find two families of regions $\Omega_1(\zeta; s), \Omega_2(\zeta; s) \subseteq \mathbb{R}_+^n \times \mathbb{R}_+^n$ such that $\Omega_{\text{eig}}(\zeta) \supseteq \Omega_1(\Theta(\zeta); s) \cap \Omega_2(\Theta(\zeta); s)$ for all $0 < s < c_2(\log n)^{1/8}$, by union bound and relaxing the requirement that $\mathbf{X}_{\mathcal{M}'}$ (resp. $\mathbf{Y}_{\mathcal{W}'}$) is in \mathcal{R}_1 , we will obtain

$$\begin{aligned} E &\leq \int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_1(\Theta(\zeta); \bar{t}), \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt \\ &\quad + \int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_2(\Theta(\zeta); \bar{t}), \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt \\ &\leq \int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_1(\Theta(\zeta); \bar{t}), \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt \\ &\quad + \int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_2(\Theta(\zeta); \bar{t}), \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt. \end{aligned}$$

Rewriting the probabilities through conditioning and further relaxing the requirement gives

$$\begin{aligned} E &\leq \int_0^\infty \mathbb{P}((\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_1(\Theta(\zeta); \bar{t}) | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1) \\ &\quad \cdot \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt \\ &\quad + \int_0^\infty \mathbb{P}((\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_2(\Theta(\zeta); \bar{t}) | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \\ &\quad \cdot \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt. \quad (100) \end{aligned}$$

Due to the symmetry between the two integrals, it then suffices to bound one of the two integrals (e.g., the latter) by showing

$$\sup_{0 < t < c_2(\log n)^{1/8}} \mathbb{P}((\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \notin \Omega_2(\Theta(\zeta); t) | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, \mathbf{X}_{\mathcal{M}'} \in \mathcal{R}_1) \leq \exp(-\Theta(\zeta^2 n)) \quad (101)$$

and

$$\int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot ne^{-nt} dt \leq e^{o(n) + o_\delta(n)} \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}, \quad (102)$$

from which the desired upper bound immediately follows. Recognizing that

$$\int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot n e^{-nt} dt = \mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_1}(\mathbf{Y}_{\mathcal{W}'})],$$

we reduce (102) to Proposition 6.6. Our road map is to first find the desirable choices for Ω_1 and Ω_2 and establish (101), and then prove Proposition 6.6 in Appendix C.4. Note that Proposition 6.6 is in fact independent of our choice of Ω_1 and Ω_2 , but we will develop useful intermediate results to prepare for its proof.

Concretely, we consider events Ω_1 and Ω_2 as follows:

$$\Omega_1(\zeta; t) := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \max_{i \in [n]} \left| (1 - \delta) \frac{\mathbf{M}_{i,\cdot} \cdot \mathbf{y}}{t \mathbf{M}_{i,\cdot} \cdot (\mathbf{M}^\top \mathbf{x})_{\mathcal{W}'}}^{-1} - 1 \right| > \zeta \right\}, \quad (103)$$

$$\Omega_2(\zeta; t) := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \max_{j \in [n]} \left| (1 - \delta) \frac{\mathbf{M}_{\cdot,j} \cdot \mathbf{x}}{t \mathbf{M}_{\cdot,j} \cdot (\mathbf{M} \mathbf{y})_{\mathcal{M}'}}^{-1} - 1 \right| > \zeta \right\}, \quad (104)$$

where $\mathbf{M}_{i,\cdot}$ and $\mathbf{M}_{\cdot,j}$ denote the i -th row and the j -th column of \mathbf{M} , respectively; inverse is applied coordinate-wise on vectors; and \mathbf{v}_S denotes the n -dimensional vector obtained by zeroing out the i -th component v_i of $\mathbf{v} \in \mathbb{R}_+^n$ for all $i \in [n] \setminus S$ (with this operation performed after coordinate-wise inverse). We first verify the following lemma.

Lemma C.1. *There exist absolute constants $\zeta_0, \delta_0 > 0$ and $k_1, k_2 > 0$ such that for all $\zeta \in (0, \zeta_0)$, $\delta \in (0, \delta_0)$, and $t > 0$ we have*

$$\Omega_1(\zeta; t) \cap \Omega_2(\zeta; t) \subseteq \Omega_{\text{eig}}(k_1 \delta + k_2 \zeta). \quad (105)$$

Proof. Let $\mathbf{d} = \mathbf{M}^\top \mathbf{x}$ and $\mathbf{e} = \mathbf{M} \mathbf{y}$. Under the event that $(\mathbf{x}, \mathbf{y}) \in \Omega_1(\zeta; t) \cap \Omega_2(\zeta; t)$, we have

$$\begin{aligned} \frac{1}{d_j} &= \frac{1}{\mathbf{M}_{\cdot,j} \cdot \mathbf{x}} \stackrel{(i)}{\leq} \frac{1 - \delta}{(1 - \zeta) t \mathbf{M}_{\cdot,j} \cdot \mathbf{e}_{\mathcal{M}'}}^{-1} \stackrel{(ii)}{\leq} \frac{1 - \delta}{(1 - \zeta) t (1 + 2C^2 \delta) \mathbf{M}_{\cdot,j} \cdot \mathbf{e}^{-1}} \\ &\stackrel{(iii)}{\leq} \frac{1 - \delta}{(1 - \zeta) t (1 + 2C^2 \delta)} \sum_{i=1}^n m_{ij} e_i = \frac{1 - \delta}{(1 - \zeta) t (1 + 2C^2 \delta)} \sum_{i=1}^n m_{ij} \mathbf{M}_{i,\cdot} \cdot \mathbf{y} \\ &\stackrel{(iv)}{\leq} \frac{1 - \delta}{(1 - \zeta) t (1 + 2C^2 \delta)} \sum_{i=1}^n m_{ij} \left(\frac{1 + \zeta}{1 - \delta} t \mathbf{M}_{i,\cdot} \cdot \mathbf{d}_{\mathcal{W}'}}^{-1} \right) \leq \frac{1 + \zeta}{(1 - \zeta) (1 + 2C^2 \delta)} (\mathbf{M}^\top \mathbf{M} \mathbf{d}^{-1})_j, \end{aligned} \quad (106)$$

where (i) uses the definition of $\Omega_2(\zeta; t)$; (ii) uses the fact that \mathbf{M} and \mathbf{e} both have bounded ratios (at most C) among their entries and assumed $\delta < 1/2$; (iii) is due to Jensen's inequality (or equivalently, harmonic-mean-arithmetic-mean inequality); and (iv) uses the definition of $\Omega_1(\zeta; t)$.

Recall our assumption that \mathbf{M} has entries bounded on $[1/(Cn), C/n]$. It is straightforward to verify that for any vector $\mathbf{v} \in \mathbb{R}_+^n$ with $\bar{v} = \frac{1}{n} \sum_{i=1}^n v_i$, we have $\max_{i \in [n]} (\mathbf{M} \mathbf{v})_i - \bar{v} \leq \max_{i \in [n]} v_i - \frac{1}{Cn} \cdot n(\max_{i \in [n]} v_i - \bar{v}) - \bar{v} = (1 - C^{-1})(\max_{i \in [n]} v_i - \bar{v})$. In the case of $\mathbf{v} = \mathbf{d}^{-1}$, this implies that

$$\begin{aligned} (1 - C^{-1})^2 (d_{i^*}^{-1} - \bar{d}_{(H)}^{-1}) &\geq \max_{i \in [n]} (\mathbf{M}^\top \mathbf{M} \mathbf{d}^{-1})_i - \bar{d}_{(H)}^{-1} \\ &\geq (\mathbf{M}^\top \mathbf{M} \mathbf{d}^{-1})_{i^*} - \bar{d}_{(H)}^{-1} \geq (1 + 2C^2 \delta) \frac{1 - \zeta}{1 + \zeta} d_{i^*}^{-1} - \bar{d}_{(H)}^{-1}, \end{aligned} \quad (107)$$

where $i^* = \operatorname{argmin}_{i \in [n]} d_i$ and $\bar{d}_H = (n^{-1} \sum_{i=1}^n d_i^{-1})^{-1}$ is the harmonic mean of d_1, \dots, d_n . Solving (107) gives

$$\frac{d_{i^*}^{-1} - \bar{d}_{(H)}^{-1}}{\bar{d}_{(H)}^{-1}} \leq \Theta(\delta) + \frac{2\zeta}{1 - \zeta - (1 - C^{-2})^2(1 + \zeta)} \leq \Theta(\delta + \zeta) \quad (108)$$

with hidden constants independent of δ and ζ , granted that ζ is sufficiently small. Hence, for all but $\sqrt{\delta + \zeta}n$ indices $i \in [n]$, we have $1 - \Theta(\sqrt{\delta + \zeta}) \leq \frac{\bar{d}_{(H)}}{d_i} \leq 1 + \Theta(\delta + \zeta)$, implying that $(\mathbf{x}, \mathbf{y}) \in \Omega_{\text{eig}}(\Theta(\delta + \zeta))$. \square

Let $\mathbf{D} = \mathbf{M}^\top \mathbf{X}_{\mathcal{M}'}$ and $\mathbf{E} = \mathbf{M} \mathbf{Y}_{\mathcal{W}'}$. Note that \mathbf{D} and \mathbf{E} both have bounded ratios among their components due to the bounded ratio assumption on \mathbf{M} , and in addition $\|\mathbf{D}\|_1 = \|\mathbf{X}_{\mathcal{M}'}\|_1$ and $\|\mathbf{E}\|_1 = \|\mathbf{Y}_{\mathcal{W}'}\|_1$. By Lemma B.2, whenever $t \leq c_2(\log n)^{1/8}$, we have for each column $\mathbf{M}_{\cdot j}$ of \mathbf{M} , $j = 1, \dots, n$,

$$\mathbb{P} \left(\left| (1 - \delta) \frac{\mathbf{M}_{\cdot j} \cdot \mathbf{X}_{\mathcal{M}'}}{t \mathbf{M}_{\cdot j} \cdot \mathbf{E}_{\mathcal{M}'}} - 1 \right| > \zeta \mid \mathbf{X}_{\mathcal{M}'} \cdot \mathbf{E} < t, \|\mathbf{E}\|_1 \geq c_1 \log n \right) \leq \exp(-\Theta(n\zeta^2)), \quad (109)$$

where we note that the effective dimension of $\mathbf{X}_{\mathcal{M}'}$ is $n - \lfloor \delta n \rfloor$ instead of n . By a union bound over $j \in [n]$, this gives

$$\mathbb{P} \left(\max_{j \in [n]} \left| (1 - \delta) \frac{\mathbf{M}_{\cdot j} \cdot \mathbf{X}_{\mathcal{M}'}}{t \mathbf{M}_{\cdot j} \cdot \mathbf{E}_{\mathcal{M}'}} - 1 \right| > \zeta, \|\mathbf{E}\|_1 \geq c_1 \log n \mid \mathbf{X}_{\mathcal{M}'} \cdot \mathbf{E} < t \right) \leq \exp(-\Theta(n\zeta^2)), \quad (110)$$

which is simply (101).

C.3 Proof of Corollary 6.10

We now prove Corollary 6.10, restated below.

Corollary 6.10. *For $\delta > 0$ sufficiently small, there exists a choice of $\zeta = \zeta(\delta) > 0$ such that $\zeta \rightarrow 0$ as $\delta \rightarrow 0$ and that*

$$\mathbb{P}(\exists \mu \text{ stable}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \Omega_{\text{eig}}(\zeta)) \lesssim e^{-n^c} \quad (29)$$

asymptotically as $n \rightarrow \infty$.

Proof. Summing over all partial matchings with size $n - \lfloor \delta n \rfloor$ gives

$$\begin{aligned} & \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M}, \mathcal{W}' \subseteq \mathcal{W} \\ |\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor}} \sum_{\substack{\mu': \mathcal{M}' \rightarrow \mathcal{W}' \\ \text{bijection}}} \mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}(\mu'), \mathbf{Y}_{\mathcal{W}'}(\mu')) \cdot \mathbb{1}_{\mathcal{R} \setminus \Omega_{\text{eig}}(\zeta)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'})] \\ & \leq \exp(o_\delta(n) - \Theta(\zeta^2 n)) \cdot \frac{(\delta n)!}{n!} \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M}, \mathcal{W}' \subseteq \mathcal{W} \\ |\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor}} \sum_{\substack{\mu': \mathcal{M}' \rightarrow \mathcal{W}' \\ \text{bijection}}} \prod_{i \in \mathcal{M}'} (nm_{i, \mu'(i)}). \end{aligned} \quad (111)$$

To bound the summation, notice that

$$\begin{aligned}
\sum_{\substack{\mathcal{M}' \subseteq \mathcal{M}, \mathcal{W}' \subseteq \mathcal{W} \\ |\mathcal{M}'| = |\mathcal{W}'| = n - \lfloor \delta n \rfloor}} \sum_{\substack{\mu': \mathcal{M}' \rightarrow \mathcal{W}' \\ \text{bijection}}} \prod_{i \in \mathcal{M}'} (nm_{i, \mu'(i)}) &= \frac{1}{(\lfloor \delta n \rfloor)!} \sum_{\substack{\mu: \mathcal{M} \rightarrow \mathcal{W} \\ \text{bijection}}} \sum_{\substack{\mathcal{M}' \subseteq \mathcal{M} \\ |\mathcal{M}'| = n - \lfloor \delta n \rfloor}} \prod_{i \in \mathcal{M}'} (nm_{i, \mu'(i)}) \\
&\leq \frac{1}{(\lfloor \delta n \rfloor)!} \sum_{\substack{\mu: \mathcal{M} \rightarrow \mathcal{W} \\ \text{bijection}}} \binom{n}{\lfloor \delta n \rfloor} C^{\lfloor \delta n \rfloor} \prod_{i \in \mathcal{M}} (nm_{i, \mu(i)}) \\
&= \frac{e^{o_\delta(n)}}{(\delta n)!} n^n \text{Perm}(\mathbf{M}). \tag{112}
\end{aligned}$$

Under the assumption that the bistochastic matrix \mathbf{M} is of moderate deviation (cf. (McCullagh, 2014, Section 3)), we know that $n^n \text{Perm}(\mathbf{M}) = O(n!)$. Hence, the quantity in (111) is bounded by $\exp(o_\delta(n) - \Theta(\zeta^2 n))$. Invoking Lemma 6.4 finishes the proof. \square

C.4 Proof of Proposition 6.6

In this section, we present the proof of Proposition 6.6, restated below.

Proposition 6.6. *For a fixed a partial matching μ' on \mathcal{M}' and \mathcal{W}' of size $n - \lfloor \delta n \rfloor$,*

$$\mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_1}(\mathbf{Y}_{\mathcal{W}'})] \leq e^{o(n) + o_\delta(n)} \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}. \tag{26}$$

Denote the target expectation by E and express it as an integral of tail probability

$$E = \int_0^\infty \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1) \cdot n e^{-nt} dt, \tag{113}$$

where $\bar{t} = t \wedge (c_2(\log n)^{1/8})$. It suffices to upper bound probabilities of the form $\mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < \bar{t}, \mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1)$ for all $t \in (0, c_2(\log n)^{1/8})$. We will go one step further and prove a stronger result by relaxing the $\mathbf{Y}_{\mathcal{W}'} \in \mathcal{R}_1$ condition, which will eventually translate to a bound on $\mathbb{E}[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'})]$.

Lemma C.2. *There exists a positive constant β such that, for any $t \in (0, c_2(\log n)^{1/8})$,*

$$\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn}, \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn} \mid \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \leq 0.1 \tag{114}$$

for n sufficiently large.

Proof. Let p denote the target probability. We have

$$\begin{aligned}
p &= \mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn} \mid \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn}, \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \\
&\quad \cdot \mathbb{P}\left(\|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn} \mid \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \\
&\leq \mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn} \mid \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn}, \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \\
&= \frac{\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn}, \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t \mid \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn}\right)}{\mathbb{P}\left(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t \mid \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn}\right)} \\
&\leq \frac{\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn} \mid \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn}\right)}{\mathbb{P}\left(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t \mid \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \beta\sqrt{tn}\right)} \\
&\leq \frac{\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn}\right)}{\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \sqrt{tn}/(C\beta)\right)} = \mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn} \mid \|\mathbf{X}_{\mathcal{M}'}\|_1 \leq (C\beta)^{-1}\sqrt{tn}\right), \tag{115}
\end{aligned}$$

where the last inequality follows from the independence between \mathbf{X} and \mathbf{Y} and the fact that $n\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} \leq C \|\mathbf{X}_{\mathcal{M}'}\|_1 \|\mathbf{Y}_{\mathcal{W}'}\|_1$. By choosing $\beta = (2C)^{-1/2}$, the upper bound becomes

$$\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \beta\sqrt{tn} \mid \|\mathbf{X}_{\mathcal{M}'}\|_1 \leq 2\beta\sqrt{tn}\right).$$

A direct invocation of Lemma B.2 implies an $\exp(-\Theta(n))$ upper bound for this probability. \square

Lemma C.3. *There exists a positive constant γ such that, for any $t \in (0, c_2(\log n)^{1/8})$,*

$$\mathbb{P}\left(\|\mathbf{Y}_{\mathcal{W}'}\|_1 \geq \gamma n(\log n)^{-7/8} \mid \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \leq 0.1 \quad (116)$$

Proof. To free ourselves from always carrying the notation for the partial matching, let us observe that, once we relinquish the condition on the bistochasticity of \mathbf{M} , it becomes irrelevant that μ' is a partial matching between $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$ (instead of a complete one between \mathcal{M} and \mathcal{W}), since the difference in the market size $|\mathcal{M}| = |\mathcal{W}| = n$ and $|\mathcal{M}'| = |\mathcal{W}'| = n - \delta n$ does not affect the final asymptotics in the Lemma. Hence, it suffices to establish a version of (116) with $\mathbf{X}_{\mathcal{M}'}$ and $\mathbf{Y}_{\mathcal{W}'}$ replaced by non-truncated value vectors \mathbf{X} and \mathbf{Y} in a complete (instead of partial) matching μ , as long as we do not rely on bistochasticity of \mathbf{M} .

In the simplified notation, let $\mathbf{Z} = \mathbf{a} \circ \mathbf{X}$ and $\mathbf{W} = \mathbf{b} \circ \mathbf{Y}$ with $\mathbf{a} = (a_{i,\mu(i)})_{i \in [n]}$ and $\mathbf{b} = (b_{j,\mu^{-1}(j)})_{j \in [n]}$, so that $\mathbf{Z}, \mathbf{W} \sim \text{Exp}(1)^n$ and are independent. Moreover, let $R = \|\mathbf{Z}\|_1$ and $\mathbf{U} = R^{-1}\mathbf{Z}$ so that, as is well known, $R \sim \Gamma(n, 1)$, $\mathbf{U} \sim \text{Unif}(\Delta_{n-1})$, and R and \mathbf{U} are independent. Similarly, let $S = \|\mathbf{W}\|_1$, and $\mathbf{V} = S^{-1}\mathbf{W}$. Then

$$\mathbf{X}^\top \mathbf{M} \mathbf{Y} = \mathbf{Z}^\top \text{diag}(\mathbf{a}^{-1}) \mathbf{M} \text{diag}(\mathbf{b}^{-1}) \mathbf{W} = R \mathbf{S} \mathbf{U}^\top \tilde{\mathbf{M}} \mathbf{V},$$

where $\tilde{\mathbf{M}} := \text{diag}(\mathbf{a}^{-1}) \mathbf{M} \text{diag}(\mathbf{b}^{-1})$ again has entries bounded on $[1/(C^2n), C^2/n]$. Since $\|\mathbf{Y}\|_1 = \Theta(S)$, it suffice to find a positive constant γ such that

$$\mathbb{P}\left(S \geq \gamma n(\log n)^{-7/8} \mid R \mathbf{S} \mathbf{U}^\top \tilde{\mathbf{M}} \mathbf{V} < t\right) < 0.1 \quad (117)$$

for all n sufficiently large and $t \in (0, c_2(\log n)^{1/8})$. Note that $1/(C^2n) \leq \mathbf{U}^\top \tilde{\mathbf{M}} \mathbf{V} \leq C^2/n$ a.s. By conditional on all possible values of $\mathbf{U}^\top \tilde{\mathbf{M}} \mathbf{V}$, it suffices to show that for all $t' \in (0, c_2 C^2(\log n)^{1/8})$

$$\mathbb{P}\left(S \geq \gamma n(\log n)^{-7/8} \mid R S < t'n\right) < 0.1 \quad (118)$$

asymptotically.

First, we write $\mathbb{P}(S \geq \gamma n(\log n)^{-7/8}, R S < t'n)$ as

$$\mathbb{P}(S \geq \gamma n(\log n)^{-7/8}, R S < t'n) = \int_{\gamma n(\log n)^{-7/8}}^{\infty} G(t'n/s) g(s) ds,$$

where $g(x) = \frac{x^{n-1}e^{-x}}{(n-1)!}$ is the probability density function of $\Gamma(n, 1)$ and G is the corresponding CDF. Since $t'n/s \ll n$, we may use Lemma B.1 to upper bound $G(t'n/s)$, giving

$$\begin{aligned} \mathbb{P}(S \geq \gamma n(\log n)^{-7/8}, R S < t'n) &\leq \int_{\gamma n(\log n)^{-7/8}}^{\infty} \left(\frac{t'e}{s}\right)^n \frac{s^{n-1}e^{-s}}{(n-1)!} ds \\ &= \frac{(t'e)^n}{(n-1)!} \int_{\gamma n(\log n)^{-7/8}}^{\infty} \frac{e^{-s}}{s} ds \\ &\leq \frac{(t'e)^n}{(n-1)!} e^{-\gamma n(\log n)^{-7/8}}. \end{aligned} \quad (119)$$

Next, we lower bound $\mathbb{P}(RS < t'n)$ by

$$\mathbb{P}\left(n^{1/2} \leq S \leq n^{2/3}, RS < t'n\right) = \int_{n^{1/2}}^{n^{2/3}} G(t'n/s)g(s)ds.$$

Note that $t'n/s \leq O(n^{2/3})$ for all $s \in [n^{1/2}, n^{2/3}]$. Using the lower bound in Lemma B.1, we have

$$\begin{aligned} \mathbb{P}\left(n^{1/2} \leq S \leq n^{2/3}, RS < t'n\right) &\geq e^{-O(n^{2/3})} \int_{n^{1/2}}^{n^{2/3}} \left(\frac{t'e}{s}\right)^n \frac{s^{n-1}e^{-s}}{(n-1)!} ds \\ &= e^{-O(n^{2/3})} \frac{(t'e)^n}{(n-1)!} \int_{n^{1/2}}^{n^{2/3}} \frac{e^{-s}}{s} ds \\ &\geq \frac{(t'e)^n}{(n-1)!} n^{-2/3} e^{-O(n^{2/3})}. \end{aligned} \quad (120)$$

Comparing (119) with (120) establishes (118) and hence finishes the proof. \square

Remark C.4. Note that this lemma should be treated only as a technical result about the typical behavior of $\mathbf{X}_{\mathcal{M}'}$ and $\mathbf{Y}_{\mathcal{W}'}$ when $q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) = \exp(-n\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'})$ is large, and should not be confused with any attempt to bound the number of stable (partial) matchings with women's total values in a certain range. For example, one might hope to replace $\gamma n(\log n)^{-7/8}$ with $\gamma n(\log n)^{-1}$ in the proof to conclude that stable matchings with $\|\mathbf{Y}_\delta\|_1 \in [n^{1/2}, n^{2/3}]$ are over $e^{n^{2/3}}$ times more common than those with $\|\mathbf{Y}_\delta\|_1 \geq \Omega(n(\log n)^{-1})$. This, however, is generally not true as we know in the classic case with uniformly random preferences. To see why this fact is not contradictory to our proof, recall from Proposition 6.3 (see Section 4 and Appendix A) that $q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'})$ is only a good approximation to $p_{\mu'}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'})$ when, among other conditions, $\|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \Theta(n(\log n)^{-7/8})$; even then, the approximation is only valid up to an $e^{o(n)}$ factor. As the ratio between (119) and (120) is only $e^{o(n)}$, the quality of approximation is insufficient for us to rule out the possibility for a (partial) stable matching to have $\|\mathbf{Y}_{\mathcal{W}'}\|_1 \geq \Theta(n(\log n)^{-1})$: the man-optimal stable matching obtained from the man-proposing deferred acceptance algorithm will be such an example.

Corollary C.5. There exists a positive constant γ' such that, for any $t \in (0, c_2(\log n)^{1/8})$,

$$\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \gamma't(\log n)^{7/8}, \|\mathbf{Y}_{\mathcal{W}'}\|_1 \geq \beta\sqrt{tn} \mid \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \leq 0.2 \quad (121)$$

for n sufficiently large, where β is the constant appearing in Lemma C.2.

Proof. Note that $\|\mathbf{M} \mathbf{Y}_{\mathcal{W}'}\|_1 = \|\mathbf{Y}_{\mathcal{W}'}\|_1 \gtrsim \sqrt{tn} \gg t$ for $t \lesssim (\log n)^{1/8}$. For any \mathbf{y} supported on coordinates indexed by \mathcal{W}' with $t \ll \|\mathbf{y}\|_1 \leq \gamma n(\log n)^{-7/8}$, Lemma B.2 implies

$$\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq 0.9 \frac{t}{n - \lfloor \delta n \rfloor} \left\| (\mathbf{M} \mathbf{y})_{\mathcal{M}'}^{-1} \right\|_1 \mid \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, \mathbf{Y}_{\mathcal{W}'} = \mathbf{y}\right) \leq e^{-\Theta(n)}. \quad (122)$$

Plugging in $\left\| (\mathbf{M} \mathbf{y})_{\mathcal{M}'}^{-1} \right\|_1 \geq (n - \lfloor \delta n \rfloor) \frac{n}{C\|\mathbf{y}\|_1} \geq \frac{n - \lfloor \delta n \rfloor}{C\gamma} (\log n)^{7/8}$ gives

$$\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq 0.9(C\gamma)^{-1} t(\log n)^{7/8} \mid \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, \mathbf{Y}_{\mathcal{W}'} = \mathbf{y}\right) \leq e^{-\Theta(n)}. \quad (123)$$

Marginalizing over all relevant values of \mathbf{y} implies

$$\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \leq \gamma't(\log n)^{7/8}, \beta\sqrt{tn} \leq \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \gamma n(\log n)^{-7/8} \mid \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \leq e^{-\Theta(n)} \quad (124)$$

with $\gamma' = 0.9(C\gamma)^{-1}$. Combining this with Lemma C.3 completes the proof. \square

Corollary C.6. For any $t \in (0, c_2(\log n)^{1/8})$ and n sufficiently large,

$$\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \wedge \|\mathbf{Y}_{\mathcal{W}'}\|_1 \geq \gamma' t (\log n)^{7/8}, \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right) \geq \frac{1}{2} \mathbb{P}\left(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t\right). \quad (125)$$

Proof. This is a direct consequence of Lemma C.2 and Corollary C.5 \square

We are now ready to state the proof of Proposition 6.6.

Proof of Proposition 6.6. Define events

$$\begin{aligned} A_1(t) &: \|\mathbf{X}_{\mathcal{M}'}\|_1 \geq \gamma' t (\log n)^{1/8}, & A_2(t) &: \|\mathbf{Y}_{\mathcal{W}'}\|_1 \geq \gamma' t (\log n)^{1/8}, \\ B_1(t) &: (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \Omega_1(\zeta; t), & B_2(t) &: (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \Omega_2(\zeta; t), \end{aligned}$$

where Ω_1 and Ω_2 are defined in Appendix C.2, and ζ is to be specified later. We have

$$\begin{aligned} \frac{1}{2} \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t) &\leq \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t), A_2(t)) \\ &\leq \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t), A_2(t), B_1(t), B_2(t)) \\ &\quad + \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t), B_1(t)^c) + \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_2(t), B_2(t)^c) \\ &\leq \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t), A_2(t), B_1(t), B_2(t)) \\ &\quad + \mathbb{P}(B_1(t)^c | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t)) \cdot \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t) \\ &\quad + \mathbb{P}(B_2(t)^c | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_2(t)) \cdot \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t). \end{aligned} \quad (126)$$

For any fixed $\delta > 0$ and $\zeta = o_\delta(1)$, $\mathbb{P}(B_1(t)^c | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t)) \rightarrow 0$ by Lemma B.2; in particular, we may assume that $\mathbb{P}(B_1(t)^c | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t))$ (and by symmetry $\mathbb{P}(B_2(t)^c | \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_2(t))$) is at most $1/8$. Thus,

$$\mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t) \leq 4 \mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t), A_2(t), B_1(t), B_2(t)). \quad (127)$$

By Lemma C.1, $B_1(t)$ and $B_2(t)$ together imply that $n \mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} = (1 + o_\delta(1)) \|\mathbf{X}_{\mathcal{M}'}\|_1 \|\mathbf{Y}_{\mathcal{W}'}\|_1$. Further, along with the events $\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t$, $A_1(t)$, and $A_2(t)$, they imply $t (\log n)^{1/8} \lesssim \|\mathbf{Y}_{\mathcal{W}'}\|_1 \lesssim n (\log n)^{-1/8}$. Hence,

$$\begin{aligned} &\mathbb{P}(\mathbf{X}_{\mathcal{M}'}^\top \mathbf{M} \mathbf{Y}_{\mathcal{W}'} < t, A_1(t), A_2(t), B_1(t), B_2(t)) \\ &\leq \mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \frac{nt}{1 + o_{\delta, \zeta}(1)}, t (\log n)^{1/8} \leq \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \Theta(n (\log n)^{-1/8})\right) \\ &\leq \mathbb{E}\left[\mathbb{P}\left(\|\mathbf{X}_{\mathcal{M}'}\|_1 \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \frac{nt}{1 + o_{\delta, \zeta}(1)}, t (\log n)^{1/8} \leq \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \Theta(n (\log n)^{-1/8}) \middle| \|\mathbf{Y}_{\mathcal{W}'}\|_1\right)\right] \\ &\leq e^{o_\delta(n)} \left(\frac{ent}{n - \lfloor \delta n \rfloor}\right)^{n - \lfloor \delta n \rfloor} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} \\ &\quad \cdot \mathbb{E}\left[\|\mathbf{Y}_{\mathcal{W}'}\|_1^{-n + \lfloor \delta n \rfloor}; t (\log n)^{1/8} \leq \|\mathbf{Y}_{\mathcal{W}'}\|_1 \leq \Theta(n (\log n)^{-1/8})\right]. \end{aligned} \quad (128)$$

It is straightforward, albeit a bit tedious, to explicitly bound the expectation term in (128) by

$$(\Theta(\log n) - \log t) e^{n - \lfloor \delta n \rfloor} \prod_{i \in \mathcal{M}'} b_{\mu'(i), i},$$

again using Lemma B.1. Carrying out the integral over t in (113) finishes the proof. \square

C.5 Proof of Lemma 7.1

In this section, we present the proof of Lemma 7.1, restated below.

Lemma 7.1. *Fix any $\delta, \zeta > 0$. Let μ' be a partial matching of size $n - \lfloor \delta n \rfloor$ on $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$. For any $\epsilon > 0$, consider*

$$\Omega_{\text{emp}}(\epsilon) := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \exists \lambda \in \mathbb{R}_+, \|\hat{\mathcal{F}}(\mathbf{x}) - F_\lambda\|_\infty \leq \epsilon + \Theta(\delta + \sqrt{\zeta}) \right\}. \quad (30)$$

Then

$$\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega_{\text{eig}}(\zeta) \setminus \Omega_{\text{emp}}(\epsilon)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \right] \leq \exp(o_\delta(n) - \Theta(\epsilon^2 n)) \cdot \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}, \quad (31)$$

where again the implicit constants are uniform over all $\mathcal{M}', \mathcal{W}'$, and μ' .

Proof. In light of Proposition 6.6, it suffices to show that for all $\mathbf{y} \in \text{Proj}_y(\mathcal{R} \cap \Omega_{\text{eig}}(\zeta))$, we have

$$\frac{\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega_{\text{eig}}(\zeta) \setminus \Omega_{\text{emp}}(\epsilon)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \mid \mathbf{Y}_{\mathcal{W}'} = \mathbf{y} \right]}{\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \mid \mathbf{Y}_{\mathcal{W}'} = \mathbf{y} \right]} \leq \exp(-\Theta(\epsilon^2 n)). \quad (129)$$

It then follows that

$$\begin{aligned} & \mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega_{\text{eig}}(\zeta) \setminus \Omega_{\text{emp}}(\epsilon)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \cap \Omega_{\text{eig}}(\zeta) \setminus \Omega_{\text{emp}}(\epsilon)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \mid \mathbf{Y}_{\mathcal{W}'} \right] \right] \\ &\leq \mathbb{E} \left[\exp(-\Theta(\epsilon^2 n)) \cdot \mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \mid \mathbf{Y}_{\mathcal{W}'} \right] \cdot \mathbb{1}_{\text{Proj}_y(\mathcal{R} \cap \Omega_{\text{eig}}(\zeta))}(\mathbf{Y}_{\mathcal{W}'}) \right] \\ &\leq \exp(-\Theta(\epsilon^2 n)) \cdot \mathbb{E} \left[\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \mid \mathbf{Y}_{\mathcal{W}'} \right] \cdot \mathbb{1}_{\mathcal{R}_1}(\mathbf{Y}_{\mathcal{W}'}) \right] \\ &\leq \exp(-\Theta(\epsilon^2 n)) \cdot \mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_2}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R}_1}(\mathbf{Y}_{\mathcal{W}'}) \right], \end{aligned} \quad (130)$$

and Proposition 6.6 immediately implies the desired bound.

To show (129), notice that the quotient in the left-hand side is simply

$$\mathbb{P}_{\mathbf{X} \sim \otimes_{i=1}^n \text{Exp}(a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i)} \left((\mathbf{X}_{\mathcal{M}'}, \mathbf{y}) \in \mathcal{R} \cap \Omega_{\text{eig}}(\zeta) \setminus \Omega_{\text{emp}}(\epsilon) \right). \quad (131)$$

Recall that $(\mathbf{X}_{\mathcal{M}'})_i = X_i$ for $i \in \mathcal{M}'$ and $(\mathbf{X}_{\mathcal{M}'})_i = 0$ for $i \notin \mathcal{M}'$. For any $\mathbf{y} \in \text{Proj}_y(\mathcal{R} \cap \Omega_{\text{eig}}(\zeta))$, there must exist $\hat{y} \in \mathbb{R}_+$ such that for all but at most $\sqrt{\zeta}n$ indices $i \in [n]$ we have $|\mathbf{M}\mathbf{y}_i - \hat{y}| \leq \sqrt{\zeta}\hat{y}$. In other words, under the distribution $\mathbf{X} \sim \otimes_{i=1}^n \text{Exp}(a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i)$, for all but at most $(\delta + \sqrt{\zeta})n$ indices $i \in [n]$, we have $n\hat{y}X_i \sim \text{Exp}(\lambda_i)$ for some

$$\lambda_i = \frac{a_{i, \mu'(i)}}{n\hat{y}} + \frac{(\mathbf{M}\mathbf{y})_i}{\hat{y}} = 1 + \Theta(\sqrt{\zeta}) + \Theta(1/\log n),$$

where we used the fact that $n\hat{y} \geq \Theta(\|\mathbf{y}\|_1) \geq \Theta(\log n)$ as implied by $\mathbf{y} \in \text{Proj}_y(\mathcal{R} \cap \Omega_{\text{eig}}(\zeta))$. The generalized Dvoretzky–Kiefer–Wolfowitz (DKW) inequality (see Lemma B.4) for independent and nearly-identically distributed random variables implies that the probability (131) is upper bounded by

$$\mathbb{P}_{\mathbf{X} \sim \otimes_{i=1}^n \text{Exp}(a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i)} \left(\|\hat{\mathcal{F}}(\mathbf{x}) - F_{n\hat{y}}\|_\infty > \epsilon + \Theta(\delta + \sqrt{\zeta}) \right) \leq \exp(-\Theta(\epsilon^2 n)), \quad (132)$$

which finishes the proof. \square

C.6 Proof of Theorem 7.4

Heuristically, we would expect the rank R_i for a man to be proportional to his value $X_i(\mu)$. We will see below that this is approximately the case when $x_i \ll 1$. There are, however, going to be some x_i of constant order, making it hard for us to say anything exact about R_i . But as we will soon see, for all but a $o(1)$ fraction of the n men, we will indeed have $x_i = o(1)$. As we are concerned with the empirical distribution, such small fraction becomes negligible in the limit and can be simply ignored. This heuristics is formalized in the next Lemma.

Lemma C.7. *Fix any $\delta > 0$. Let μ' be a partial matching of size $n - \lfloor \delta n \rfloor$ on $\mathcal{M}' \subseteq \mathcal{M}$ and $\mathcal{W}' \subseteq \mathcal{W}$. For any $0 < \xi < \rho < 1$, consider $\Omega_{\text{tail}}(\xi, \rho)$ defined as*

$$\left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^n \mathbb{1} \left\{ n x_i (\mathbf{M}\mathbf{y})_i \notin (F^{-1}(\xi/2)/2, F^{-1}(1-\xi/2)) \right\} \leq \lfloor \delta n \rfloor + \rho(n - \lfloor \delta n \rfloor) \right\}. \quad (133)$$

Then

$$\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \setminus \Omega_{\text{tail}}(\xi, \rho)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \right] \leq \exp(o_\delta(n) - \Theta(D(\rho \parallel \xi)n)) \cdot \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}, \quad (134)$$

where $D(q \parallel p)$ denotes the KL divergence from $\text{Bern}(p)$ to $\text{Bern}(q)$.

Proof. The proof entirely mirrors that of Lemma 7.1. It suffices to show that for all $\mathbf{y} \in \text{Proj}_{\mathbf{y}}(\mathcal{R})$ we have

$$\frac{\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \cdot \mathbb{1}_{\mathcal{R} \setminus \Omega_{\text{tail}}(\xi, \rho)}(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \mid \mathbf{Y}_{\mathcal{W}'} = \mathbf{y} \right]}{\mathbb{E} \left[q(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \mid \mathbf{Y}_{\mathcal{W}'} = \mathbf{y} \right]} \leq \exp(-\Theta(D(\rho \parallel \xi)n)). \quad (135)$$

The quotient is simply

$$\mathbb{P}_{\mathbf{X} \sim \prod_{i=1}^n \text{Exp}(a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i)} \left((\mathbf{X}_{\mathcal{M}'}, \mathbf{y}) \in \mathcal{R} \setminus \Omega_{\text{tail}}(\xi, \rho) \right) \leq \mathbb{P} \left((\mathbf{X}_{\mathcal{M}'}, \mathbf{y}) \notin \Omega_{\text{tail}}(\xi, \rho) \right), \quad (136)$$

where we will have $\mathbf{X} \sim \prod_{i=1}^n \text{Exp}(a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i)$ for the rest of this proof. Note that under this specified distribution, $\left((a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i) X_i \right)_{i \in \mathcal{M}'}$ are $n - \lfloor \delta n \rfloor$ i.i.d. samples from $\text{Exp}(1)$, each falling outside the interval $(F^{-1}(\xi/2), F^{-1}(1 - \xi/2))$ with probability precisely ξ . Hence,

$$\begin{aligned} & \mathbb{P} \left(\sum_{i \in \mathcal{M}'} \mathbb{1} \left\{ X_i (a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i) \notin (F^{-1}(\xi/2), F^{-1}(1 - \xi/2)) \right\} \leq \rho(n - \lfloor \delta n \rfloor) \right) \\ &= \mathbb{P}_{Z \sim \text{Bin}(n - \lfloor \delta n \rfloor, \xi)} (Z > \rho(n - \lfloor \delta n \rfloor)) \leq \exp(-D(\rho \parallel \xi)(n - \lfloor \delta n \rfloor)) \end{aligned} \quad (137)$$

by the Hoeffding bound for binomial distribution. Since $n(\mathbf{M}\mathbf{y})_i \leq a_{i, \mu'(i)} + n(\mathbf{M}\mathbf{y})_i \leq 2n(\mathbf{M}\mathbf{y})_i$ across all $i \in \mathcal{M}'$ for $\mathbf{y} \in \mathcal{R}_1$ and n sufficiently large, the probability (137) upper bounds $\mathbb{P}((\mathbf{X}_{\mathcal{M}'}, \mathbf{y}) \notin \Omega_{\text{tail}}(\xi, \rho))$. This establishes (135) and concludes the proof. \square

By fixing some small δ, ρ and choosing ξ sufficiently small, we can make $D(\rho \parallel \xi)$ arbitrarily large and obtain the following Corollary.

Corollary C.8. *For any $0 < \delta, \rho < 1/2$, there exists a choice of $\xi > 0$ such that*

$$\mathbb{P}(\exists \mu \in \mathcal{S}, (\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \tilde{\Omega}_{\text{tail}}(\xi, \rho)) \lesssim e^{-n^c} \quad (138)$$

asymptotically as $n \rightarrow \infty$, where $\tilde{\Omega}_{\text{tail}}(\xi, \rho)$ is defined as

$$\left\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^n : \sum_{i=1}^n \mathbb{1} \left\{ x_i \notin \left(F^{-1}(\xi/2) \frac{(\log n)^{7/8}}{C^2 \bar{c}_1 n}, 2F^{-1}(1-\xi/2) \frac{C^2}{\bar{c}_1 \log n} \right) \right\} \leq (\delta + \rho)n \right\}. \quad (139)$$

That is, with high probability, no stable matchings μ have more than $\delta + \rho$ fraction of the men's post-truncation values outside an interval $(\Theta(n^{-1}(\log n)^{7/8}), \Theta(1/\log n))$.

Proof. Observe that $\mathcal{R} \setminus \tilde{\Omega}_{\text{tail}}(\xi, \rho) \subseteq \mathcal{R} \setminus \Omega_{\text{tail}}(\xi, \rho)$ by our definition of \mathcal{R} and the boundedness assumption on the entries of \mathbf{M} . Again, invoking Lemma 6.4 using inequality (134) in Lemma C.7 yields the $\Theta(e^{-n^c})$ upper bound on $\mathbb{P}((\mathbf{X}_\delta(\mu), \mathbf{Y}_\delta(\mu)) \notin \tilde{\Omega}_{\text{tail}}(\xi, \rho))$. \square

Remark C.9. Recall that in a market with uniform preferences, the man-optimal (and woman-pessimal) stable matching realizes an average rank of $\Theta(\log n)$ for men and $\Theta(n/\log n)$ for women. Under the heuristics that values multiplied by n roughly correspond to ranks (which we will formalize below), Lemma C.7 nicely matches our expectation that even in the most extreme cases, few individuals will strike a rank better (smaller) than $\Theta((\log n)^{7/8})$ or worse (larger) than $\Theta(n/\log n)$. The lower bound can be refined to $\Theta(\log n / \log \log n)$ with a more careful analysis.

Now let us consider a specific partial matching μ' of size $n - \lfloor \delta n \rfloor$ between \mathcal{M}' and \mathcal{W}' and condition on μ' being stable with value vectors $(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) = (\mathbf{x}, \mathbf{y}) \in \tilde{\Omega}_{\text{tail}}(\xi, \rho)$. That is, there exists a subset $\bar{\mathcal{M}}' \subseteq \mathcal{M}'$ with $|\bar{\mathcal{M}}'| \geq (1 - \delta - \rho)n$ such that $\Theta(n^{-1}(\log n)^{7/8}) \leq (X_{\mathcal{M}'})_i \leq \Theta(1/\log n)$ for all $i \in \bar{\mathcal{M}}'$. By symmetry, we may further assume that there exists $\bar{\mathcal{W}}' \subseteq \mathcal{W}'$ with $|\bar{\mathcal{W}}'| \geq (1 - \delta - \rho)n$ such that $\Theta(n^{-1}(\log n)^{7/8}) \leq (Y_{\mathcal{W}'})_j \leq \Theta(1/\log n)$ for all $j \in \bar{\mathcal{W}}'$. We want to show that, for $i \in \bar{\mathcal{M}}'$, the *pre-truncation* rank R_i of man m_i (i.e., over the entire market, including the $\lfloor \delta n \rfloor$ women outside \mathcal{M}') is well characterized by his value $X_{i, u'(i)}$ in the matching, up to some proper scaling. From now on, we will consider some $i \in \bar{\mathcal{M}}'$ with value $X_{i, u'(i)} = x_i$, and write

$$R_i = 1 + \sum_{j \neq \mu'(i)} \mathbb{1}_{[0, x_i]}(X_{ij}). \quad (140)$$

The condition that μ' is stable requires $(X_{ij}, Y_{ji}) \notin [0, x_i] \times [0, y_j]$ for all $j \in \mathcal{W}' \setminus \{\mu'(i)\}$. Thus, for all $j \in \mathcal{W}' \setminus \{\mu'(i)\}$,

$$\begin{aligned} \mathbb{P}(X_{ij} \leq x_i | \mu' \text{ stable}, (\mathbf{X}_{\mathcal{M}'})_i = x_i, (\mathbf{Y}_{\mathcal{W}'})_j = y_j) &= \frac{\mathbb{P}(X_{ij} \leq x_i, Y_{ji} > y_j)}{1 - \mathbb{P}(X_{ij} \leq x_i, Y_{ji} \leq y_j)} \\ &= \frac{(1 - e^{-a_{ij}x_i})e^{-b_{ji}y_j}}{1 - (1 - e^{-a_{ij}x_i})(1 - e^{-b_{ji}y_j})}, \end{aligned} \quad (141)$$

and for all $j \in \mathcal{W} \setminus \mathcal{W}'$ (so $(\mathbf{Y}_{\mathcal{W}'})_j = 0$),

$$\mathbb{P}(X_{ij} \leq x_i | \mu' \text{ stable}, (\mathbf{X}_{\mathcal{M}'})_i = x_i) = 1 - e^{-a_{ij}x_i}. \quad (142)$$

Define

$$p_{ij} = \begin{cases} 1 & \text{when } j = \mu'(i), \\ \frac{(1 - e^{-a_{ij}x_i})e^{-b_{ji}y_j}}{1 - (1 - e^{-a_{ij}x_i})(1 - e^{-b_{ji}y_j})} & \text{when } j \in \mathcal{W}' \setminus \{\mu'(i)\}, \\ 1 - e^{-a_{ij}x_i} & \text{when } j \in \mathcal{W} \setminus \mathcal{W}', \end{cases}$$

and $I_{ij} \sim \text{Bern}(p_{ij})$ independently for $i \in [n]$ so that $R_i = \sum_{j=1}^n I_{ij}$ conditional on $(\mathbf{X}_{\mathcal{M}'})_i = X_i = x_i$. Note that for any $j \neq \mu'(i)$ and $j \notin \bar{\mathcal{W}}'$, we always have

$$p_{ij} \leq 1 - e^{-a_{ij}x_i} \leq a_{ij}x_i \quad (143)$$

and

$$\begin{aligned} p_{ij} &\geq \frac{(1 - e^{-a_{ij}x_i})e^{-b_{ji}y_j}}{1 - (1 - e^{-a_{ij}x_i})(1 - e^{-b_{ji}y_j})} \geq (1 - e^{-a_{ij}x_i})e^{-b_{ji}y_j} \\ &\geq e^{-\Theta(\frac{1}{\log n})} \left(1 - \Theta\left(\frac{1}{\log n}\right)\right) a_{ij}x_i = (1 - o(1))a_{ij}x_i. \end{aligned} \quad (144)$$

For $j \in \bar{\mathcal{W}}' \setminus \{\mu'(i)\}$, p_{ij} admits the same upper bound (143) and the trivial lower bound of zero. Hence, conditional on

$$\mu' \text{ stable and } (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) = (\mathbf{x}, \mathbf{y}) \in \tilde{\Omega}_{\text{tail}}(\xi, \rho) \cap \tilde{\Omega}_{\text{emp}}(\epsilon) \cap \mathcal{R} \quad (\dagger)$$

for any (fixed) $\xi, \rho, \epsilon > 0$, we have the stochastic dominance

$$1 + \sum_{j \notin \bar{\mathcal{W}}' \cup \{\mu'(i)\}} \underline{I}_{ij} \leq R_i \leq 1 + \sum_{j \neq \mu'(i)} \bar{I}_{ij}, \quad (145)$$

where $\underline{I}_{ij} \sim \text{Bern}((1 - o(1))a_{ij}x_i)$ and $\bar{I}_{ij} \sim \text{Bern}(a_{ij}x_i)$. Since $i \in \bar{\mathcal{M}}'$ by our assumption and thus $\Theta((\log n)^{7/8}/n) \leq x_i \leq \Theta(1/\log n)$, the expectation of R_i/x_i can be upper bounded by

$$\mathbb{E} \left[\frac{R_i}{x_i} \middle| (\dagger) \right] \leq \frac{1}{x_i} + \sum_{j \neq \mu'(i)} a_{ij} = (1 + o(1)) \sum_{j=1}^n a_{ij} \quad (146)$$

and lower bounded by

$$\mathbb{E} \left[\frac{R_i}{x_i} \middle| (\dagger) \right] \geq (1 - o(1)) \sum_{j \in \bar{\mathcal{W}}'} a_{ij} = (1 - \Theta(\delta)) \sum_{j=1}^n a_{ij}. \quad (147)$$

Similarly, we may bound the variance of R_i/x_i by

$$\text{Var} \left(\frac{R_i}{x_i} \middle| (\dagger) \right) \leq \sum_{j \neq \mu'(i)} a_{ij}(1 - a_{ij}x_i) \leq \sum_{j=1}^n a_{ij}. \quad (148)$$

Hence, we have

$$1 - \Theta(\delta) \leq \mathbb{E} \left[\frac{R_i}{x_i \sum_{j=1}^n a_{ij}} \middle| (\dagger) \right] \leq 1 + o(1) \quad \text{and} \quad \text{Var} \left(\frac{R_i}{x_i \sum_{j=1}^n a_{ij}} \middle| (\dagger) \right) \leq \Theta(n^{-1}), \quad (149)$$

with these quantities conditionally independent for all $i \in \bar{\mathcal{M}}'$ and the hidden constants depending only on C , implying concentration of R_i around $x_i \sum_{j=1}^n a_{ij}$ in the following sense.

Proposition C.10. *Conditional on (\dagger) , for any fixed $\theta > 0$ and $\delta, \rho, \gamma \in (0, 1/2)$, we have*

$$\mathbb{P} \left(\sum_{i=1}^n \mathbb{1}_{(\theta + \Theta(\delta), \infty)} \left(\left| \frac{R_i}{x_i \sum_{j=1}^n a_{ij}} - 1 \right| \geq (\delta + \rho + \gamma)n \middle| (\dagger) \right) \lesssim \mathbb{P}_{N \sim \text{Poi}(\Theta(\theta^{-2}))}(N \geq \gamma n) \leq e^{-\Theta(\gamma n)}. \quad (150)$$

Proof. By Chebyshev's inequality and (149), $\mathbb{P} \left(\left| \frac{R_i}{x_i \sum_{j=1}^n a_{ij}} - \mathbb{E} \left[\frac{R_i}{x_i \sum_{j=1}^n a_{ij}} \middle| (\dagger) \right] \right| \geq \theta \middle| (\dagger) \right) \leq \Theta((n\theta^2)^{-1})$ for all $i \in \bar{\mathcal{M}}'$. Hence, by conditional independence of the ranks, $\sum_{i \in \bar{\mathcal{M}}'} \mathbb{1}_{(\theta + \Theta(\delta), \infty)} \left(\left| \frac{R_i}{x_i \sum_{j=1}^n a_{ij}} - 1 \right| \geq (\delta + \rho + \gamma)n \right)$ is stochastically dominated by $\text{Bin}(n, \Theta((n\theta^2)^{-1}))$, which converges to $\text{Poi}(\Theta(\theta^{-2}))$ in total variance. The Proposition follows from the well known tail bound for $N \sim \text{Poi}(\lambda)$ that $\mathbb{P}(N \geq \lambda + t) \leq \exp\left(-\frac{t^2}{2(\lambda+t)}\right)$, which implies $\mathbb{P}_{N \sim \text{Poi}(\Theta(\theta^{-2}))}(N \geq \gamma n) \leq \exp\left(-\frac{(\gamma n - \Theta(\theta^{-2}))^2}{2\gamma n}\right) \lesssim \exp(-\gamma n/2)$. \square

Proof of Theorem 7.4. Note that in Corollary C.8 and Corollary C.8, ρ can be chosen arbitrarily small once δ is fixed. In particular, we may always guarantee $\rho \leq \delta$. Similarly, we may assume $\theta \leq \delta$ in Proposition C.10. Thus,

$$\mathbb{P}\left(\sum_{i=1}^n \mathbb{1}_{(\Theta(\delta), \infty)}\left(\left|\frac{R_i}{x_i w_i} - 1\right|\right) \geq (2\delta + \gamma)n \middle| (\dagger)\right) \leq e^{-\Theta(\gamma n)}, \quad (151)$$

where $w_i = \sum_{j=1}^n a_{ij}$ is the fitness value of man m_i . Marginalizing over all pairs of relevant value vectors $(\mathbf{x}, \mathbf{y}) \in \tilde{\Omega}_{\text{tail}}(\xi, \rho) \cap \tilde{\Omega}_{\text{emp}}(\epsilon) \cap \mathcal{R}$ in the condition (\dagger) , we obtain

$$\mathbb{P}\left(\mathcal{E}_{\text{ratio}}(\delta, \gamma) \middle| (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \tilde{\Omega}_{\text{tail}}(\xi, \rho) \cap \tilde{\Omega}_{\text{emp}}(\epsilon) \cap \mathcal{R}\right) \leq e^{-\Theta(\gamma n)}, \quad (152)$$

where $\mathcal{E}_{\text{ratio}}(\delta, \gamma)$ denotes the undesirable event that $\sum_{i=1}^n \mathbb{1}_{(\Theta(\delta), \infty)}\left(\left|\frac{R_i}{(\mathbf{X}_{\mathcal{M}'})_i w_i} - 1\right|\right) \geq (2\delta + \gamma)n$ for the partial matching μ' . By Proposition 6.6,

$$\begin{aligned} \mathbb{P}((\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \tilde{\Omega}_{\text{tail}}(\xi, \rho) \cap \Omega_{\text{emp}}(\epsilon) \cap \mathcal{R}) &\leq \mathbb{P}(\mu' \text{ stable}, (\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \mathcal{R}) \\ &\leq e^{o(n)+o_\delta(n)} \frac{(\delta n)!}{n!} \prod_{i \in \mathcal{M}'} a_{i, \mu'(i)} b_{\mu'(i), i}. \end{aligned} \quad (153)$$

By choosing $\gamma = \gamma(\delta) = o_\delta(1)$ sufficiently large (relative to δ) and following a similar computation as in Lemma 6.4 and Corollary 6.7, we can ensure that with probability $1 - \Theta(e^{-n^\epsilon})$ there exists no stable partial matching μ' where both $\mathcal{E}_{\text{ratio}}(\delta, \gamma)$ and $(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \tilde{\Omega}_{\text{emp}}(\epsilon_0(\delta)) \cap \mathcal{R}$ happen, where the function ϵ_0 is defined in the proof of 7.2. Notice that by repeated uses of the triangle inequality,

$$(\mathbf{X}_{\mathcal{M}'}, \mathbf{Y}_{\mathcal{W}'}) \in \tilde{\Omega}_{\text{emp}}(\epsilon_0(\delta)), \mathcal{E}_{\text{ratio}}(\delta, \gamma)^c \implies \|\hat{\mathcal{F}}(\mathbf{w}^{-1} \circ \mathbf{R}(\mu')) - F_\lambda\|_\infty \leq \Theta(\delta) + \gamma(\delta) + \epsilon_0(\delta) = o_\delta(1) \quad (154)$$

for the choice of $\lambda = \|\mathbf{Y}_{\mathcal{W}'}\|_1$. Combining this with (35) and Proposition 6.1, we conclude that with probability $1 - \Theta(e^{-n^\epsilon})$, all stable matchings $\mu \in \mathcal{S}$ induce δ -truncated stable partial matchings μ_δ with $\|\hat{\mathcal{F}}(\mathbf{w}^{-1} \circ \mathbf{R}(\mu_\delta)) - F_{\lambda(\mu)}\|_\infty = o_\delta(1)$, where $\lambda(\mu) = \|\mathbf{Y}_\delta(\mu)\|_1$. The δ -truncation affects the distance by at most δ , which can be absorbed into the $o_\delta(1)$ upper bound. Thus, by choosing δ sufficiently small relative to any fixed $\epsilon > 0$, we complete our proof of Theorem 7.4. \square

C.7 Proofs of Theorem 3.4 and Corollary 3.5

Theorem 3.4. *Assume $\alpha < n^{-\eta}$ for some constant $\eta > 1/2$. Then, as $n \rightarrow \infty$,*

$$\max_{\mu \in \mathcal{S}_\alpha} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\mathbf{X}(\mu)) - F_\lambda\|_\infty \xrightarrow{P} 0 \quad \text{and} \quad \max_{\mu \in \mathcal{S}_\alpha} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\phi^{-1} \circ \mathbf{R}(\mu)) - F_\lambda\|_\infty \xrightarrow{P} 0. \quad (6)$$

Proof. There are $\binom{n}{\alpha n}^2 = \exp(2h_b(\alpha)n + O(\ln n))$ sub-markets of size at least $(1-\alpha)n$, where $h_b(p) = -p \log p - (1-p) \log(1-p)$ is the binary entropy function. Under Assumption 2.2 for the whole market, each of such sub-markets also satisfies Assumption 2.2. Fix any $\epsilon > 0$. By Theorem 7.2, each of such sub-markets can only contain a stable matching with men's empirical distribution of value deviating from the family of exponential distributions by at least $\epsilon/2$ in Kolmogorov-Smirnov distance with probability at most $1 - \exp(-n^c)$ for any fixed $c \in (0, 1/2)$. Whenever $\alpha < n^{-\eta}$ for some $\eta > 1/2$, we have $h_b(\alpha)n < n^{1-\eta}$. Choosing $c \in (1-\eta, 1/2)$ and applying union bound over all relevant sub-markets gives the first part of (6), since the additional α fraction of the market affects the empirical distribution by at most $\alpha \rightarrow 0$. The second part follows analogously from Theorem 7.4. \square

Corollary 3.5. *Consider a market consisting of $n - k$ men and n women, where $k < n^\beta$ for some constant $\beta < 1/2$. Assume that the contiguity condition holds as in Assumption 2.2. Then, as $n \rightarrow \infty$,*

$$\max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\mathbf{X}(\mu)) - F_\lambda\|_\infty \xrightarrow{p} 0 \quad \text{and} \quad \max_{\mu \in \mathcal{S}} \inf_{\lambda \in \mathbb{R}_+} \|\hat{\mathcal{F}}(\phi^{-1} \circ \mathbf{R}(\mu)) - F_\lambda\|_\infty \xrightarrow{p} 0. \quad (7)$$

Proof. The proof is entirely the same as the proof of Theorem 3.4. The union bound covers all sub-markets of size $n - k$, that is, consisting all the men and a subset of the women. The rest is the same. \square