Descending the Stable Matching Lattice: How many Strategic Agents are required to turn Pessimality to OptIMALity?

Ndiamé Ndiaye∗ Sergey Norin† Adrian Vetta‡

Abstract

The set of stable matchings induces a distributive lattice. The supremum of the stable matching lattice is the boy-optimal (girl-pessimal) stable matching and the infimum is the girl-optimal (boy-pessimal) stable matching. The classical boy-proposal deferred-acceptance algorithm returns the supremum of the lattice, that is, the boy-optimal stable matching. In this paper, we study the smallest group of girls, called the minimum winning coalition of girls, that can act strategically, but independently, to force the boy-proposal deferred-acceptance algorithm to output the girl-optimal stable matching. We characterize the minimum winning coalition in terms of stable matching rotations and show that its cardinality can take on any value between 0 and ⌊n/2⌋, for instances with n boys and n girls. Our two main results concern the random matching model. First, the expected cardinality of the minimum winning coalition is small, specifically (1/2 + o(1)) log n. This resolves a conjecture of Kupfer [14]. Second, in contrast, a randomly selected coalition must contain nearly every girl to ensure it is a winning coalition asymptotically almost surely. Equivalently, for any ε > 0, the probability a random group of (1 − ε)n girls is not a winning coalition is at least δ(ε) > 0.

1 Introduction

We study the stable matching problem with n boys and n girls. Each boy has a preference ranking over the girls and vice versa. A matching is stable if there is no boy-girl pair that prefer each other over their current partners in the matching. A stable matching always exists and can be found by the deferred-acceptance algorithm [5]. Furthermore, the set of stable matchings forms a lattice whose supremum matches each boy to his best stable-partner and each girl to her worst stable-partner. This matching is called the boy-optimal (girl-pessimal) stable matching. Conversely, the infimum of the lattice matches each boy to his worst stable-partner and each girl to her best stable-partner. Consequently this matching is called the girl-optimal (boy-pessimal) stable matching.

Interestingly, the deferred-acceptance algorithm outputs the optimal stable matching for the proposing side. Perhaps surprisingly, the choice of which side makes the proposal can make a significant difference. For example, for the random matching model, where the preference list of each boy and girl is sampled uniformly and independently, Pittel [16] showed the boy-proposal deferred acceptance algorithm assigns the boys with much better ranking partners than the girls. Specifically, with high probability, the sum of the partner ranks is close to n log n for the boys and close to \( \frac{n^2}{\log n} \) for the girls. Hence, on average, each boy ranks his partner at position log n at the boy-optimal

∗Department of Mathematics & Statistics, McGill University: ndiame.ndiaye@mail.mcgill.ca
†Department of Mathematics & Statistics, McGill University: sergey.norin@mcgill.ca
‡School of Computer Science and Dept. of Mathematics & Statistics, McGill University: adrian.vetta@mcgill.ca
stable matching while each girl only ranks her partner at position \( \frac{n}{\log n} \). Consequently, collectively the girls may have a much higher preference for the infimum (girl-optimal) stable matching than the supremum (girl-pessimal) stable matching output by the boy-proposal deferred-acceptance algorithm.

Remarkably, Ashlagi et al. [1] proved that in an unbalanced market with one fewer girls than boys this advantage to the boys is reversed. In the random matching model, with high probability, each girl is matched to a boy she ranks at \( \log n \) on average and each boy is matched to a girl he ranks at \( \frac{n}{\log n} \) on average, even using the boy-proposal deferred-acceptance algorithm.\(^1\) Kupfer [14] then showed a similar effect arises in a balanced market in which exactly one girl acts strategically. The expected rank of the partner of each girl improves to \( O(\log^4 n) \) while the expected rank of the partner of each boy deteriorates to \( \Omega(\frac{n}{\log^{2+\varepsilon} n}) \). Thus, just one strategic girl suffices for the stable matching output by the boy-proposal deferred-acceptance algorithm to change from the supremum of the lattice to a stable matching “close” to the infimum. But how many strategic girls are required to guarantee the infimum itself is output? Kupfer [14] conjectured that \( O(\log n) \) girls suffice in expectation. In this paper we prove this conjecture. More precisely, we show that the minimum number of strategic girls required is \( \frac{1}{2} \log n + O(\log \log n) = (\frac{1}{2}+o(1)) \log n \) in expectation. Consequently, the expected cardinality of the optimal winning coalition of girls is relatively small.

Conversely, a random coalition of girls must be extremely large, namely of cardinality \( n - o(n) \), if it is to be a winning coalition with high probability. We prove that, for any \( \varepsilon > 0 \), the probability a random group of \( (1 - \varepsilon)n \) girls is not a winning coalition is at least a constant.

1.1 Overview

In Section 2, we present the relevant background on the stable matching problem, in particular, concerning the stable matching lattice and the rotation poset. In Section 3 we provide a characterization of winning coalitions of girls in terms of minimal rotations in the rotation poset. This allows us to show that for general stable matching instances the cardinality of the minimum winning coalition may take on every integral value between a lower bound of 0 and an upper bound of \( \left\lfloor \frac{n}{2} \right\rfloor \). In Section 4, we present an example to illustrate the relevant stable matching concepts and ideas used in the paper. In Section 5, we present the random matching model studied for the main results of the paper. Our first main result is given in Section 6 and shows that in random instances the cardinality of the minimum winning coalition is much closer to the lower bound than the upper bound. Specifically, in the random matching model, the expected cardinality of the minimum winning coalition is \( \frac{1}{2} \log n + O(\log \log n) \). Our second main result is presented in Section 7 and shows that for a randomly selected coalition to be a winning coalition with probability \( 1 - o(1) \), it must have cardinality \( n - o(n) \).

2 The Stable Matching Problem

Here we review the stable matching problem and the concepts and results relevant to this paper. The reader is referred to the book [9] by Gusfield and Irving for a comprehensive introduction to stable matchings. An illustrative example, defined in Table 1, will also be presented in Section 4. We remark that this example is deferred until all the relevant concepts have been defined (indeed,
it will be clearer to present the structural properties of the example in a different order than how
they are defined in this review).

We are given a set \( B = \{ b_1, b_2, \ldots, b_n \} \) of boys and a set \( G = \{ g_1, g_2, \ldots, g_n \} \) of girls. Every
boy \( b \in B \) has a preference ranking \( \succ_b \) over the girls; similarly, every girl \( g \in G \) has a preference
ranking \( \succ_g \) over the boys. Now let \( \mu \) be a (perfect) matching between the boys and girls. We say
that boy \( b \) is matched to girl \( \mu(b) \) in the matching \( \mu \); similarly, girl \( g \) is matched to boy \( \mu(g) \). Boy \( b \)
and girl \( g \) form a blocking pair \( \{ b, g \} \) if they prefer each other to their partners in the matching \( \mu \);
that is \( g \succ_b \mu(b) \) and \( b \succ_g \mu(g) \). A matching \( \mu \) that contains no blocking pair is called stable;
otherwise it is unstable. In the stable matching problem, the task is to find a stable matching.

2.1 The Deferred-Acceptance Algorithm

The first question to answer is whether or not a stable matching is guaranteed to exist. Indeed a
stable matching always exists, as shown in the seminal work of Gale and Shapley [5]. Their proof
was constructive; the deferred-acceptance algorithm, described in Algorithm 1, outputs a stable
matching.

\textbf{Algorithm 1: Deferred-Acceptance (Boy-Proposal Version)}

\begin{align*}
\text{while there is an unmatched boy } b & \text{ do} \\
\quad & \text{Let } b \text{ propose to his favourite girl } g \text{ who has not yet rejected him;} \\
\quad & \text{if } g \text{ is unmatched then} \\
\quad & \quad g \text{ provisionally matches with } b; \\
\quad & \text{else if } g \text{ is provisionally matched to } \hat{b} & \text{then} \\
\quad & \quad g \text{ provisionally matches to her favourite of } b \text{ and } \hat{b}, \text{ and rejects the other;}
\end{align*}

The key observation here is that only a girl can reject a provisional match. Thus, from a girl’s
perspective, her provisional match can only improve as the algorithm runs. It follows that the
defined-acceptance algorithm terminates when every girl has received at least one proposal. In
addition, from a boy’s perspective, his provisional match can only get worse as the algorithm runs.
Indeed, it would be pointless for a boy to propose to girl who has already rejected him. Thus, each
boy will make at most \( n \) proposals. Furthermore, because each boy makes proposals in decreasing
order of preference, every girl must eventually receive a proposal. Thus the deferred-acceptance
algorithm must terminate with a perfect matching \( \mu \). At this point all provisional matches are
made permanent. But why will this permanent set of matches \( \mu \) form a stable matching? The proof
is simple and informative, so we include it for completeness.

\textbf{Theorem 2.1 (Gale and Shapley 1962 [5]).} The deferred-acceptance algorithm outputs a stable
matching.

\textbf{Proof.} Suppose \( \{ b, g \} \) is a blocking pair for \( \mu \). Then boy \( b \) prefers girl \( g \) over girl \( \hat{g} = \mu(b) \), that
is \( g \succ_b \mu(b) \). So \( b \) must have proposed to \( g \) before proposing to \( \hat{g} \). Then \( g \) must have rejected \( b \).
Either she rejected \( b \) at the time of the proposal or she provisionally accepted his offer but later
rejected him after receiving a better offer. As her provisional partner only improves over time, it
3.4 is ill-specified. In particular, which unmatched boy is selected to make the next proposal? Theorem 2.3 explains the laxity of our description. It is irrelevant which unmatched boy is selected to make the next proposal. That is, \( \mu(g) \succ g b \), and so \( \{b, g\} \) is not a blocking pair.

2.2 The Stable Matching Lattice

So a stable matching always exists. In fact, there may be an exponential number of stable matchings [12]; see Theorem 3.4 for an example. The set \( \mathcal{M} \) of all stable matchings forms a poset \( (\mathcal{M}, \succeq) \) whose order \( \succeq \) is defined via the preference lists of the boys. Specifically, \( \mu_1 \succeq \mu_2 \) if and only if every boy weakly prefers their partner in the stable matching \( \mu_1 \) to their partner in the stable matching \( \mu_2 \); that is \( \mu_1(b) \succeq_b \mu_2(b) \), for every boy \( b \).

Conway (see Knuth [12]) observed that the poset \( (\mathcal{M}, \succeq) \) is in fact a distributive lattice. Thus, by the lattice property, each pair of stable matchings \( \mu_1 \) and \( \mu_2 \) has a join (least upper bound) and a meet (greatest lower bound) in the lattice. Moreover, the join \( \hat{\mu} = \mu_1 \lor \mu_2 \) has the remarkable property that each boy \( b \) is matched to his most preferred partner amongst the girls \( \mu_1(b) \) and \( \mu_2(b) \). Similarly, in the meet \( \hat{\mu} = \mu_1 \land \mu_2 \) each boy is matched to his least preferred partner amongst the girls \( \mu_1(b) \) and \( \mu_2(b) \). In particular, in the supremum \( 1 = \bigvee_{\mu \in \mathcal{M}} \mu \) of the lattice each boy is matched to his most preferred partner from any stable matching (called his best stable-partner). Accordingly, the matching \( 1 \) is called the boy-optimal stable matching. On the other hand, in the infimum \( 0 = \bigwedge_{\mu \in \mathcal{M}} \mu \) of the lattice each boy is matched to his least preferred partner from any stable matching (called his worst stable-partner). Accordingly, the matching \( 0 \) is called the boy-pessimal stable matching.

Theorem 2.2. [5] The deferred-acceptance algorithm outputs the boy-optimal stable matching.

Proof. If not, let \( b \) be the first boy rejected by a stable partner, say \( g \), during the course of the deferred-acceptance algorithm. Assume there is a stable matching \( \hat{\mu} \) in which the pair \((b, g)\) is matched and assume that \( g \) rejects \( b \) in favour of the boy \( \hat{b} \). By assumption \( \hat{b} \) was the first boy rejected by a stable partner so boy \( b \) had not been rejected by any stable-partner when \( g \) rejected \( b \). Thus \( \hat{b} \) prefers \( g \) over any stable partner. In particular, he prefers \( g \) over his stable partner \( \hat{\mu}(b) \neq g \). But then \( g \succ_{b, \hat{b}} \hat{\mu}(b) \) and \( \hat{b} \succ_g \hat{\mu}(g) = b \). Hence, \( \{b, g\} \) is a blocking pair for the matching \( \hat{\mu} \), a contradiction.

The reader may have observed that the description of the deferred-acceptance algorithm given in Algorithm 1 is ill-specified. In particular, which unmatched boy is selected to make the next proposal? Theorem 2.2 explains the laxity of our description. It is irrelevant which unmatched boy is chosen in each step, the final outcome is guaranteed to be the boy-optimal stable matching! In fact, the original description of the algorithm by Gale and Shapley [5] allowed for simultaneous proposals by unmatched boys — again this has no effect on the stable matching output.

The inverse poset \( (\mathcal{M}, \preceq) \) is also of fundamental interest. Indeed, McVitie and Wilson [15] made the surprising observation that \( (\mathcal{M}, \preceq) \) is the lattice defined using the preference lists of the girls rather than the boys. That is, every boy weakly prefers their partner in the stable matching \( \mu_1 \) to their partner in the stable matching \( \mu_2 \) if and only if every girl weakly prefers their partner in the stable matching \( \mu_2 \) to their partner in the stable matching \( \mu_1 \).

Theorem 2.3. [15] If \( \mu_1 \succeq \mu_2 \) in the lattice \( (\mathcal{M}, \succeq) \) then every girl weakly prefers \( \mu_2 \) over \( \mu_1 \).

Proof. Assume there is a girl \( g \) who prefers boy \( b = \mu_1(g) \) over boy \( \mu_2(g) \). But, by assumption, boy \( b \) prefers \( g = \mu_1(b) \) over girl \( \mu_2(b) \). Thus \( \{b, g\} \) is a blocking pair for the matching \( \mu_2 \), a contradiction.
Consequently, the boy-optimal stable matching 1 is also the girl-pessimal stable matching and the boy-pessimal stable matching 0 is the girl-optimal stable matching.

For our example, the set of stable matchings and the stable matching lattice are shown in Table 3 and Figure 3 of Section 4, respectively.

2.3 The Rotation Poset

Recall that the lattice \((\mathcal{M}, \succeq)\) is a distributive lattice. This is important because the fundamental theorem for finite distributive lattices of Birkhoff [2] states that associated with any distributive lattice \(\mathcal{L}\) is a unique auxiliary poset \(\mathcal{P}(\mathcal{L})\). Specifically, the order ideals (or down-sets) of the auxiliary poset \(\mathcal{P}\), ordered by inclusion, form the lattice \(\mathcal{L}\). We refer the reader to the book of Stanley [19] for details on the fundamental theorem for finite distributive lattices. For our purposes, however, it is sufficient to note that the auxiliary poset \(\mathcal{P}\) for the stable matching lattice \((\mathcal{M}, \succeq)\) has an elegant combinatorial description that is very amenable in studying stable matchings.

In particular, the auxiliary poset for the stable matching lattice is called the rotation poset \(\mathcal{P} = (\mathcal{R}, \succeq)\) and was first discovered by Irving and Leather [10]. The elements of the auxiliary poset are rotations. Informally, given a stable matching \(\mu\), a rotation will rearrange the partners of a suitably chosen subset of the boys in a circular fashion to produce another stable matching. Formally, a rotation \(R \in \mathcal{R}\) is a subset of the pairs in the stable matching \(\mu\), \(R = [(b_0, g_0), (b_1, g_1), \ldots, (b_k, g_k)]\), such that for each boy \(b_i\), the girl \(g_{i+1}\) is the first girl after his current stable-partner \(g_i\) on his preference list who would accept a proposal from him. That is, \(g_{i+1}\) prefers boy \(b_i\) over her current partner boy \(b_{i+1}\) and every girl \(g\) that boy \(b_i\) ranks on his list between \(g_i\) and \(g_{i+1}\) prefers her current partner in \(\mu\) over \(b_i\).

In this case, we say that \(R\) is a rotation exposed by the stable matching \(\mu\). Let \(\hat{\mu} = \mu \otimes R\) be the perfect matching obtained by matching boy \(b_i\) with the girl \(g_{i+1}\mod k+1\), for each \(0 \leq i \leq k\), with all other matches the same as in \(\mu\). Irving and Leather [10] showed that \(\hat{\mu}\) is also a stable matching. More importantly they proved:

**Theorem 2.4.** [10] The matching \(\hat{\mu}\) is covered\(^2\) by \(\mu\) in the Hasse diagram of the stable matching lattice if and only if \(\hat{\mu} = \mu \otimes R\) for some rotation \(R\) exposed by \(\mu\).

Theorem 2.4 implies that we may traverse the stable matching lattice \((\mathcal{M}, \succeq)\) using rotations. As stated, we may also derive a poset \(\mathcal{P} = (\mathcal{R}, \succeq)\) whose elements are rotations. Let \(\mathcal{R}_\mu\) be the set of all rotations exposed in \(\mu\). Then \(\mathcal{R} = \bigcup_{\mu \in \mathcal{M}} \mathcal{R}_\mu\) is the set of all rotations. We then define the partial order \(\succeq\) as follows. Let \(R_1 \succeq R_2\) in \(\mathcal{P}\) if and only if for any stable matching \(\mu_1 \in \{\mu \in \mathcal{M} : R_1 \in \mathcal{R}_\mu\}\) and any stable matching \(\mu_2 \in \{\mu \in \mathcal{M} : R_2 \in \mathcal{R}_\mu\}\), either \(\mu_1\) and \(\mu_2\) are incomparable or \(\mu_1 \succeq \mu_2\) in \((\mathcal{M}, \succeq)\). This rotation poset \(\mathcal{P} = (\mathcal{R}, \succeq)\) is the auxiliary poset for the stable matching lattice \((\mathcal{M}, \succeq)\); see Gusfield and Irving [9]. In particular, there is a bijection between stable matchings and antichains of the rotation poset.

The set of rotations and the rotation poset for our running example are illustrated in Table 2 and Figure 2 of Section 4, respectively.

We remark that, unlike the stable matching lattice, the cardinality of the rotation poset is always polynomial. Specifically, any boy-girl pair \(\{b, g\}\) can appear in at most one rotation [10]. It immediately follows that the rotation poset has at most \(O(n^2)\) elements; in fact, Gusfield [8] showed how to find all the rotations in \(O(n^2)\) time.

\(^2\)We say \(y\) is covered by \(x\) in a poset if \(x \succeq y\) and there is no element \(z\) such that \(x \succeq z \succeq y\).
2.4 The Rotation Graph

For any stable matching $\mu = \{(b_1, g_1), (b_2, g_2), \ldots, (b_n, g_n)\}$ we define an auxiliary directed graph $H(\mu)$. This graph, which we call the (exposed) rotation graph, has a vertex $i$ for each boy $b_i$. There is an arc from $i$ to $j$ if the next girl on $b_i$’s list to prefer $b_i$ over her current partner is $g_j$. If for some $b_i$, no such girl exists, then $i$ has out-degree 0; otherwise it has out-degree 1. By definition, the rotations exposed in $\mu$ are exactly the cycles of $H(\mu)$. (See Figure 1 in Section 4 for the rotation-graph $H(1)$ for the running example.) For example, if $\mu = 1$ then $H(1)$ consists of the set of rotations exposed in the boy-optimal stable matching. We call these the maximal rotations.

A rotation $R$ exposed in $\mu$ is minimal if $\mu \otimes R = 0$. Equivalently, the minimal rotations are the set of rotations exposed in the girl-optimal stable matching when ordering using the preferences of the girls rather than the boys.

3 Incentives in the Stable Matchings Problem

Intuitively, because the deferred-acceptance algorithm outputs the boy-optimal stable matching, there is no incentive for a boy not to propose to the girls in order of preference. This fact was formally proven by Dubins and Freedman [3]. On the other hand, because the stable matching is girl-pessimal, it can be beneficial for a girl to strategize. Indeed, Roth [18] showed that no stable matching mechanism exists that is incentive compatible for every participant.

3.1 The Minimum Winning Coalition of Girls

The structure of the stable matching lattice $\mathcal{L}$ is extremely useful in understanding the incentives that arise in the stable matching problem. For example, the following structure will be of importance in this paper. Let $F \subseteq G$ be a group of girls and let $\mathcal{M}_F$ be the collection of stable matchings where every girl in $F$ is matched to their best stable-partner. Given the aforementioned properties of the join and meet operation in the stable matching lattice, it is easy to verify that $\mathcal{L}_F = (\mathcal{M}_F, \succeq)$ is also a lattice. Thus, $\mathcal{L}_F$ has a supremum $1_F$ which is the boy-optimal stable matching given that every girl in $F$ is matched to their best stable-partner. Similarly, $\mathcal{L}_F$ has an infimum $0_F$ which is the boy-pessimal stable matching given that every girl in $F$ is matched to their best stable-partner.

Observe that $0_F$ is the girl-optimal stable-matching 0, for any subset $F$ of the girls.

Why is this useful here? Well, imagine that each girl in $F$ rejects anyone who is not their best stable-partner. Then the deferred-acceptance algorithm will output the stable matching $1_F$; see also the works of Gale and Sotomayor [6] on strong equilibria and of Gonczarowski [7] on blacklisting. Of course, if $F = G$ then both $1_G$ and $0_G$ must match every girl to their optimal stable partner so $1_G = 0_G = 0$.

We will call any $F \subseteq G$ such that $1_F = 0$ a winning coalition and the smallest such group is called a minimum winning coalition. Winning coalitions can be found using the rotation poset.

**Theorem 3.1.** A set of girls is a winning coalition if and only if it contains at least one girl from each minimal rotation in the rotation poset $(\mathcal{R}, \succeq)$

**Proof.** Let $G^2$ be the set of girls who have at least two stable-partners. For each girl $g_j \in G^2$, let $\mathcal{M}_j$ be the set of stable matchings in which she is not matched to her best stable-partner. Then $\mathcal{L}_j = (\mathcal{M}_j, \succeq)$ is a lattice with supremum $1^j$ and infimum $0^j$. Observe that $1^j = 1$ and $0^j \neq 0$. 

Now let \( \{\mu_1, \mu_2, \ldots, \mu_k\} \) be the minimal stable-matchings in the poset \((\mathcal{M} \setminus \mathbf{0}, \succeq)\). That is, \( \{\mu_1, \mu_2, \ldots, \mu_k\} \) is the set of matchings such that for any \( i \in \{1, \ldots, k\} \) and any stable matching \( \mu \not\in \{\mu_1, \mu_2, \ldots, \mu_k\} \cup \{\mathbf{0}\} \), there is a boy who strictly prefers \( \mu \) over \( \mu_i \). For each \( 1 \leq \ell \leq k \), observe that \( \mu_\ell = \mathbf{0}^\ell \) for some girl \( g_j \in G^2 \). But, if \( \mu_\ell = \mathbf{0}^\ell \) then girl \( g_j \) must be matched to her best stable-partner in \( \mu_i \) for any \( i \neq \ell \). Otherwise, because \( \mu_\ell \land \mu_i = \mathbf{0} \) in the stable matching lattice \( \mathcal{L} \), it would be the case that girl \( g_j \) is not matched to her best stable-partner in \( \mathbf{0} \), a contradiction. Let \( U_\ell \) be the set of girls who are not matched to their best stable-partner in \( \mu_\ell \). Thus, the sets \( U_1, \ldots, U_k \) are non-empty and disjoint.

Let \( F \) be any group of girls that contains at least one girl from each set \( U_\ell \), for \( 1 \leq \ell \leq k \). We claim that \( 1_F = \mathbf{0} \) and, consequently, \( F \) is a winning coalition. For each \( 1 \leq \ell \leq k \), at least one girl \( g_j \) in \( F \) is not matched to best stable-partner in \( \mu_\ell \). Thus \( \mu_\ell = \mathbf{0}^\ell \) for some girl \( g_j \in G^2 \). It follows immediately that there is a unique stable-matching, namely the girl-optimal stable matching \( \mathbf{0} \), that matches every girl in \( F \) to their best stable-partner. Hence, \( 1_F = \mathbf{0} \) as claimed.

Conversely, let \( F \) be any group of girls such that for some \( \ell \), \( F \cap U_\ell = \emptyset \). We claim that \( 1_F \neq \mathbf{0} \) and, consequently, \( F \) is a losing coalition. By definition, every girl in \( F \) is matched to their best stable-partner in \( \mu_\ell \). But then, by definition, \( 1_F \geq \mu_\ell \) as claimed.

Finally, observe that our definition of \( U_\ell \) is exactly the set of girls in the unique rotation exposed in \( \mu_i \) which is a minimal rotation of the rotation poset \((\mathcal{R}, \succeq)\) which proves the statement of the theorem.

Theorem 3.1 allows us to find a minimum winning coalition.

**Corollary 3.2.** The cardinality of the minimum winning coalition is equal to the cardinality of the set of minimal rotations in the rotation poset \((\mathcal{R}, \succeq)\).

Section 4 provides an illustration of how rotations correspond to stable matchings and gives a minimum winning coalition for the running example.

### 3.2 Efficiency and Extremal Properties

From the structure inherent in Theorem 3.1 and Corollary 3.2 we can make several straight-forward deductions regarding winning coalitions.

First, Theorem 3.1 implies that we have a polynomial algorithm to verify winning coalitions. Likewise Corollary 3.2 implies that we have a polynomial time algorithm to compute the minimum winning coalition. In fact, the techniques of Gusfield [8] (see also [9]) can now be used to solve both problems in \( O(n^2) \) time.

Second, we can upper bound the cardinality of the minimum winning coalition.

**Lemma 3.3.** In any stable matching problem the minimum winning coalition has cardinality at most \( \left\lfloor \frac{n}{2} \right\rfloor \).

**Proof.** Consider the minimal stable-matchings \( \{\mu_1, \mu_2, \ldots, \mu_k\} \) in the poset \((\mathcal{M} \setminus \mathbf{0}, \succeq)\). We claim \( k \leq \left\lfloor \frac{n}{2} \right\rfloor \). To prove this, observe that since \( \forall \ell \in [k] \mu_\ell \not\in \{\mu_1, \mu_2, \ldots, \mu_k\} \cup \{\mathbf{0}\} \) there must be at least two girls who are not matched to their best stable-partners in the stable matching \( \mu_\ell \). Furthermore, recall that each girl is matched to their best stable-partner in every matching \( \{\mu_1, \mu_2, \ldots, \mu_k\} \) except at most one. It immediately follows that \( k \leq \left\lfloor \frac{n}{2} \right\rfloor \). \( \square \)
Can this upper bound on the cardinality of the minimum winning coalition ever be obtained? The answer is yes. In fact, every integer between 0 and \( \lfloor \frac{n}{2} \rfloor \) can be the cardinality of the smallest winning coalition.

**Theorem 3.4.** For each \( 0 \leq k \leq \lfloor \frac{n}{2} \rfloor \) there exists a stable matching instance where the minimum winning coalition has cardinality exactly \( k \).

**Proof.** Take any \( 0 \leq k \leq \lfloor \frac{n}{2} \rfloor \). We construct a stable matching instance where the minimum winning coalition has cardinality exactly \( k \) as follows. For \( 2k + 1 \leq \ell \leq n \), let boy \( b_\ell \) and girl \( g_\ell \) rank each other top of their preference lists – the other rankings in their preference lists may be arbitrary. Thus, boy \( b_\ell \) and girl \( g_\ell \) must be matched together in every stable matching.

For \( 1 \leq \ell \leq k \), let boy \( b_{2\ell-1} \) rank girl \( g_{2\ell-1} \) first and girl \( g_{2\ell} \) second and let boy \( b_{2\ell} \) rank girl \( g_{2\ell} \) first and girl \( g_{2\ell-1} \) second. In contrast, let girl \( g_{2\ell-1} \) rank boy \( b_{2\ell} \) first and boy \( b_{2\ell-1} \) second and let girl \( g_{2\ell} \) rank boy \( b_{2\ell-1} \) first and boy \( b_{2\ell} \) second. Again, all other rankings may be arbitrary.

It is then easy to verify that two possibilities arise. In any stable matching for each \( 1 \leq \ell \leq k \) either (i) both boys \( b_{2\ell-1} \) and \( b_{2\ell} \) are matched to their best stable-partners, namely girls \( g_{2\ell-1} \) and \( g_{2\ell} \), respectively, or (ii) both boys \( b_{2\ell-1} \) and \( b_{2\ell} \) are matched to their worst stable-partners, namely girls \( g_{2\ell} \) and \( g_{2\ell-1} \), respectively.

But this implies that to obtain the girl-optimal stable matching at least one girl from the pair \( \{g_{2\ell-1}, g_{2\ell}\} \) must misreport her preferences, for each \( 1 \leq \ell \leq k \). One girl from each of these pairs is also sufficient to output the girl-optimal stable matching. Thus the minimum winning coalition has cardinality exactly \( k \). \[\square\]

We remark that the instances constructed in the proof of Theorem 3.4 have \( 2^k \) stable matchings. As \( k \) can be as large as \( \lfloor \frac{n}{2} \rfloor \), this gives a simple proof of the well known fact that the number of stable matchings may be exponential in the number of participants [12].

We now have all the tools required to address the main questions in this paper. We will first, as promised, illustrate these tools using an example.
4 An Illustrative Example

Here we present an example to illustrate the main concepts covered in the paper. This stable matching instance is derived from an example constructed by Irving et al. [11]. There are eight boys and eight girls whose preference rankings are shown in Table 1.

Table 1: A Stable Matching Instance.

| Boy | Rank | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----|------|---|---|---|---|---|---|---|---|
| b₁  |      | g₄| g₃| g₂| g₁| g₅| g₇| g₆|   |
| b₂  |      | g₃| g₇| g₅| g₈| g₆| g₄| g₁| g₂|
| b₃  |      | g₇| g₅| g₈| g₃| g₆| g₂| g₁| g₄|
| b₄  |      | g₆| g₄| g₂| g₇| g₃| g₁| g₅| g₈|
| b₅  |      | g₈| g₇| g₁| g₅| g₆| g₄| g₃| g₂|
| b₆  |      | g₅| g₄| g₇| g₆| g₂| g₈| g₃| g₁|
| b₇  |      | g₁| g₄| g₅| g₆| g₂| g₈| g₃| g₇|
| b₈  |      | g₂| g₅| g₄| g₃| g₇| g₈| g₁| g₆|

| Girl | Rank | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|------|------|---|---|---|---|---|---|---|---|
| g₁   |      | b₃| b₁| b₅| b₇| b₄| b₂| b₈| b₆|
| g₂   |      | b₆| b₁| b₃| b₄| b₈| b₇| b₅| b₂|
| g₃   |      | b₇| b₄| b₃| b₆| b₅| b₁| b₂| b₈|
| g₄   |      | b₅| b₃| b₈| b₂| b₆| b₃| b₄| b₇|
| g₅   |      | b₄| b₁| b₃| b₈| b₇| b₃| b₆| b₅|
| g₆   |      | b₆| b₂| b₅| b₇| b₈| b₄| b₃| b₁|
| g₇   |      | b₇| b₈| b₁| b₆| b₂| b₃| b₄| b₅|
| g₈   |      | b₂| b₆| b₇| b₁| b₈| b₃| b₄| b₅|

This instance has 23 stable matchings. To see this, let’s begin by running the deferred acceptance algorithm (Algorithm 1) to find the boy-optimal stable matching. Observe that all the boys have different first preferences. Thus, the boys will consecutively each propose to their first choice who will temporarily accept; once all the boys have proposed these matches will become permanent. Thus, the boy-optimal matching \( M₁ \) simply matches each boy with his favourite girl. To find the remaining stable matchings, we swap partners using rotations.

We start by finding the rotations exposed at \( \mathbf{1} = M₁ \). To do this we simply search for the for directed cycles in the (exposed) rotation graph \( H(\mathbf{1}) \). Recall, this graph has eight vertices, one for each boy. The first vertex \( \mathbf{1} \) corresponds to boy \( b₁ \) who is matched to girl \( g₄ \) in \( M₁ \). If girl \( g₄ \) breaks up with boy \( b₁ \) then he will propose to his second choice, girl \( g₃ \). She prefers him over her current partner boy \( b₂ \) so will accept this offer. Thus boy \( b₁ \) will gain the current partner of boy \( b₂ \); hence there is an arc \((1, 2)\) in the rotation graph \( H(\mathbf{1}) \). To find the outgoing arc at vertex \( \mathbf{2} \) assume that \( g₃ \) does break-up with boy \( b₂ \). He will then propose his second choice, girl \( g₇ \). She prefers him over her current partner \( b₃ \) so will accept this offer. Thus, there is an arc at \((2, 3)\). Now \((b₃, g₇) \in M₁\) and if girl \( g₇ \) breaks-up with boy \( b₃ \) then he will next propose to \( g₅ \). She will accept as she prefers \( b₃ \) over her current partner boy \( b₆ \). So the rotation graph contains the arc \((3, 6)\). Next consider boy \( b₄ \). If
his partner, girl $g_6$ breaks up with him then he will propose to girl $g_4$. She will reject this proposal as she prefers her current partner $b_1$ over $b_1$. So $b_4$ will then propose to his third choice $g_2$ and this proposal will be accepted as she prefers him over her current partner $b_8$. Hence $H(1)$ contains the arc $(4,8)$. Continuing in this fashion, the reader can verify that the rotation graph $H(1)$ is as shown in Figure 1.

![Figure 1: The (Exposed) Rotation Graph $H(1)$ at the Boy-Optimal Stable Matching.](image)

Observe that the rotation graph $H(M_1)$ contains a single cycle $\{v_1, v_2, v_3, v_6\}$. Consequently, there is exactly one exposed rotation, namely $\rho_1 = \{(b_1,g_4),(b_2,g_3),(b_5,g_7),(b_6,g_5)\}$. We remark that this is the unique maximal rotation for this stable matching instance. Thus from $M_1 = (g_4,g_3,g_7,g_6,g_8,g_5,g_1,g_2)$ we may create one new stable matching by performing the rotation $\rho_1$. Specifically, we rotate the partners of the boys $\{b_1,b_2,b_3,b_6\}$. This gives the stable matching $M_2 = (g_3,g_7,g_5,g_6,g_8,g_4,g_1,g_2)$.

Similarly, for $M_2$ the rotation graph $H(M_2)$ contains a single directed cycle $\{v_3,v_5,v_7\}$. Rotating the partners of boys $\{b_3,b_5,b_7\}$ then produces the stable matching $M_3 = (g_3,g_7,g_8,g_6,g_1,g_4,g_5,g_2)$. The rotation graph $H(M_3)$ contains two directed cycle correspond to $\rho_3$ and $\rho_4$. These are specified, along with all the other rotations in Table 2.

| Rotation | Rotation |
|----------|----------|
| $\rho_1$ | $(b_1,g_4),(b_2,g_3),(b_5,g_7),(b_6,g_5)$ |
| $\rho_2$ | $(b_3,g_5),(b_5,g_8),(b_7,g_1)$ |
| $\rho_3$ | $(b_4,g_6),(b_8,g_2),(b_7,g_5)$ |
| $\rho_4$ | $(b_1,g_3),(b_5,g_8)$ |
| $\rho_5$ | $(b_2,g_7),(b_8,g_5),(b_6,g_4)$ |
| $\rho_6$ | $(b_3,g_3),(b_1,g_2)$ |
| $\rho_7$ | $(b_1,g_5),(b_5,g_1),(b_7,g_6)$ |
| $\rho_8$ | $(b_5,g_6),(b_8,g_4),(b_6,g_7)$ |
| $\rho_9$ | $(b_2,g_5),(b_7,g_8),(b_4,g_3)$ |
| $\rho_{10}$ | $(b_1,g_1),(b_5,g_2)$ |

These ten rotations form the rotation poset $(\mathcal{R},\succeq)$ whose Hasse diagram is given in Figure 2. As shown the rotation $\rho_1$ is the unique maximal rotation, the rotation exposed at the boy-optimal stable matching $1 = M_1$. On the other hand, $\rho_8,\rho_9$ and $\rho_{10}$ are the minimal rotations. These rotations lead to the boy-pessimal (girl-optimal) stable matching $0 = M_{23}$.

Applying these rotations in the appropriate order allows us to generate all 23 stable matchings given in Table 3.

These stable matchings then form the stable matching lattice $(\mathcal{M},\succeq)$ whose Hasse diagram is illustrated in Figure 3. In this diagram, each edge is labelled by the rotation that transforms the
upper stable matching into the lower stable matching. For example, the rotation $\rho_0$ is exposed at $M_{15}$ and applying it produces the matching $M_{18}$; similarly, the rotation $\rho_4$ is exposed at $M_5$ and induces $M_6$.

Recall, by the fundamental theorem for finite distributive lattices, the stable matching lattice
$(\mathcal{M}, \succeq)$ has an auxiliary poset whose order ideals, ordered by inclusion, form $\mathcal{M}$. We claimed that this auxiliary poset is the rotation poset $\mathcal{P} = (\mathcal{R}, \succeq)$. By inspection of Figure 2 and Figure 3, the reader may verify that this is indeed the case for this stable matching instance. In particular, we can see the correspondence between minimal stable matchings in $\mathcal{M} \setminus \{0\}$ and minimal rotations. In this case there are three such minimal matchings, namely $\{M_{20}, M_{21}, M_{22}\}$ and three minimal rotations, namely, $\{\rho_8, \rho_9, \rho_{10}\}$.

Finally, Theorem 3.1, tells us that the minimum winning coalition has cardinality equal the number of minimal rotations. In this instance, these three rotations consist of the girls $\{g_1, g_2\}$, $\{g_3, g_5, g_8\}$ and $\{g_4, g_6, g_7\}$, respectively. A minimum winning coalition must contain exactly one element of each of these groups. For example, the three girls $\{g_1, g_3, g_4\}$ form a minimum winning coalition and allow us to descend all the way from 1 to 0. If we select only one girl we descend the lattice only as far down as $M_{17}$, $M_{18}$ or $M_{19}$. If we select two girls we can descend only as far down as $M_{20}$, $M_{21}$, or $M_{22}$. 

Figure 3: The Hasse Diagram of the Stable Matching Lattice $(\mathcal{M}, \succeq)$. 
5 The Random Matching Model

For the rest of the paper we use the random matching model which was first studied by Wilson [20] and subsequently examined in detail by Knuth, Pittel and coauthors [12, 16, 13, 17]. Here the preference ranking of each boy and each girl is drawn uniformly and independently from the symmetric group $S_n$. Specifically, each preference ranking is a random permutation of the set $[n] = \{1, 2, \ldots, n\}$.

We may now state the two main results of the paper. First, in the random matching model, the expected cardinality of the minimum winning coalition is $O(\log n)$.

**Theorem 5.1.** In the random matching model, the expected cardinality of the minimum winning coalition $F$ is

$$\mathbb{E}(|F|) = \frac{1}{2} \log(n) + O(\log \log n)$$

So the minimum winning coalition is small. Surprisingly, in sharp contrast, our second result states that a random coalition must contain nearly every girl if it is to form a winning coalition with high probability. Equivalently:

**Theorem 5.2.** In the random matching model, $\forall \varepsilon > 0$, $\exists \delta(\varepsilon) > 0$ such that for a random coalition $F$ of cardinality $(1 - \varepsilon) \cdot n$ the probability that $F$ is not a winning coalition is at least $\delta(\varepsilon)$.

To prove these results, recall Theorem 3.1 which states that a winning coalition $F$ must intersect each minimal rotation in the rotation poset $(\mathcal{R}, \geq)$. Thus, for Theorem 5.1 it suffices to show that the expected number of minimal rotations is $O(\log n)$. To show Theorem 5.2 we must lower bound the probability that a randomly chosen coalition of girls contains at least one girl in each minimal rotation. Our approach is to show the likelihood of a small cardinality minimal rotation is high. In particular, we prove there is a minimal rotation containing exactly two girls with constant probability. It immediately follows that a random coalition must contain nearly all the girls if it is to be a winning coalition with high probability.

5.1 Overview of the Proofs

So our proofs require that we study the set of minimal rotations in the random matching model. The following two “tricks” will be useful in performing our analyses. First, instead of minimal rotations we may, in fact, study the set $\mathcal{R}^\text{max}$ of maximal rotations, that is the rotations that are exposed at the boy-optimal stable matching $1$. This is equivalent because Theorem 2.3 tells us that the inverse lattice $(\mathcal{M}, \leq)$ is the stable matching lattice ordered according to the preferences of the girls. This symmetry implies that the behaviour of minimal rotations is identical to the behaviour of maximal rotations as the maximal rotations of one lattice are the minimal rotations of the other. But why is the switch to maximal rotations from minimal rotations helpful? Simply put, as we are using the boy proposal version of the deferred acceptance algorithm, we obtain the boy-optimal stable matching and, consequently, it is more convenient to reason about the rotations exposed at $1$, that is the maximal rotations.

Second, it will be convenient to view the deferred acceptance algorithm with random preferences in an alternative manner. In particular, instead of generating the preference rankings in advance, we may generate them dynamically. Specifically, when a boy $b$ is selected to make a proposal he asks a girl $g$ chosen uniformly at random. If $b$ has already proposed to $g$ then this proposal is immediately
rejected; such a proposal is termed redundant. Meanwhile, $g$ maintains a preference ranking only for the boys that have proposed to her. Thus if this is the $k$th distinct proposal made to girl $g$ then she assigns to $b$ a rank chosen uniformly at random among $\{1, \ldots, k\}$. In particular, in the deferred acceptance algorithm $g$ accepts the proposal with probability $1/k$. As explained by Knuth et al. [13], this process is equivalent to randomly generating the preference rankings independently in advance. Furthermore, recall from Theorem 2.2 that the deferred acceptance algorithm will output the boy-optimal stable matching regardless of the order of proposals. It follows that, for the purposes of analysis, we may assume the algorithm selects the unmatched boy with the lowest index to make the next proposal.

So our task now is to investigate the properties of maximal rotations, that is directed cycles in the rotation graph $H(1)$. Intuitively, this relates to the study of directed cycles in random graphs with out-degrees exactly one. But there is one major problem. In random graphs the choice of out-neighbour is independent for each vertex. But in the rotation graph $H(1)$ this independence is lost. In particular, the arcs in $H(1)$ share intricate dependencies and specifically depend on who made and who received each proposal in obtaining the boy-optimal stable matching $1$. Moreover, a vertex may even have out-degree zero in $H(1)$. Essentially, the remainder of paper is devoted to showing that the myriad of dependencies that arise are collectively of small total consequence. It will then follow that the expected number or maximal rotations and the minimum cardinality of a maximal rotation both behave in a predictable manner, similar to that of directed cycles in random graphs with out-degrees exactly one. Namely, the expected number of cycles is close to $\frac{\log n}{2}$ and the existence a cycle of size two with constant probability $[4]$.

Consequently, to study maximal rotations we must consider $H(1)$. We do this via a two-phase approach. In the first phase we calculate the boy-optimal stable matching $1$, without loss of generality, $1 = \{(b_1, g_1), (b_2, g_2), \ldots, (b_n, g_n)\}$. This of course can be found by running the boy-proposal deferred acceptance algorithm. In the second phase, we calculate the rotation graph $H(1)$. But, as explained in Section 2.3 and illustrated in the example of Section 4, we can find the rotations by running the boy-proposal deferred acceptance algorithm longer.

In fact, to calculate (i) the expected number of maximal rotations and (ii) the probability that there is a maximum rotation of cardinality 2, we will not need the entire rotation graph $H(1)$ only subgraphs of it. Moreover, the subgraphs we require will be different in each case. Consequently, the second phases required to prove Theorem 5.1 and Theorem 5.2 will each be slightly different. These distinct second phases will be described in detail in Section 6 and Section 7, respectively. They both, however, share fundamental properties which will be exploited in shortening the subsequent proofs.

### 5.2 A Technical Tool for Counters

Before describing the two algorithms, we present a technical lemma that we will use repeated in analyzing the deviations that arise in their application. To formalize the lemma, we require the notion of a state. The state of the algorithm at any point is the record of all the (random) choices made so far: the sequence of proposals and the preference rankings generated by the girls. Thus we are working in the probability space $(\Omega, P)$ of all possible states $\Omega$ of the algorithm and the probabilities of reaching them.

We index the intermediate states of the algorithm by the number of proposals made to reach it. Let $\Omega_t$ denote the set of all possible states of the procedure after $t$ proposals. Thus $\Omega_t$ can be thought of as a partition of $\Omega$, and the partition $\Omega_{t+1}$ refines the partition $\Omega_t$ for any $t$. A random
variable \(X_t\) is \(\Omega_t\)-measurable if \(X_t\) is determined by the algorithm state after \(t\) proposals, that is \(X\) is constant on each part of \(\Omega_t\). We say that a sequence \((X_t)_{t \geq 0}\) of random variables is a counter if \(X_t\) is \(\Omega_t\)-measurable and \(X_t - X_{t-1} \in \{0,1\}\). Thus counters count the number of certain events occurring over the course of the algorithm. As an example, the number of successful proposals among the first \(t\) proposals is a counter.

Our main tool is Lemma 5.3 below which is used to control large deviations of counters. Let \(B_{k,p}\) be a random variable which follows a binomial distribution with parameters \(k\) and \(p\). We say that a collection of states \(\mathcal{G}\) is monotone if for every state \(S \not\in \mathcal{G}\) we have \(S' \not\in \mathcal{G}\) for every state \(S'\) that can be reached from \(S\). For example, the collection of states in which every girl received at most one proposal is monotone. Let \(\{S_t | t \in \mathbb{N}\}\) be the sequence of random variables corresponding to the state of the algorithm at time \(t\).

**Lemma 5.3.** Let \(\mathcal{G}\) be a monotone collection of states and let \((X_t)_{t \geq 0}\) be a counter.

(a) If \(P(X_{t+1} - X_t = 1 | S_t = S_t') \geq p\) for every state \(S_t' \in \Omega_t \cap \mathcal{G}\), for any \(t' \in [t,t+k]\), then, for any \(\lambda \geq 0\) and any \(k \geq 1\),

\[
P((X_{t+k} - X_t \leq \lambda) \land (S_{t+k-1} \in \mathcal{G}) | S_t = S_t') \leq P(B_{k,p} \leq \lambda).
\]

(b) If \(P(X_{t+1} - X_t = 1 | S_t = S_t') \leq p\) for every state \(S_t' \in \Omega_t \cap \mathcal{G}\), for any \(t' \in [t,t+k]\), then, for any \(\lambda \geq 0\) and any \(k \geq 1\),

\[
P((X_{t+k} - X_t \geq \lambda) \land (S_{t+k-1} \in \mathcal{G}) | S_t = S_t') \leq P(B_{k,p} \geq \lambda).
\]

This implies that if we can bound the probability of the counter being incremented tightly enough then the behaviour of the counter will be similar to the behaviour of a binomial random variable. Evidently, this will be useful because binomial random variables are much simpler to work with.

**Proof.** We prove (a) by induction on \(k\). The base case \(k = 1\) is immediate. For the induction step, note that if \(S_t \not\in \mathcal{G}\) then the left side of (a) is zero. Thus we may assume \(S_t \in \mathcal{G}\) and hence \(P(X_{t+1} - X_t = 1 | S_t) \geq p\). By the induction hypothesis, we then have

\[
P(X_{t+k} - X_{t+1} \leq \lambda - 1 | (X_{t+1} - X_t = 1) \land S_t = S_t') \leq P(B_{k-1,p} \leq \lambda - 1)
\]

and

\[
P(X_{t+k} - X_{t+1} \leq \lambda | (X_{t+1} - X_t = 0) \land S_t = S_t') \leq P(B_{k-1,p} \leq \lambda).
\]

Since \(P(B_{k-1,p} \leq \lambda - 1) \leq P(B_{k-1,p} \leq \lambda)\), combining these three inequalities gives

\[
P(X_{t+k} - X_t \leq \lambda | S_t = S_t') = P(X_{t+1} - X_t = 1 | S_t = S_t') \cdot P(B_{k-1,p} \leq \lambda - 1)
\]

\[
+ P(X_{t+1} - X_t = 0 | S_t = S_t') \cdot P(B_{k-1,p} \leq \lambda)
\]

\[
\leq pP(B_{k-1,p} \leq \lambda - 1) + (1 - p)P(B_{k-1,p} \leq \lambda)
\]

(1)

where we obtain (1) by noting the following:

\[
P(B_{k-1,p} \leq \lambda - 1) \leq P(B_{k-1,p} \leq \lambda)
\]

\[
\implies p \cdot [P(B_{k-1,p} \leq \lambda - 1) - P(B_{k-1,p} \leq \lambda)] \leq 0
\]

\[
\implies pP(B_{k-1,p} \leq \lambda - 1) + (1 - p)P(B_{k-1,p} \leq \lambda) \leq P(B_{k,p} \leq \lambda).
\]

The proof of (b) is completely analogous. □
In our subsequent analyses we will combine Lemma 5.3 with the following well-known Chernoff bounds that control deviations of $B_{k,p}$ from the mean.

**Lemma 5.4.** For $0 \leq \delta \leq 1$,

\[
P(B_{k,p} \geq (1 + \delta)pk) \leq \exp\left(-\frac{\delta^2 pk}{3}\right) \quad \text{and} \quad P(B_{k,p} \leq (1 - \delta)pk) \leq \exp\left(-\frac{\delta^2 pk}{2}\right).
\]

### 6 Minimum Winning Coalitions

In this section, we will evaluate the expected cardinality of the minimum winning coalition. Recall, it suffices to find the expected number of directed cycles, $R^{\text{max}}$, in the rotation graph $H(1)$. To do this, it will be useful to describe the cardinality of $R^{\text{max}}$ in a more manipulable form. Specifically, for any boy $b_i$ define a variable

\[
Z_i = \begin{cases} 
\frac{1}{|R|} & \text{if } b_i \text{ is in a maximal rotation } R \\
0 & \text{if } b_i \text{ is not in a maximal rotation}
\end{cases}
\]

Then we obtain that:

\[
|R^{\text{max}}| = \sum_{R \in R^{\text{max}}} 1 = \sum_{R \in R^{\text{max}}} \sum_{(b,g) \in R} \frac{1}{|R|} = \sum_{i=1}^{n} Z_i
\]

By linearity of expectation, the expected cardinality of the minimum winning coalition $F$ is

\[
E(|F|) = E(|R^{\text{max}}|) = E\left(\sum_{i=1}^{n} Z_i\right) = \sum_{i=1}^{n} E(Z_i)
\]

As discussed in Section 5, the difficulty in computing $E(|F|)$ is the myriad of dependencies that arise in the formation of the rotations in $R^{\text{max}}$. Equation 2 is extremely useful in this regard. To quantify the dependency effects, rather than count expected rotations directly, it allows to focus simply on computing $E(Z_i)$.

#### 6.1 Generating Maximal Rotations from the Rotation Graph

Ergo, our task now is to evaluate $E(Z_i)$. For this we study a two-phase randomized algorithm, henceforth referred to as the algorithm, for generating the potential maximal rotation containing a given boy. The first phase computes the boy-optimal stable matching $1 = \{(b_1, g_1), (b_2, g_2), \ldots, (b_n, g_n)\}$. In the second phase we use a variation of the deferred acceptance algorithm to generate arcs in (a subgraph of) the rotation graph and generate a random variable $Z$.

The second phase starts with a randomly selected boy $i_1$ who makes uniformly random proposals until the first time he proposes to a girl $g_j$ who prefers him over her partner $b_j$ in the boy-optimal stable matching. The boy $b_j$ will make the next sequence of proposals. The process terminates if we find a maximal rotation. Moreover, if this rotation is completed because girl $g_i$ receives and accepts a proposal then we have found a maximal rotation containing boy $i_1$. In this case we also update $Z$. Formally, we initialize the second-phase by:

- Choose $i_1$ from $\{1, 2, \ldots, n\}$ uniformly at random.
• Initialize the potential cycle in the rotation digraph containing $i_1$ by setting $R = [i_1]$.

Once $R = [i_1, \ldots, i_k]$ is found, we generate the arc of the rotation digraph emanating from $i_k$, as follows.

• Let boy $b_{i_k}$ make uniformly random proposals until the first time he proposes to a girl $g_j$ such that $g_j$ ranks $b_{i_k}$ higher than $b_j$. That is, $g_j$ ranks $b_{i_k}$ higher than her pessimal stable partner.
  
  - If $j \notin R$ then we set $i_{k+1} = j$, $R = [i_1, \ldots, i_k, i_{k+1}]$, and recurse.
  - If $j \in R$ then we terminate the procedure. We set $Z = \frac{1}{|R|}$, if $j = i_1$, and $Z = 0$, otherwise.

• If, instead, boy $b_{i_k}$ gets rejected by all the girls then the vertex $i_k$ has no-outgoing arcs in the rotation graph. Thus, $b_{i_k}$ belongs to no maximal rotation, so we terminate the procedure and set $Z = 0$.

We emphasize that as the second phase runs, we do not change any assigned partnerships. Specifically, when a girl receives a proposal we always compare her rank for the proposing boy to the rank of her pessimal partner, regardless of any other proposals she may have received during the second phase. Now $Z = Z_{i_1}$ where $i_1$ was chosen uniformly at random. The next lemma is then implied by (2) as the expectation of $Z$ is the average of the expectations of $Z_{i_1}$.

**Lemma 6.1.** $E(|F|) = n \cdot E(Z)$ where $Z$ is the random variable generated by the algorithm.

Recall, $b_{i_1}$ is in a maximal rotation if and only if the rotation graph of the boy optimal stable matching has a cycle containing $b_{i_1}$. Observe that every connected component of a directed graph in which each vertex has out-degree 1 contains exactly one cycle. Hence, if we find a cycle in the same connected component as $b_i$ but which does not contain him then $b_{i_k}$ is not in a maximal rotation. Then, since $|F| = \sum_{j=1}^{n} Z_{ij}$, we get $E(|F|) = \sum_{j=1}^{n} E(Z_{ij}) = n \cdot E(Z)$.

### 6.2 Properties of the Two-Phase Algorithm

We now present a series of properties that arise with high enough probability during the two-phase process. In particular, the process does not deviate too far from its expected behaviour. For example, the running time of each phase is not much longer than expected, no girl receives too many proposals, and no boy makes too many proposals. To formalize this, let $T_1$ and $T_2$ be the number of proposals made in the first and second phases, respectively, and let $T = T_1 + T_2$. Further, let a run be a sequence of consecutive proposals made by the same boy in the same phase. Now consider the following properties that may apply to a state:

I. The algorithm has not terminated.

II. If the algorithm is in the first phase then $t \leq 5n \log n$. If the algorithm is in the second phase then $T_1 \leq 5n \log n$

III. If the algorithm has not found a rotation yet then $t \leq T_1 + \sqrt{n} \log^3 n$.

IV. Each girl has received at most $21 \log n$ proposals.

V. Each boy started at most $21 \log n$ runs.
VI. Each run contained at most $111 \log^2 n$ proposals.

VII. Each boy has made at most $\log^4 n$ proposals.

Informally, we want the states of the algorithm to satisfy these property because they imply that:

- No girl receives too many proposals compared to the others girls. Consequently, the event of having already accepted a proposal in the second phase does not impact significantly the probability of accepting another proposal later.

- No boy makes too many proposals. Consequently, we do not need to worry about redundant proposals.

- The second phase is significantly shorter than the first phase. Consequently, the probability of a proposal being accepted at the start of the second phase and the probability of a proposal being accepted at the end of the second phase are very similar.

Let $G$ be the set of all states that satisfy properties I to VII. We call these good states. Any state that is not good is bad. Clearly, $G$ is monotone. Let $G^*$ denote the event $S_{T-1} \in G$, that is, the event that the algorithm is in a good state the period before it terminates. Let $G^*$ be the complement of $G$.

We remark that, for technical reasons, we will assume the second-phase terminates if $n \log n$ proposals are made during that phase. This assumption is superfluous here by condition III, which states the second phase has at most $\sqrt{n \log^3 n}$ proposals. However, the assumption is useful as it will allow the following lemma to also apply for the modified second-phase algorithm that we use in Section 7.

**Lemma 6.2.** For $n$ sufficiently large, $P(G^*) \geq 1 - O(n^{-4})$.

**Proof.** It suffices to show the probability is $O(1/n^4)$ of reaching a state $S_k \in \Omega_k$ such that (i) the algorithm has not yet terminated, (ii) $S_k$ is bad, and (iii) all the states preceding $S_k$ are good.

Note that, for $n$ sufficiently large, conditions II and our bound on the length of the second phase imply that $k \leq 6n \log n$ for any such state. Furthermore, again for sufficiently large $n$, conditions V and VI together imply VII; thus, VII cannot be the only condition violated by $S_k$. Hence, it suffices to verify that the probability of reaching such a state violating one of the conditions II-VI is small.

First consider condition II. Recall the first phase terminates when every girl has received a proposal. So if $S_k$ violates II then $k \geq 5n \log n$ and at least one girl still has not received a proposal. By definition, each proposal is directed at girl $g$ with probability $1/n$, for each $g$. So, by Lemma 5.3 applied to the counter $(X_{g,k})_{k \geq 0}$, where $X_{g,k}$ is the number of proposals received by $g$ by the time $k$,

$$P(X_{g,k} = 0) \leq \left(1 - \frac{1}{n}\right)^k \leq e^{-k/n} \leq \frac{1}{n^5}.$$  

Thus, by the union bound, $S_k$ violates II with probability at most $1/n^4$, as desired.

Next consider condition IV. If $s = \lceil 21 \log n \rceil$ then the probability that a girl $g$ received at least $s$ proposals is at most

$$P(X_{g,k} \geq s) \leq \binom{k}{s} \left(\frac{1}{n}\right)^s \leq \left(\frac{ek}{sn}\right)^s \leq \left(\frac{6en \log n}{21 \log n} n\right)^s \leq \left(\frac{6e}{21}\right)^s \leq n^{21 \log(6e/21)} \leq \frac{1}{n^{5.3}}.$$
and so, by the union bound, $S_k$ violates IV with probability at most $1/n^4$.

The proof of V is similar. Take a boy $b$. Apart from his first run (and possibly the first state of second phase), $b$ can only start a run if the girl $g$ he had been matched to received a proposal in the previous round. This occurs with probability at most $1/n$ conditioned on the previous state. Thus, analogously to the argument above, $S_k$ violates condition V with probability at most $1/n^4$.

Now consider VI and set $s = 111 \log^2 n$. For sufficiently large $n$, the proposal following a good state is non-redundant with probability considerably greater than $21/22$ by V. Because each girl has received at most $21 \log n$ proposals by IV, the probability that a proposal is accepted conditioned on the previous good state is at least $1/22 \log n$. By Lemma 5.3 applied to the counter $X_t$ equal to the number of proposals accepted by time $t$,

$$P \left( (X_k - X_{k-s} = 0) \wedge (S_{k-s} \in G) | S_{k-s} = S_{k-s} \right) \leq \left( 1 - \frac{1}{22 \log n} \right)^s \leq \exp \left( -\frac{s}{22 \log n} \right) \leq \frac{1}{n^{5.04}}$$

for any good state $S_{k-s}$. In particular, for any such state $S_{k-s}$, the probability is at most $1/5n^{0.4}$, for sufficiently large $n$. By the union bound taken over possible choices of $k \leq 5n \log n$, the probability that we reach $S_k$ violating VI is at most $1/n^4$.

Set $\ell = \lceil \frac{1}{2} \sqrt{\frac{n}{\log n}} \rceil$. If $S_k$ violates III then the second phase been running for at least $2\ell$ steps without finding a cycle, and we have previously reached a state $S \in \Omega_{k-\ell}$ such the second phase contained at least $\frac{\ell}{111 \log^2 n}$ runs before $S$ due to VI. Therefore, the potential cycle $R$ generated in $S$ may contain at least $\frac{\ell}{111 \log^2 n}$ elements. In each subsequent step starting in a good state, the probability that a non-redundant proposal is made to a girl $g_i$ with $i \in R$ is at least $\frac{\ell}{n \log^3 n}$. Further, such a proposal is then accepted, terminating the process, with probability at least $\frac{1}{21 \log n}$. However, to reach a state $S_k$ the algorithm must continue for at least $\ell$ more steps. By Lemma 5.3, the probability that this happens starting with any given state $S$ as above is at most

$$\left( 1 - \frac{\ell}{21n \log^4 n} \right)^\ell \leq \exp \left( -\frac{\ell^2}{21n \log^4 n} \right) \leq \frac{1}{n^6}$$

and thus the probability of reaching $S_k$ violating III is at most $1/n^4$.

So, with high probability, we are in a good state the period before the algorithm terminates. It follows that the magnitude of the expected number of maximal rotations can be evaluated by consideration of good states.

### 6.3 The Expected Cardinality of the Minimum Winning Coalition

Now, to calculate the expected number of maximal rotations we must analyze in more detail the second phase of the algorithm. In particular, this section is devoted to the proof of the following lemma.

**Lemma 6.3.** Let $S_* = S_{T_1}$ be the terminal state of the first phase. If $P (G_* | S_* = S_*) \leq \frac{1}{n^4}$ then

$$\mathbb{E} (Z | S_* = S_*) = \frac{\log n}{2n} + O \left( \frac{\log \log n}{n} \right).$$
Before embarking on the proof of Lemma 6.3, we remark that that our first main result, Theorem 5.1, readily follows from it via Lemmas 6.1 and 6.2. It is also worth noting that III implies that the second phase has at most $\sqrt{n} \log^3 n$ proposals when $G_s$ occurs, due to the fact that we stop once we find our first cycle.

**Proof of Theorem 5.1 (modulo Lemma 6.3).** Let $G_{ss}$ denote the set of the terminal states $S_s$ of the first phase of the algorithm satisfying $P(G_s | S_s = S_s) \leq \frac{1}{n}$. Then $P(S_s \notin G_{ss}) = O(1/n)$ by Lemma 6.2. Since $0 \leq Z \leq 1$, by Lemmas 6.1 and 6.3 we have

$$
\mathbb{E}(|F|) = n\mathbb{E}(Z) = n\mathbb{E}(Z | S_s \in G_{ss})(1 - P(S_s \notin G_{ss})) + n\mathbb{E}(Z | S_s \notin G_{ss})P(S_s \notin G_{ss}) = n\mathbb{E}(Z | S_s \in G_{ss}) + nP(S_s \notin G_{ss})(\mathbb{E}(Z | S_s \notin G_{ss}) - \mathbb{E}(Z | S_s \in G_{ss})) = \frac{1}{2}\log n + O(\log^2 n) + O(1) \quad \square
$$

So let’s prove Lemma 6.3. For the remainder of the section we fix $S_s = S_s$ satisfying the conditions of the lemma. Let $\rho_i$ be the number of non-redundant proposals received by girl $g_i$ in $S_s$. Set $\rho = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\rho_i + 1}$. As $S_s$ is good, we have $\rho_i \leq 21 \log n$ for every girl $g_i$; so, $\rho \geq \frac{1}{22 \log n}$. We evaluate $\mathbb{E}(Z)$ separately for every choice of initial vertex $i_1$ of $R$ in the following lemma:

**Lemma 6.4.** For every $1 \leq i \leq n$ we have

$$
\mathbb{E}(Z | S_s = S_s \land (i_1 = i)) = \frac{1}{n\rho(\rho_i + 1)} \left( \frac{1}{2}\log n + O(\log^2 n) \right).
$$

Since $\frac{1}{n} \sum_{i=1}^{n} \frac{1}{\rho_i + 1} = \frac{1}{n\rho} \sum_{i=1}^{n} \frac{1}{\rho_i + 1} = \frac{n}{\rho} = \frac{1}{n}$, this lemma implies Lemma 6.3. To prove Lemma 6.4, we may assume that we reached the state $S_s$, chose $i_1 = i$, and that the probabilities of the subsequent events are scaled accordingly. Note that, by Lemma 6.2, we have $P(G_s) \geq 1 - 1/n^2$.

We relabel the states of our process $(S_0 = S_s, S_1, \ldots, S_{t}, \ldots)$, so that $S_t$ is the state attained after $t$ proposals have been made in the second phase. Let $R_t$ denote the (random) set $R$ generated in the state $S_t$. Let $X_t = |R_t|$ be the associated counter. First we show that any proposal made after a good state increases $X_t$ with probability close to $\rho$.

**Lemma 6.5.** For any good state $S_t$, we have:

$$
P(X_{t+1} = 1 | S_t = S_t) \in \left[ \rho - n^{-1/3}, \rho \right].
$$

**Proof.** Let $\rho_{i,t'}$ be the number of non-redundant proposals received by girl $g_i$ in a state $S_{t'}$ preceding $S_t$, and set

$$
\rho(t') = \frac{1}{n} \sum_{i \in R_{t'}} \frac{1}{\rho_{i,t'} + 1}.
$$

Then $\rho(0) \geq \rho - 1/n$ and $\rho(t' + 1) \geq \rho(t') - 1/n$, for every $0 \leq t' \leq t - 1$, as $R_{t'}$ increases by at most one vertex in any step.

Let $b$ be the boy making the proposal following $S_t$, and let $B$ be the set of girls $b$ has already proposed to. Then $|B| \leq \log^4 n$, as $S_t$ is good. The probability that the next proposal is accepted by a girl not in $R_t$, thus increasing $X_t$, is then

$$
\frac{1}{n} \sum_{i \in R_t \setminus B} \frac{1}{\rho_{i,t} + 1} \leq \rho,
$$

20
immediately implying the upper bound. On the other hand, this probability is lower bounded by

\[ \rho(t) - \frac{|B|}{n} \geq \rho - \frac{t + 1}{n} - \frac{\log^4 n}{n} \geq \rho - n^{-1/3} \]

where the last bound holds as \( S_t \) is good, and so \( t \leq \sqrt{n} \log^3 n \).

\[ \text{Lemma 6.6.} \]

1. For \( t \geq \log^5 n \), \( P \left( \left| X_t - \rho t \right| \geq \frac{\rho t}{\log n} \right) \land (S_{t-1} \in \mathcal{G}) \leq \frac{1}{n^2} \).

2. For \( t \geq \frac{400 \log \log n}{\rho} \), \( P \left( X_t \leq \frac{1}{2} \rho t \right) \land (S_{t-1} \in \mathcal{G}) \leq \frac{1}{2 \log^* n} \).

\[ \text{Proof.} \] Let \( \delta \geq \log^{-1} n \). Recall that \( \rho \geq \frac{1}{22} \log^{-1} n \), and so

\[ \rho - n^{-1/3} \geq \left( 1 - \frac{\delta}{2} \right) \rho \geq \frac{1}{25} \log^{-1} n. \]

Combining Lemma 5.4 with Lemma 5.3 (where \( t = 0 \) and \( k = t \)) and Lemma 6.5 gives

\[ P \left( X_t \leq (1 - \delta) \rho t \land (S_{t-1} \in \mathcal{G}) \right) \leq P \left( B_{t, \rho - n^{-1/3}} \leq (1 - \delta) \rho t \right) \]

\[ \leq P \left( B_{t, \rho - n^{-1/3}} \leq \left( 1 - \frac{\delta}{2} \right) (\rho - n^{-1/3}) t \right) \]

\[ \leq \exp \left( -\frac{1}{2} \left( \frac{\delta}{2} \right)^2 (\rho - n^{-1/3}) t \right). \]  

(3)

If \( \delta = \log^{-1} n \) and \( t \geq \log^5 n \) then (3) is upper bounded by \( \exp \left( -\frac{\log^2 n}{200} \right) < \frac{1}{2n^2} \). Meanwhile for \( \delta = 1/2 \) and \( t \geq \frac{C \log \log n}{\rho} \), the last term of (3) can instead be upper bounded by \( \log^{-C \rho} n \). This proves the stated bounds on lower deviation.

The inequality

\[ P \left( \left( X_t \geq \left( 1 + \frac{1}{\log n} \right) \rho t \right) \land (S_{t-1} \in \mathcal{G}) \right) \leq \frac{1}{2n^2} \]

for \( t \geq \log^5 n \) is derived in the same manner. \[ \square \]

We also need the following two easy lemmas.

\[ \text{Lemma 6.7.} \] We have \( P \left( T_2 \leq \sqrt{n} / \log n \right) \leq 1/2 \log^2 n \).

\[ \text{Proof.} \] Note that \( X_t \leq t + 1 \), and so for any \( t \leq \sqrt{n} \log^{-1} n - 1 \), the probability that the next proposal is directed to a girl with index in \( R_t \) is at most \( \frac{1}{\sqrt{n} \log n} \). Therefore, the probability that the second phase terminates after exactly \( t \) proposals is at most \( \frac{1}{\sqrt{n} \log n} \) for every such \( t \). The lemma follows by applying the union bound. \[ \square \]

Let \( \{ b_j \}_{j \in J} \) be the set of boys who have proposed to \( g_i \) by the end of the first phase.

\[ \text{Lemma 6.8.} \] The probability that at least one of the first \( \sqrt{n} \log^3 n \) proposals of the second phase is directed to a girl \( g_j \) with \( j \in J \) is at most \( n^{-1/3} \).
Proof. As $S_*$ is good, we have $|J| \leq 21 \log n$. Thus, this lemma follows by applying the union bound analogously to Lemma 6.7. We omit the details.

**Proof of Lemma 6.4.** Let’s begin by proving the lower bound. Let $L_t$ denote the collection of states $S_t$ such that

- $\log^5 n \leq t \leq \sqrt{n} / \log n$
- $S_t \in G$, in particular the algorithm has not yet terminated,
- $X_t \leq \left(1 + \frac{1}{\log n}\right) \rho t$,
- every girl $g_j$ with $j \in J$ received no proposal in the second phase so far.

It follows from Lemmas 6.6, 6.7 and 6.8 that $P(S_t \notin L_t) \leq \log^{-5} n$, for any $400 \rho^{-1} \log \log n \leq t \leq \sqrt{n} \log^{-1} n$. As any state $S_t \in L_t$ is good and satisfies $t \geq \log^5 n$, the boy $b_i$ has already finished the run which started the second phase. Moreover, no other boy who has previously proposed to $g_i$ has lost his partner and had an opportunity to make a proposal. Thus, if the next proposal is directed at $g_i$, which happens with probability $1/n$, it is non redundant. Such a proposal is accepted with probability $1/\rho_i + 1$. In such a case, the algorithm terminates and outputs $Z = 1/X_t$. By Lemma 6.5, considering only the contributions of outcomes when the process terminates immediately following a state in $L_t$ we get the following lower bound on the expected value of $Z$.

$$E(Z | S_\ast = S_\ast \land (i_1 = i)) \geq \left(1 - \frac{1}{\log^2 n}\right) \frac{1}{n(\rho_i + 1)} \sum_{t=\log^5 n}^{\sqrt{n} \log n} \frac{1}{t} \left(1 + \frac{1}{\log n}\right) \rho t$$

$$= \left(1 - \frac{1}{\log n}\right) \frac{1}{n\rho(\rho_i + 1)} \sum_{t=\log^5 n}^{\sqrt{n} \log n} \frac{1}{t}$$

$$\geq \left(1 - \frac{2}{\log n}\right) \frac{1}{n\rho(\rho_i + 1)} \left(\log \left(\frac{\sqrt{n}}{\log n}\right) - \log(\log^5 n) - O(1)\right)$$

$$= \frac{1}{n\rho(\rho_i + 1)} \left(\frac{1}{2} \log n - O(\log \log n)\right)$$

Next we prove the upper bound. Let $U_t$ denote the collection of states $S_t$ such that

- $400 \rho^{-1} \log \log n \leq t \leq \sqrt{n} \log^3 n$
- $S_t \in G$ or the algorithm has terminated,
- $X_t \geq \frac{1}{2} \rho t$.
- $X_t \geq \left(1 - \frac{1}{\log n}\right) \rho t$, if $t \geq \log^5 n$,

It follows from Lemma 6.6 that:

$$\begin{cases} 
P(S_t \notin U_t) \leq \log^{-5} n \text{ for } 400 \rho^{-1} \log \log n \leq t \leq \log^5 n \n P(S_t \notin U_t) \leq \frac{1}{n} \text{ for } \log^5 n \leq t \leq \sqrt{n} \log^3 n
\end{cases}$$

22
Noting that the process terminates and outputs $Z = 1/X_t$ immediately following any given state $S_t$ with probability at most $\frac{1}{n(\rho_i + 1)}$, we obtain the desired upper bound on the expected value of $Z$, as follows:

$$
\mathbb{E}(Z|S_* = S_* \land (i_1 = i)) \leq \frac{1}{n(\rho_i + 1)} \sum_{t \leq 400\rho^{-1} \log \log n} \frac{1}{n(\rho_i + 1)} \sum_{t = \log^5 n}^{\log^5 n} \left( \frac{1}{\log^5 n} + \frac{2}{\rho t} \right) 
$$

$$
+ \frac{1}{n(\rho_i + 1)} \sum_{t = \log^5 n}^{\sqrt{\pi \log^3 n}} \left( \frac{1}{n} + \frac{1}{\rho t} \right) + P(T_2 \geq \sqrt{n} \log^3 n)
$$

$$
= \frac{1}{n(\rho_i + 1)} \sum_{t \leq \sqrt{\pi \log^3 n}} \frac{1}{\rho t} + O\left( \frac{\log \log n}{n \rho(\rho_i + 1)} \right)
$$

$$
= \frac{1}{n \rho(\rho_i + 1)} \left( \frac{1}{2} \log n + O(\log \log n) \right).
$$

This completes the proof of Lemma 6.4 and thus of Lemma 6.3. Our first main result, Theorem 5.1, immediately follows. \qed

7 Random Winning Coalitions

In this section, we consider the case where the girls in the coalition are themselves randomly selected. Our task now is to prove that almost every girl must be selected if we wish to obtain a winning coalition asymptotically almost surely. To do this, it will suffice to prove that there is a maximal rotation of cardinality two with constant probability.

7.1 Generating Maximal Rotations from the Rotation Graph

Let $Z'$ be a random variable counting the number of maximal rotations of cardinality two. Again, to analyze $Z'$ we use a two-phase algorithm. The first phase is the same as before. We simply generate the boy-optimal stable matching $1 = \{(b_1, g_1), (b_2, g_2), \ldots, (b_n, g_n)\}$. But the second phase is slightly different. Previously we had to evaluate the expected number of maximal rotations and, to achieve that, it sufficed to end the second phase once we had found one rotation. Now, because we are interested in maximal rotations of cardinality two we will extend the second phase and terminate only when and if we find rotation of cardinality two.

So now in the second phase we use the following algorithm to generate the random variable $Z'$, initialized at 0:

- Choose $i_1$ from $\{1, 2, \ldots, n\}$ uniformly at random.
- Initialize the set of indices of boys who have made proposals in the second phase with $\mathcal{I} = \{i_1\}$.
- Set $\text{tar} = \infty$.

For motivation, at any step, girl $g_{\text{tar}}$ can be viewed as the target girl. If she accepts the next proposal then this will complete a rotation of cardinality two. Observe that we initialize $\text{tar} = \infty$ as it is impossible to complete a rotation in the first step.
To complete the description of the second-phase, assume we have \( \mathcal{I} = \{i_1, \ldots, i_k\} \). If \( k < \frac{n}{2} \) and less than \( n \log n \) proposals in total have been made then we generate the next arc of the rotation digraph starting at \( i_k \), as follows:

- Let boy \( b_{i_k} \) make uniformly random proposals until the first time he proposes to a girl \( g_j \) such that \( g_j \) ranks \( b_{i_k} \) higher than \( b_j \).
  - If \( j = \text{tar} \) then increment \( Z' \) by 1. Recurse.
  - If \( j \in \mathcal{I} \setminus \{\text{tar}\} \) then pick \( i_{k+1} \) from \( \{1, 2, \ldots, n\} \setminus \mathcal{I} \) uniformly at random. Set \( \mathcal{I} = \{i_1, \ldots, i_k, i_{k+1}\} \) and \( \text{tar} = \infty \). Recurse.
  - If \( j \notin \mathcal{I} \) then set \( i_k = j, \text{tar} = i_k, \mathcal{I} = \{i_1, \ldots, i_k, i_{k+1}\} \). Recurse.

- If, instead, boy \( b_{i_k} \) gets rejected by all the girls then return \( Z' = 0 \)

**Lemma 7.1.** The probability of the existence of a maximal rotation of size two is lower bounded by \( P(Z' \geq 1) \).

**Proof.** Observe that \( Z' \) is only incremented when we find a pair \( (i, j) \) such that the next girls to accept proposals from \( b_i \) and \( b_j \), respectively, are \( g_j \) and \( g_i \). This implies that \( Z' \geq 1 \) can only arise when there is a maximal rotation of cardinality 2. \( \square \)

Therefore, our aim is to prove that \( P(Z' \geq 1) = \Omega(1) \), where \( Z' \) is the random variable generated by the algorithm.

### 7.2 Bounding the Number of Proposals

Our objective now is to show that the behaviour of this new two-phase algorithm does not deviate too much from its expected behaviour. Specifically, we show it satisfies a series of properties with sufficiently high probability. As before, let \( T_1 \) and \( T_2 \) be the number of proposals made in the first and second phases, respectively, and let \( T = T_1 + T_2 \). Properties I to VII are as defined in Section 6. But now we require several more properties. To describe these, let \( p_{S_t} \) denote the probability of the next proposal being accepted when in state \( S_t \). We are interested in the following five properties that may apply to a state in the second phase:

- VIII. \( t \geq \frac{1}{2} n \log n \)
- IX. No more than \( n^{\frac{9}{10}} \) girls have received less than \( \frac{1}{4} \log n \) proposals.
- X. No more than \( \sqrt{n} \) girls have received a redundant proposal.
- XI. \( \{p_t | T_1 \leq \tau \leq t\} \subseteq \left[ \frac{1}{22 \log n}, \frac{5}{\log n} \right] \)
- XII. \( T_2 \geq \frac{1}{20} n \log n \)

Let \( G' \) be the set of all good states in the second phase satisfying these conditions. Like \( G \), \( G' \) is monotone. Let \( G^* \) denote the event \( S_{T-1} \in G' \), that is the event that the algorithm in a good state satisfying these conditions the period before it terminates.

**Lemma 7.2.** For \( n \) sufficiently large, \( P(G^*) \geq 1 - o(1) \).
Proof. Given the algorithm has not terminated, properties II to VII hold with high probability by the same argument as in Lemma 6.2. Therefore, it is enough to show that properties VIII to XII hold almost surely conditioned on II to VII.

Recall $T_1$ is the number of proposals until each girl receives at least one proposal. Thus $T_1$ is just the random variable for a coupon collector’s problem. Let $t^i$ be the number of proposals needed to collect the $i^{th}$ coupon after the first $i-1$ coupons have already been collected. So the $t^i$ are independent geometric random variables with parameters $\frac{n-(i-1)}{n}$ which sum to $T_1$. Thus the expectation of $T_1$ is $\sum_{i=1}^{n} E(t^i) = \sum_{i=1}^{n} \frac{n}{n-(i-1)} = n \log n + o(n \log n)$ with variance $\sum_{i=1}^{n} \text{Var}(t^i) = \sum_{i=1}^{n} \left( \left( \frac{n}{n-(i-1)} \right)^2 - \frac{n}{n-(i-1)} \right) = n^2 \sum_{j=1}^{n} \frac{1}{j^2} - n \sum_{j=1}^{n} \frac{1}{j}$. Consequently, the variance is bounded above by $\frac{\pi^2}{6} \cdot n^2$. Applying Chebyshev’s inequality then gives:

$$P\left( T_1 \leq \frac{1}{2} n \log n \right) \leq P\left( \left| T_1 - E(T_1) \right| \geq \frac{1}{3} n \log n \right) \leq \frac{\pi^2}{6 \left( \frac{1}{3} \log n \right)^2} = \frac{3\pi^2}{2 \log^2 n}$$

This proves VIII occurs almost surely as $t \geq T_1$ in the second phase.

Let $X_i$ be the indicator of whether girl $i$ has received at most $\frac{1}{4} \log n$ proposals after $t \geq \frac{1}{2} n \log n$ total proposals. Let $Y_i$ be the number of different proposals she has received in that amount of time. Then, by Markov’s inequality:

$$P\left( \sum_{i=1}^{n} X_i \geq n^{\frac{9}{10}} \right) \leq \frac{E(\sum_{i=1}^{n} X_i)}{n^{\frac{9}{10}}} = n^{\frac{13}{10}} \cdot P(X_1 = 1) = n^{\frac{13}{10}} \cdot P \left( Y_1 \leq \frac{1}{4} \log n \right)$$

Since $Y_i$ is binomial and has expectation greater than $\frac{1}{4} \log n$, applying a Chernoff bound gives:

$$P \left( Y_1 \leq \left( 1 - \frac{1}{2} \right) \frac{1}{2} \log n \right) \leq e^{-\frac{1}{8} \log n} = n^{-\frac{1}{8}}$$

Thus,

$$P \left( \sum_{i=1}^{n} X_i \geq n^{\frac{9}{10}} \right) \leq n^{\frac{13}{10} - \frac{1}{8}} = n^{\frac{11}{8}}$$

This proves that IX occurs almost surely.

Next let $R$ denote the number of redundant proposals made in the first phase. We know that, throughout the first phase, the set of girls that any boy has proposed to has cardinality at most $\log^4 n$. Thus, the probability that any proposal is redundant is bounded above by $\frac{\log^4 n}{n}$. This implies that the expected number of redundant proposals is at most:

$$\frac{\log^4 n}{n} \cdot 5n \log n = 5 \log^5 n$$

In particular, let $R_G$ be the number of girls who have received redundant proposals. Clearly $R_G \leq R$. Thus, by Markov’s inequality:

$$P \left( R_G \geq \sqrt{n} \right) \leq P \left( R \geq \sqrt{n} \right) \leq \frac{5 \log^5 n}{\sqrt{n}}$$

This proves that X occurs almost surely.
Now let $X$ be the set of girls who have received at least $\frac{1}{4} \log n$ distinct proposals. By VIII and IX, no more than $\frac{n}{2} \sqrt{n}$ girls have received less than $\frac{1}{4} \log n$ proposals by the end of the first phase. By X no more than $\sqrt{n}$ of the remaining girls have received redundant proposals. It follows that $|X| \geq n - \sqrt{n} - \frac{n}{10}$. Recall, $p_{S_t}$ is the probability the next proposal being accepted and let $\rho_i$ be the number of proposals received by $g_i$ at the end of the first phase. So, when VIII to X occur:

$$p_{S_t} \leq \frac{1}{n} \sum_{g \in G} \frac{1}{\rho_g + 1} \leq \frac{1}{n} \sum_{g \in X} \frac{1}{\rho_g + 1} \sqrt{n} + n^\frac{9}{10} \frac{1}{n} \leq \frac{n - n^\frac{9}{10} - \sqrt{n}}{\frac{1}{4} \log n + 1 + n^{-\frac{1}{2}} + n^\frac{1}{10}} \leq \frac{5}{\log n}$$

Similarly, by VII, at least $n - \log^2 n$ girls have not received a proposal from the proposing boy. This, along with IV, implies that when VIII to X occur:

$$p_{S_t} \geq \frac{n - \log^2 n}{n} \cdot \frac{1}{21 \log n + 1} \geq \frac{1}{22 \log n}$$

Hence, XI occurs almost surely.

Finally, to show the last property, we modify the proof of Lemma 6.6 using the bounds obtained from XI. Indeed, if $t \leq \frac{5 \log n}{2}$ then the algorithm has not terminated. For $t \geq \frac{5 \log n}{2}$, denoting $X_t$ to be the number of accepted proposals in the second phase, applying Lemma 5.4 with Lemma 5.3 for $t \leq \frac{1}{20} n \log n$ gives:

$$P(X_t \geq \frac{n}{2}) \leq P(X_t \geq \frac{10}{\log n}) \leq P(B_t \geq \frac{10}{\log n}) \leq \exp \left(-t \cdot \frac{5}{3} \log n \right) = o(1)$$

Thus XII occurs almost surely.

7.3 Bounding the Probability of Missing a Rotation

We can complete the proof of our second main result in two steps. First, we show that the probability of a maximal rotation of cardinality two existing is at least a constant. The second step is then easy. If there is a maximal rotation of cardinality two then a random coalition of cardinality at most $(1 - \epsilon) \cdot n$ will not be a winning coalition with constant probability.

**Theorem 7.3.** Let $Z'_1 = 1$ when the $t^{th}$ proposal of the second phases closes a rotation of size 2 and let $Z' = \sum_{t=1}^{T_2} Z'_t$. Then $P(Z' \geq 1) = \Omega(1)$.

**Proof.** The $t^{th}$ proposal of the second phase ends a rotation of size 2 if the following five events occur:

- $E^t_1$: “The second phase has not ended and we are not in the first run of the second phase.”
- $E^t_2$: “The last proposal of the previous run was to a girl who hadn’t yet accepted a proposal.”
- $E^t_3$: “The proposal is made to the optimal partner of the boy who started the previous run.”
- $E^t_4$: “The proposal is not redundant.”
- $E^t_5$: “The proposal is accepted.”

If $111 \log^2 n \leq t \leq \frac{1}{20} n \log n$, by V and by XII we have $P(E^t_1 | S_t \in G') = 1$.

Denote $E^t_{(a,b)} = \bigwedge_{i=a}^b E^t_i$ and let $G_t$ be the set of girls who have accepted proposals in the second phase. By XI the probability of a proposal being accepted is at most $\frac{5}{\log n}$. On the other hand, by IV
the probability of a proposal being accepted by a girl who isn’t in $G_1$ is at least $\frac{1}{n} \cdot \frac{n^2}{22 \log n} = \frac{1}{44 \log n}$, since at least $\frac{n}{2}$ girls are not in $G_1$. Thus, the probability of the proposal being accepted by a girl who isn’t in $G_1$, given that it was accepted, is at least:

$$\frac{1}{44 \log n} = \frac{1}{220}$$

Therefore, $P(E'_3|E'_1 \land S_t \in \mathcal{G}') \geq \frac{1}{220}$. Now, clearly $P(E'_3|E'_{(1,2)} \land S_t \in \mathcal{G}') = \frac{1}{n}$. Moreover, by VII, $P(E'_4|E'_{(1,3)} \land S_t \in \mathcal{G}') \geq 1 - \frac{\log^4 n}{n}$ and by IV:

$$P(E'_5|E'_t \land S_t \in \mathcal{G}') \geq \frac{1}{22 \log n}$$

Recall by VI, the first run contains at most $111 \log n$ proposals. We may conclude that for large enough $n$ and $t$ such that $111 \log n \leq t \leq \frac{1}{20} n \log n$:

$$P(Z'_t = 1|S_t \in \mathcal{G}') = P(E'_{(1,5)}|S_t \in \mathcal{G}') \geq \frac{1}{220} \cdot \frac{1}{n} \cdot \left(1 - \frac{\log^4 n}{n}\right) \cdot \frac{1}{22 \log n} \geq \frac{1}{5000 n \log n}$$

Next denote the event to be $E'$. Since $\sum_{\tau=1}^{t} Z'_\tau$ is a counter, by Lemma 5.3, we obtain the following bound:

$$P(Z' = 0) \leq \sum_{S_{111 \log^2 n}} P(E') \cdot P\left(\sum_{S_{\frac{1}{20} n \log n}} Z'_\tau = 0 \land S_{\frac{1}{20} n \log n} \in \mathcal{G}' \bigg| S_{111 \log^2 n} = S_{111 \log^2 n}\right)$$

$$\leq \sum_{S_{111 \log^2 n}} P\left(B_{\frac{1}{20} n \log n - 111 \log^2(n), \frac{1}{5000 n \log n} = 0}\right)$$

$$= P\left(S_{\frac{1}{20} n \log n} \in \mathcal{G}'\right) \cdot P\left(B_{\frac{1}{20} n \log n - 111 \log^2(n), \frac{1}{5000 n \log n} = 0}\right)$$

By Lemma 7.2, $P\left(S_{\frac{1}{20} n \log n} \in \mathcal{G}'\right) = 1 - o(1)$. So it is enough to show that:

$$P\left(B_{\frac{1}{20} n \log n - 111 \log^2(n), \frac{1}{5000 n \log n} = 0}\right) < 1 - \Omega(1)$$

This follows as

$$P\left(B_{\frac{1}{20} n \log n - 111 \log^2(n), \frac{1}{5000 n \log n} = 0}\right) = \left(1 - \frac{1}{5000 n \log n}\right) \cdot \frac{1}{20 n \log n - 111 \log^2(n)}$$

$$\leq \left(1 - \frac{1}{5000 n \log n}\right) \cdot \frac{1}{20 n \log n}$$

$$= e^{-\frac{1}{10000}}$$

$$= 1 - \Omega(1)$$

We may now complete the proof of our second main result.
Proof of Theorem 5.2. The probability that in a random instance there is a rotation of size 2 is $\Omega(1)$ by Theorem 7.3. Given that there is a rotation of size 2, the following is the probability that a random set of $\lambda n$ girls misses this rotation of size 2:

\[
\frac{\binom{n-2}{\lambda n}}{\binom{n}{\lambda n}} = \frac{(n-2)!}{(n-2-\lambda n)! \cdot (\lambda n)!} \cdot \frac{(n-\lambda n)! \cdot (\lambda n)!}{n!} \\
= \frac{n-\lambda n}{n} \cdot \frac{n-1-\lambda n}{n-1} \\
\geq \left(1 - \frac{\lambda n}{n-1}\right)^2.
\]

So, if $\lambda < 1 - \varepsilon$ for any positive constant $\varepsilon$, then the probability of missing a rotation is $\Omega(1) \cdot \left(1 - \frac{\lambda n}{n-1}\right)^2 = \Omega(1)$. \hfill \Box

8 Conclusion

We have evaluated the expected cardinality of the minimum winning coalition. We believe this result is of theoretical interest and that the techniques applied may have broader applications for stable matching problems. In terms of practical value it is worth discussing the assumptions inherent in the model. The assumption of uniform and independent random preferences, while ubiquitous in the theoretical literature, is somewhat unrealistic in real-world stable matching instances. Furthermore, as presented, the model assumes full information, which is clearly not realistic in practice. However, to implement the behavioural strategy presented in this paper, the assumption of full information is not required. It needs only that a girl has a good approximation of the rank of her best stable partner. But, by the results of Pittel [16], she does know this with high probability. Consequently, a near-optimal implementation of her behavioural strategy requires knowledge only of her own preference list! This allows for a risk-free method to output a matching close in the lattice to the girl-optimal stable matching. Similarly, as discussed, although our presentation has been in terms of a coalition of girls, each girl is able to implement a near-optimal behavioural strategy independent of who the other girls in the coalition may be or what their preferences are.

References

[1] I. Ashlagi, Y. Kanoria, and J. Leshno. Unbalanced random matching markets: the stark effect of competition. Journal of Political Economy, 125(1):69–98, 2017.

[2] G. Birkhoff. Rings of sets. Duke Mathematical Journal, 3(3):443–454, 1937.

[3] L. Dubins and D. Freedman. Machiavelli and the Gale-Shapley algorithm. American Mathematical Monthly, 88(7):485–494, 1981.

[4] Philippe Flajolet and Andrew M. Odlyzko. Random mapping statistics. In Jean-Jacques Quisquater and Joos Vandewalle, editors, Advances in Cryptology — EUROCRYPT ’89, pages 329–354, Berlin, Heidelberg, 1990. Springer Berlin Heidelberg.

[5] D. Gale and L. Shapley. College admissions and the stability of marriage. American Mathematical Monthly, 69(1):9–15, 1962.
[6] D. Gale and M. Sotomayor. Ms. Machiavelli and the stable matching problem. *The American Mathematical Monthly*, 92(4):261–268, 1985.

[7] Y. Gonczarowski. Manipulation of stable matchings using minimal blacklists. In *Proceedings of the Fifteenth ACM Conference on Economics and Computation*, EC ’14, page 449. Association for Computing Machinery, 2014.

[8] D. Gusfield. Three fast algorithms for four problems in stable marriage. *SIAM Journal on Computing*, 16(1):111–128, 1987.

[9] D. Gusfield and R. Irving. *The Stable Marriage Problem: Structure and Algorithms*. MIT Press, 1989.

[10] R. Irving and P. Leather. The complexity of counting stable marriages. *SIAM Journal on Computing*, 15(3):655–667, 1986.

[11] R. Irving, P. Leather, and D. Gusfield. An efficient algorithm for the "optimal" stable marriage. *J. ACM*, 34(3):532–543, 1987.

[12] D. Knuth. Mariages stables et leurs relations avec d’autres problèmes combinatoires. *Les Presses de l’Université de Montréal*, 1982.

[13] D. Knuth, R. Motwani, and B. Pittel. Stable husbands. *Random Structures and Algorithms*, 1(1):1–14, 1990.

[14] R. Kupfer. The instability of stable matchings: The influence of one strategic agent on the matching market. *Proceedings of the 16th Conference on Web and Internet Economics (WINE)*, 2020.

[15] D. McVitie and L. Wilson. The stable marriage problem. *Communications of the ACM*, 14:486–490, 1971.

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

[17] B. Pittel. On likely solutions of a stable matching problem. *The Annals of Applied Probability*, 2(2):358–501, 1992.

[18] A. Roth. The economics of matching: stability and incentives. *Mathematics of Operations Research*, 7(4):617–628, 1982.

[19] R. Stanley. *Enumerative Combinatorics, Volume I*. Cambridge University Press, 1997.

[20] L. Wilson. An analysis of the stable marriage assignment algorithm. *BIT Numerical Mathematics*, 12:569–575, 1972.