Stable Matching with Uncertain Linear Preferences

Haris Aziz[1,2], Péter Biró[3], Serge Gaspers[2,1], Ronald de Haan[4], Nicholas Mattei[1,2], and Baharak Rastegari[5]

[1] Data61/CSIRO, Sydney, Australia. haris.aziz@data61.csiro.au, nicholas.mattei@data61.csiro.au
[2] University of New South Wales, Sydney, Australia. sergeg@cse.unsw.edu.au
[3] Hungarian Academy of Sciences, Institute of Economics. peter.biro@krtk.mta.hu
[4] Technische Universität Wien, Vienna, Austria. dehaan@ac.tuwien.ac.at
[5] School of Computing Science, University of Glasgow, Glasgow, UK. baharak.rastegari@glasgow.ac.uk

Abstract. We consider the two-sided stable matching setting in which there may be uncertainty about the agents' preferences due to limited information or communication. We consider three models of uncertainty: (1) lottery model — in which for each agent, there is a probability distribution over linear preferences, (2) compact indifference model — for each agent, a weak preference order is specified and each linear order compatible with the weak order is equally likely and (3) joint probability model — there is a lottery over preference profiles. For each of the models, we study the computational complexity of computing the stability probability of a given matching as well as finding a matching with the highest probability of being stable. We also examine more restricted problems such as deciding whether a certainly stable matching exists. We find a rich complexity landscape for these problems, indicating that the form uncertainty takes is significant.

1 Introduction

We consider a *Stable Marriage problem (SM)* in which there is a set of men and a set of women. Each man has a linear order over the women, and each woman has a linear order over the men. For the purpose of this paper we assume that the preference lists are complete, i.e., each agent finds each member of the opposite side acceptable.[6] In the stable marriage problem the goal is to compute a *stable matching*; a matching where no two agents prefer to be matched to each other rather than be matched to their current partners. Unlike most of the literature on stable matching problems [5, 10, 12], we assume that men and women may have uncertainty in their preferences which can be captured by

[6] We note that the complexity of all problems that we study are the same for complete and incomplete lists, where non-listed agents are deemed unacceptable—see Proposition 2 in the full version of the paper [1].

various probabilistic uncertainty models. We focus on *linear models* in which each possible deterministic preference profile is a set of linear orders.

Uncertainty in preferences could arise for a number of reasons both practical and epistemological. For example, an agent could express a weak order because the agent did not invest enough time or effort to differentiate between potential matches and therefore one could assume that each linear extension of the weak order is equally likely; this maps to our *compact indifference model*. In many real applications the ties are broken randomly with lotteries, e.g., in the school choice programs in New York and Boston as well as in centralized college admissions in Ireland. However, a central planner may also choose a matching that is optimal in some sense, without breaking the ties in the preference list. For instance, in Scotland they used to compute the maximum size (weakly) stable matching to allocate residents to hospitals [10]. We argue that another natural solution could be the matching which has the highest probability of being stable after conducting a lottery. Alternatively, there may be a cost associated with eliciting preferences from the agents, so a central planner may want to only obtain and provide a recommendation based on a subset of the complete orders [2].

As another example, imagine a group of interns are admitted to a company and allocated to different projects based on their preferences and the preferences of the project leaders. Suppose that after three months the interns can switch projects if the project leaders agree; though the company would prefer not to have swaps if possible. However, both the interns and the project leaders can have better information about each other after the three months, and the assignment should also be stable with regard to the refined preferences. This example motivates our lottery and joint probability models. In the *lottery model*, the agents have *independent* probabilities over possible linear orders (e.g. each project leader has a probability distribution on possible refined rankings over the interns independently from each other). In the *joint probability model*, the probability distribution is over possible preference profiles and can thus accommodate the possibility that the preferences of the agents are refined in a correlated way (e.g. if an intern performs well in the first three months then she is likely to be highly ranked by all project leaders). Uncertainty in preferences has already been studied in voting [7] and for cooperative games [9]. Ehlers and Massó [3] considers many-to-one matching markets under a Bayesian setting. Similarly, in auction theory, it is standard to examine Bayesian settings in which there is a probability distribution over the types of agents.

To illustrate the problem we describe a simple example with four agents. We write $b \succ_a c$ to say that agent a prefers b to c and assume the lottery model.

Example 1. We have two men m_1 and m_2 and two women w_1 and w_2. Each agent assigns a probability to each strict preference ordering as follows. (i) $p(w_1 \succ_{m_1} w_2) = 0.4$ and $p(w_2 \succ_{m_1} w_1) = 0.6$ (ii) $p(w_1 \succ_{m_2} w_2) = 0.0$ and $p(w_2 \succ_{m_2} w_1) = 1.0$ (iii) $p(m_1 \succ_{w_1} m_2) = 1.0$ and $p(m_2 \succ_{w_1} m_1) = 0.0$ (iv) $p(m_1 \succ_{w_2} m_2) = 0.8$ and $p(m_2 \succ_{w_2} m_1) = 0.2$. This setting admits two matchings that are stable with positive probability: $\mu_1 = \{(m_1, w_1), (m_2, w_2)\}$ and $\mu_2 = \{(m_1, w_2), (m_2, w_1)\}$. Notice that if each agent submits the preference list that s/he finds most likely to be true, then the setting admits a unique stable matching that is μ_2. The probability of μ_2 being stable, however, is 0.48 whereas the probability of μ_1 being stable is 0.52.

1.1 Uncertainty Models

We consider three different uncertainty models:

- **Lottery Model**: For each agent, we are given a probability distribution over strict preference lists.
- **Compact Indifference Model**: Each agent reports a single weak preference list that allows for ties. Each complete linear order extension of this weak order is assumed to be equally likely.
- **Joint Probability Model**: A probability distribution over preference profiles is specified.

Note that for the Lottery Model and the Joint Probability Model the representation of the input preferences can be exponentially large. However, in settings where similar models of uncertainty are used, including resident matching [2] and voting [7], a limited amount of uncertainty (i.e. small supports) is commonly expected and observed in real world data. Consequently, we consider special cases when the uncertainty is bounded in certain natural ways including the existence of only a small number of uncertain preferences and/or uncertainty on only one side of the market.

Observe that the compact indifference model can be represented as a lottery model. This is a special case of the lottery model in which each agent expresses a weak order over the candidates (similar to the SMT setting [5, 10]). However, the lottery model representation can be exponentially larger than the compact indifference model; for an agent that is indifferent among n agents on the other side of the market, there are $n!$ possible linearly ordered preferences.

1.2 Computational Problems

Given a stable marriage setting where agents have uncertain preferences, various natural computational problems arise. Let *stability probability* denote the probability that a matching is stable. We then consider the following two natural problems for each of our uncertainty models.

- MatchingWithHighestStabilityProbability: Given uncertain preferences of the agents, compute a matching with highest stability probability.
- StabilityProbability: Given a matching and uncertain preferences of the agents, what is the stability probability of the matching?

We also consider two specific problems that are simpler than StabilityProbability: (1) IsStabilityProbabilityNon-Zero — For a given matching, is its stability probability non-zero? (2) IsStabilityProbabilityOne — For a given matching, is its stability probability one?

We additionally consider problems connected to, and more restricted than, MatchingWithHighestStabilityProbability: (1) ExistsCertainlyStableMatching — Does there exist a matching that has stability

Problems	Lottery Model	Compact Indifference	Joint Probability
StabilityProbability	#P-complete	?	in P
	in P for all three models if 1 side is certain		
IsStabilityProbabilityNon-Zero	NP-complete	in P	in P
IsStabilityProbabilityOne	in P	in P	in P
ExistsPossiblyStableMatching	in P	in P	in P
ExistsCertainlyStableMatching	in P	in P	NP-complete
MatchingWithHighestStabilityProb	?	NP-hard	NP-hard
	in P for all models if 1 side is certain and there is $O(1)$ number of uncertain agents		

Table 1. Summary of results.

probability one? (2) ExistsPossiblyStableMatching — Does there exist a matching that has non-zero stability probability?

Note that ExistsPossiblyStableMatching is straightforward to answer for any of the three uncertainty models we consider here, since there exists a stable matching for each deterministic preference profile that is a possible realization of the uncertain preferences.

1.3 Results

Table 1 summarizes our main findings. Note that the complexity of each problem is considered with respect to the input size, and that under the lottery and joint probability models the input size could be exponential in n, namely $O(n! \cdot 2n)$ for the lottery model and $O((n!)^{2n})$ for the joint probability model, where n is the number of agents on either side of the market. The complete version of the sketched or missing proofs can be found in the full version of the paper [1].

We point out that StabilityProbability is #P-complete for the lottery model even when each agent has at most two possible preferences, but in P if one side has certain preferences. Additionally, we show that IsStabilityProbabilityNon-Zero is in P for the lottery model if each agent has at most two possible preferences. Note that StabilityProbability is open for the compact indifference model when both sides may be uncertain, and we also do not know the complexity of MatchingWithHighestStabilityProbility in the lottery model, except when only a constant number of agents are uncertain on the same side of the market.

2 Preliminaries

In the Stable Marriage problem, there are two sets of agents. Let M denote a set of n men and W a set of n women. We use the term *agents* when making statements that apply to both men and women, and the term *candidates* to refer to the agents on the opposite side of the market to that of an agent under consideration. Each agent has a linearly ordered preference over the candidates. An agent may be uncertain about his/her linear preference ordering. Let L denote the *uncertain preference profile* for all agents. We denote by $I = (M, W, L)$ an instance of a *Stable Marriage problem with Uncertain Linear Preferences (SMULP)*.

We say that a given uncertainty model is *independent* if any uncertain preference profile L under the model can be written as a product of uncertain preferences L_a for all agents a, where all L_a's are independent. Note that the lottery and the compact indifference models are both independent, but the joint probability model is not.

A *matching* μ is a pairing of men and women such that each man is paired with at most one woman and vice versa; defining a list of (man, woman) pairs (m, w). We use $\mu(m)$ to denote the woman w that is matched to m and $\mu(w)$ to denote the match for w. Given linearly ordered preferences, a matching is *stable* if there is no pair (m, w) not in μ where m prefers w to his current partner in μ, i.e., $w \succ_m \mu(m)$, and vice versa. If such a pair exists, it constitutes a *blocking pair*; as the pair would prefer to defect and match with each other rather than stay with their partner in μ. Given an instance of SMULP, a matching is *certainly stable* if it is stable with probability 1.

The following extensions of SM will come in handy in proving our results. The *Stable Marriage problem with Partially ordered lists (SMP)* is an extension of SM in which agents' preferences are partial orders over the candidates. The *Stable Marriage problem with Ties (SMT)* is a special case of SMP in which incomparability is transitive and is interpreted as indifference. Therefore, in SMT each agent partitions the candidates into different ties (equivalence classes), is indifferent between the candidates in the same tie, and has strict preference ordering over the ties. In some practical settings some agents may find some candidates unacceptable and prefer to remain unmatched than to get matched to the unacceptable ones. *SMP with Incomplete lists (SMPI)* and *SMT with Incomplete lists (SMTI)* capture these scenarios where each agent's partially ordered list contains only his/her acceptable candidates. A matching is *super-stable* in an instance of SMPI if it is stable w.r.t. all linear extensions of the partially ordered lists.

We define the *certainly preferred* relation $\succ_a^{\text{cert}}$ for agent a. We write $b \succ_a^{\text{cert}} c$ if and only if agent a prefers b over c with probability 1. Based on the certainly preferred relation, we can define a dominance relation D: $D_m(w) = \{w\} \cup \{w' : w' \succ_m^{\text{cert}} w\}$; $D_w(m) = \{m\} \cup \{m' : m' \succ_w^{\text{cert}} m\}$. Based on the notion of the dominance relation, we present a useful characterization of certainly stable matchings for independent uncertainty models.

Lemma 1. *A matching μ is certainly stable for an independent uncertainty model if and only if for each pair $\{m, w\}$, $\mu(m) \in D_m(w)$ or $\mu(w) \in D_w(m)$.*

We point out that certainly preferred relation can be computed in polynomial time for all three models studied in this paper.

Certainly stable matchings are closely related to the notion of super-stable matchings [4, 8]. In fact we can define a certainly stable matching using a terminology similar to that of super-stability. Given a matching μ and an unmatched pair $\{m, w\}$, we say that $\{m, w\}$ *very weakly blocks (blocks)* μ if $\mu(m) \not\succ_m^{\text{cert}} w$ and $\mu(w) \not\succ_w^{\text{cert}} m$. The next claim then follows from Lemma 1.

Proposition 1. *A matching μ is certainly stable for an independent uncertainty model if and only if it admits no very weakly blocking pair.*

3 General Results

In this section, we present some general results that apply to multiple uncertainty models. First we show that ExistsCertainlyStableMatching can be solved in polynomial time for any independent uncertainty model including lottery and compact indifference. Second, when the number of uncertain agents is constant and one side of the market is certain, then we can solve MatchingWithHighestStabilityProbability efficiently for each of the linear models.

3.1 An Algorithm for the Lottery and Compact Indifference Models

Theorem 1. *For any independent uncertainty model in which the certainly preferred relation is transitive and can be computed in polynomial time,* ExistsCertainlyStableMatching *can be solved in polynomial time.*

Proof sketch. We prove this by reducing ExistsCertainlyStableMatching to the problem of deciding whether an instance of SMP admits a super-stable matching. The latter problem can be solved in polynomial time using algorithm SUPER-SMP in [11].

Let $I = (M, W, L)$ be an instance of ExistsCertainlyStableMatching under an independent uncertainty model, assuming that the certainly preferred relation is transitive and can be computed in polynomial time. We construct an instance $I' = (M, W, p)$ of SMP, in polynomial time, as follows. The set of men and women are unchanged. To create the partial preference ordering p_a for each agent a we do the following. W.l.o.g., assume that a is a man m. For every pair of women w_1 and w_2 (i) if $w_1 \succ_m^{\text{cert}} w_2$ then $(w_1, w_2) \in p_m$, denoting that m (strictly) prefers w_1 to w_2 in I', (ii) if $w_2 \succ_m^{\text{cert}} w_1$ then $(w_2, w_1) \in p_m$, denoting that m (strictly) prefers w_2 to w_1 in I'. We claim, and show, that I' admits a super-stable matching iff I has a certainly stable matching. □

3.2 An Algorithm for a Constant Number of Uncertain Agents

Theorem 2. *When the number of uncertain agents is constant and one side of the market is certain then* MatchingWithHighestStabilityProbability *is polynomial-time solvable for each of the linear models.*

Proof sketch. Let $I = (M, W, L)$ be an instance of MatchingWithHighest-StabilityProbability and let $X \subseteq M$ be the set of uncertain agents with $|X| = k$ for a constant k. We consider all the possible matchings between X and W, where their total number is $K = n(n-1)\dots(n-k)$. Let μ_i be such a matching for $i \in \{1 \dots K\}$. The main idea of the proof is to show that there exist an extension of μ_i to $M \cup W$ that has stability probability at least as high as any other extension of μ_i. In this way we will need to compute this probability for only a polynomial number of matchings in n, which we can do efficiently for each model when one side has certain preferences—see Theorems 3, 8 and 10, and select the one with the highest probability. □

4 Lottery Model

In this section we focus on the lottery model.

Theorem 3. *For the lottery model, if one side has certain preferences,* StabilityProbability *is polynomial-time solvable.*

Proof sketch. W.l.o.g. assume that men are certain. The stability probability of a given matching μ is equal to the probability that none of the possible blocking pairs form. The probability of one blocking pair $\{m, w\}$ forming is equal to the probability that w prefers m to $\mu(w)$ given m also prefers w to $\mu(m)$. □

Theorem 4. *For the lottery model,* IsStabilityProbabilityOne *can be solved in linear time.*

Theorem 5. *For the lottery model,* IsStabilityProbabilityNon-Zero *is polynomial-time solvable when each agent has at most two possible preference orderings.*

Proof sketch. We reduce the problem to 2SAT, that is polynomial-time solvable. For each agent and for both possible preference orderings for that agent, we introduce a variable, and we construct a 2CNF formula that encodes (1) that for each agent exactly one preference ordering is selected, and (2) that the selected preference orderings cause the given matching to be stable. Satisfying assignments then correspond to witnesses for non-zero stability probability. □

Lemma 2. *In polynomial time, we can transform any 2CNF formula φ over the variables $x_1, \dots, x_n$ to a 2CNF formula φ' over the variables $x_1, \dots, x_n, y_1, \dots, y_n$ such that (1) φ and φ' have the same number of satisfying assignments, (2) each clause of φ' contains exactly one variable x_i and one variable y_j, and (3) for any two variables, there is at most one clause in φ' that contains these variables.*

Theorem 6. *For the lottery model,* StabilityProbability *is #P-complete, even when each agent has at most two possible preferences.*

Proof. We show how to count the number of satisfying assignments for a 2CNF formula using the problem StabilityProbability for the lottery model where each agent has two possible preferences. Since this problem is #P-hard, we get #P-hardness also for StabilityProbability.

Let φ be a 2CNF formula over the variables $x_1, \ldots, x_n$. We firstly transform φ to a 2CNF formula φ' over the variables $x_1, \ldots, x_n, y_1, \ldots, y_n$ as specified by Lemma 2. We then construct an instance of StabilityProbability. The sets of agents that we consider are $\{x_1, \ldots, x_n, a_1, \ldots, a_n\}$ and $\{y_1, \ldots, y_n, b_1, \ldots, b_n\}$. The matching that we consider matches x_i to b_i and matches y_i to a_i, for each $1 \leq i \leq n$. This is depicted below. Each agent b_i has only a single possible preference, namely one where they prefer x_i over all other agents. Similarly, each agent a_i has a single possible preference where they prefer y_i over all other agents. In other words, the agents a_i and b_i are perfectly happy with the given matching. The agents x_i and y_i each have two possible preferences, that are each chosen

with probability $\frac{1}{2}$. These two possible preferences are associated with setting these variables to true or false, respectively. We describe how these preferences are constructed for the agents x_i. The construction for the preferences of the agents y_i is then entirely analogous.

Take an arbitrary agent x_i. We show how to construct the two possible preferences for agent x_i, which we denote by p_{x_i} and $p_{\neg x_i}$. Both of these possible preferences are based on the following partial ranking: $b_1 > b_2 > \cdots > b_n$, and we add some of the agents $y_1, \ldots, y_n$ to the top of this partial ranking, and the remaining agents to the bottom of this partial ranking.

To the ranking p_{x_i} we add exactly those agents y_j to the top where φ' contains a clause $(\neg x_i \vee y_j)$ or a clause $(\neg x_i \vee \neg y_j)$. All remaining agents we add to the bottom. Similarly, to the ranking $p_{\neg x_i}$ we add exactly those agents y_j to the top where φ' contains a clause $(x_i \vee y_j)$ or a clause $(x_i \vee \neg y_j)$. The rankings p_{y_i} and $p_{\neg y_i}$, for the agents y_i, are constructed entirely similarly.

Now consider a truth assignment $\alpha : \{x_1, \ldots, x_n, y_1, \ldots, y_n\} \to \{0, 1\}$, and consider the corresponding choice of preferences for the agents $x_1, \ldots, x_n, y_1, \ldots, y_n$, where for each agent x_i the preference p_{x_i} is chosen if and only if $\alpha(x_i) = 1$, and for each agent y_i the preference p_{y_i} is chosen if and only if $\alpha(y_i) = 1$. Then α satisfies φ' if and only if the corresponding choice of preferences leads to the matching being stable. Since each combination of preferences is equally likely to occur, and there are 2^{2n} many combinations of preferences, the probability that the given matching is stable is exactly $q = \frac{s}{2^{2n}}$, where s is the number of satisfying truth assignments for φ. Therefore, given q, s can be obtained by computing $s = q2^{2n}$. $\square$

If each agent is allowed to have three possible preferences, then even the following problem is NP-complete. The statement can be proved via a reduction from Exact Cover by 3-Sets (X3C).

Theorem 7. *For the lottery model,* IsStabilityProbabilityNon-Zero *is NP-complete.*

We obtain the first corollary from Theorem 7 and the second from [13, Proposition 8] and Theorem 7.

Corollary 1. *For the lottery model, unless P = NP, there exists no polynomial-time algorithm for approximating* StabilityProbability *of a given matching.*

Corollary 2. *For the lottery model, unless NP = RP, there is no FPRAS for* StabilityProbability*.*

5 Compact Indifference Model

The compact indifference model is equivalent to assuming that we are given an instance of SMT and each linear order over candidates (each possible preference ordering) is achieved by breaking ties independently at random with uniform probabilities. It is easy to show that IsStabilityProbablityNonZero, IsStabilityProbablityOne, and ExistsCertainlyStableMatching are all in P. The corresponding claims and the proof can be found in [1].

We do not yet know the complexity of computing the stability probability of a given matching under the compact indifference model, but this problem can be shown to be in P if one side has certain preferences.

Theorem 8. *In the compact indifference model, if one side has certain preferences,* StabilityProbability *is polynomial-time solvable.*

Proof. Assume, w.l.o.g., that men have certain preferences. The following procedure gives us the stability probability of any given matching μ. (1) For each uncertain woman w identify those men with whom she can potentially form a blocking pair. That is, those m such that $w \succ_m \mu(m)$ and w is indifferent between m and her partner in μ. Assume there are k of such men. The probability of w not forming a blocking pair with any men is then $\frac{1}{k+1}$. (2) Multiply the probabilities from step 1. □

We next show that MatchingWithHighestStabilityProbability is NP-hard. For an instance I of SMT and matching μ, let $p(\mu, I)$ denote the probability of μ being stable, and let $p_S(I) = max\{p(\mu, I) | \mu \text{ is a matching in } I\}$, that is the maximum probability of a matching being stable. A matching μ is said to be weakly stable if there exists a tie-breaking rule where μ is stable. Therefore a matching μ has positive probability of being stable if and only if it is weakly stable. Furthermore, if the number of possible tie-breaking is N then any weakly stable matching has a probability of being stable at least $\frac{1}{N}$.

An extreme case occurs if we have one woman only with n men, where the woman is indifferent between all men. In this case any matching (pair) has a $\frac{1}{n}$ probability of being stable. An even more unfortunate scenario is when we have n men and n women, each women is indifferent between all men, and each man ranks the women in a strict order in the same way, e.g. in the order of their indices. In this case, the probability that the first woman picks her best partner, and thus does not block any matching is $\frac{1}{n}$. Suppose that the first woman picked her best partner, the probability that the second woman also picks her best partner from the remaining $n-1$ men is $\frac{1}{n-1}$, and so on. Therefore, the probability that an arbitrary complete matching is stable is $\frac{1}{n(n-1)\ldots 2} = \frac{1}{n!}$.

Theorem 9. *For the compact indifference model* MatchingWithHighestStabilityProbability *is NP-hard, even if only one side of the market has uncertain agents.*

Proof sketch. For an instance I of SMTI, let $opt(I)$ denote the maximum size of a weakly stable matching in I. Halldorsson et al. [6] showed [in the proof of Corollary 3.4] that given an instance I of SMTI of size n, where only one side of the market has agents with indifferences and each of these agents has a single tie of size two, and any arbitrary small positive ϵ, it is NP-hard to distinguish between the following two cases: (1) $opt(I) \geq \frac{21-\epsilon}{27}n$ (2) $opt(I) < \frac{19+\epsilon}{27}n$.

When choosing ϵ so that $0 < \epsilon < \frac{1}{2}$ we can simplify the above cases to (1) $opt(I) > \frac{41}{54}n$, since $opt(I) \geq \frac{21-\epsilon}{27}n > \frac{41}{54}n$ and (2) $opt(I) < \frac{39}{54}n$, since $opt(I) < \frac{19+\epsilon}{27}n < \frac{39}{54}n$.

Therefore, the number of agents left unmatched on either side of the market is less than $\frac{13}{54}n$ in the first case and more than $\frac{15}{54}n$ in the second case. Let us now extend instance I to a larger instance of SMTI I' as follows. Besides the n men $M = \{m_1, \ldots, m_n\}$ and n women $W = \{w_1, \ldots, w_n\}$, we introduce $\frac{13}{54}n$ men $X = \{x_1, \ldots x_k\}$ and another $\frac{n}{27}$ men $Y = \{y_1, \ldots y_l\}$ and $\frac{n}{27}$ women $Z = \{z_1, \ldots z_l\}$. Furthermore, for each $y_j \in Y$, we introduce n men $Y^j = \{y_1^j, \ldots, y_n^j\}$. We create the preferences of I' as follows. The preferences of men M remain the same. For each woman $w \in W$ we append the men X and then Y at the end of her list in the order of their indices. Each man $x_i \in X$ has only all the women W in his list in the order of their indices. Furthermore, each $y_j \in Y$ has all the women W first in his preference list in the order or their indices and then z_j. Let each $z_j \in Z$ has y_j as first choice and then all the men Y^j in one tie of size n. Each man in Y^j has only z_j in his list. We will show that in case one $p_S(I') \geq \frac{1}{2^n}$, whilst in case two $p_S \leq (\frac{1}{n})^{\frac{n}{27}}$. Therefore, for $n > 2^{27}$, it is NP-hard to decide which of the two separate intervals contains the value $p_S(I')$. □

6 Joint Probability Model

In this section, we examine problems concerning the joint probability model.

Theorem 10. *For the joint probability model,* StabilityProbability *can be solved in polynomial time.*

Corollary 3. *For the joint probability model,* IsStabilityProbabilityNonZero *and* IsStabilityProbabilityOne *can be solved in polynomial time.*

For the joint probability model, the problem ExistsCertainlyStableMatching is equivalent to checking whether the intersection of the sets of stable matchings of the different preference profiles is empty or not.

Theorem 11. *For the joint probability model,* ExistsCertainlyStableMatching *is NP-complete.*

Proof sketch. The problem is in NP, since computing StabilityProbability can be done in polynomial time by Theorem 10. The proof is by reduction from 3-Colorability. Let $G = (V, E)$ be a graph specifying an instance of 3-Colorability, where $V = \{v_1, \ldots, v_n\}$. We construct an instance I of SMULP assuming the joint probability model.

For each vertex $v_i \in V$, we introduce three men $m_{i,1}, m_{i,2}, m_{i,3}$ and three women $w_{i,1}, w_{i,2}, w_{i,3}$. Then, we introduce one preference profile P_0 that ensures that every certainly stable matching matches—for each $i \in [n]$—each $m_{i,j}$ to some $w_{i,j'}$ and, vice versa, each $w_{i,j}$ to some $m_{i,j'}$, for $j, j' \in [3]$. Moreover, it ensures that for each $i \in [n]$, exactly one of three matchings between the men $m_{i,j}$ and the women $w_{i,j}$ must be used:

(1) $m_{i,1}$ is matched to $w_{i,1}$, $m_{i,2}$ is matched to $w_{i,2}$, and $m_{i,3}$ is matched to $w_{i,3}$;
(2) $m_{i,1}$ is matched to $w_{i,2}$, $m_{i,2}$ is matched to $w_{i,3}$, and $m_{i,3}$ is matched to $w_{i,1}$; or
(3) $m_{i,1}$ is matched to $w_{i,3}$, $m_{i,2}$ is matched to $w_{i,1}$, and $m_{i,3}$ is matched to $w_{i,2}$;

Intuitively, choosing one of the matchings (1)–(3) for the agents $m_{i,j}, w_{i,j}$ corresponds to coloring vertex v_i with one of the three colors in $\{1, 2, 3\}$.

Then, for each edge $e = \{v_{i_1}, v_{i_2}\} \in E$, and for each color $c \in \{1, 2, 3\}$, we introduce a preference profile $P_{e,c}$ that ensures that in any certainly stable matching, the agents $m_{i_1,j}, w_{i_1,j}$ and the agents $m_{i_2,j}, w_{i_2,j}$ cannot both be matched to each other with matching (c). We let each preference profile appear with non-zero probability (e.g., we take a uniform lottery). As a result, any certainly stable matching directly corresponds to a proper 3-coloring of G. A detailed description of the preference profiles P_0 and $P_{e,c}$ can be found in [1], as well as a proof of correctness for this reduction. □

By modifying the proof of Theorem 11, the following can also be proved.

Corollary 4. *For the joint probability model,* ExistsCertainlyStableMatching *is NP-complete, even when there are only 16 preference profiles in the lottery.*

7 Future work

First we note that we left open two outstanding questions, as described in Table 1. In this paper we focused on the problem of computing a matching with the highest stability probability. However, a similarly reasonable goal could be to minimize the expected number of blocking pairs. It would also be interesting to investigate some further realistic probability models, such as the situation when the candidates are ranked according to some noisy scores (like the SAT scores in the US college admissions). This would be a special case of the joint probability model that may turn out to be easier to solve. Finally, in a follow-up paper we are planning to investigate another probabilistic model that is based on independent pairwise comparisons.

Acknowledgments. Biró is supported by the Hungarian Academy of Sciences under its Momentum Programme (LP2016-3) and the Hungarian Scientific Research Fund, OTKA, Grant No. K108673. Rastegari was supported EPSRC grant EP/K010042/1 at the time of the submission. De Haan is supported by the Austrian Science Fund (FWF), project P26200. The authors gratefully acknowledge the support from European Cooperation in Science and Technology (COST) action IC1205. Serge Gaspers is the recipient of an Australian Research Council (ARC) Future Fellowship (FT140100048) and acknowledges support under the ARC's Discovery Projects funding scheme (DP150101134). Data61/CSIRO (formerly, NICTA) is funded by the Australian Government through the Department of Communications and the ARC through the ICT Centre of Excellence Program.

References

1. H. Aziz, P. Biró, S. Gaspers, R. de Haan, N. Mattei, and B. Rastegari. Stable matching with uncertain linear preferences. Technical Report 1607.02917, Cornell University Library, July 2016. Available at http://arxiv.org/abs/1607.02917.
2. J. Drummond and C. Boutilier. Preference elicitation and interview minimization in stable matchings. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 645–653, 2014.
3. L. Ehlers and J. Massó. Matching markets under (in)complete information. *Journal of Economic Theory*, 157:295–314, 2015.
4. D. Gusfield and R. Irving. *The Stable Marriage Problem: Structure and Algorithms.* MIT Press, 1989.
5. D. Gusfield and R. W. Irving. *The stable marriage problem: Structure and algorithms.* MIT Press, Cambridge, MA, USA, 1989.
6. M. M. Halldórsson, K. Iwama, S. Miyazaki, and H. Yanagisawa. Improved approximation results for the stable marriage problem. *ACM Trans. Algorithms*, 3(3), 2007.
7. N. Hazon, Y. Aumann, S. Kraus, and M. Wooldridge. On the evaluation of election outcomes under uncertainty. *Artificial Intelligence*, 189:1–18, 2012.
8. R. Irving. Stable marriage and indifference. *Discrete Applied Mathematics*, 48: 261–272, 1994.
9. Y. Li and V. Conitzer. Cooperative game solution concepts that maximize stability under noise. In *AAAI*, pages 979–985, 2015.
10. D. Manlove. *Algorithmics of Matching Under Preferences.* World Scientific Publishing Company, 2013.
11. B. Rastegari, A. Condon, N. Immorlica, R. Irving, and K. Leyton-Brown. Reasoning about optimal stable matching under partial information. In *Proceedings of the ACM Conference on Electronic Commerce (EC)*, pages 431–448. ACM, 2014.
12. A. E. Roth and M. A. O. Sotomayor. *Two-Sided Matching: A Study in Game Theoretic Modelling and Analysis.* Cambridge University Press, 1990.
13. D. J. A. Welsh and C. Merino. The Potts model and the Tutte polynomial. *Journal of Mathematical Physics*, 41(3):1127–1152, 2000.