

## ON RANDOM STABLE PARTITIONS

BORIS PITTEL

ABSTRACT. The stable roommates problem does not necessarily have a solution, i.e. a stable matching. We had found that, for the uniformly random instance, the expected number of solutions converges to  $e^{1/2}$  as  $n$ , the number of members, grows, and with Rob Irving we proved that the limiting probability of solvability is  $e^{1/2}/2$ , at most. Stephan Mertens's extensive numerics compelled him to conjecture that this probability is of order  $n^{-1/4}$ . Jimmy Tan introduced a notion of a stable cyclic partition, and proved existence of such a partition for every system of members' preferences, discovering that presence of odd cycles in a stable partition is equivalent to absence of a stable matching. In this paper we show that the expected number of stable partitions with odd cycles grows as  $n^{1/4}$ . However the standard deviation of that number is of order  $n^{3/8} \gg n^{1/4}$ , too large to conclude that the odd cycles exist with high probability (whp). Still, as a byproduct, we show that whp the fraction of members with more than one stable "predecessor" is of order  $n^{-1/4}$ . Furthermore, whp the average rank of a predecessor in every stable partition is of order  $n^{1/2}$ . The likely size of the largest stable matching is  $n/2 - O(n^{1/4+o(1)})$ , and the likely number of pairs of unmatched members blocking the optimal complete matching is  $O(n^{3/4+o(1)})$ .

## 1. INTRODUCTION AND MAIN RESULTS

A roommates problem instance is specified by an even integer  $n$ , number of members, and for each  $i$  ( $1 \leq i \leq n$ ) a permutation  $\sigma_i$  of the set  $[n] = \{1, 2, \dots, n\}$  in which  $i$  itself occupies position  $n$ , ( $\sigma_i(n) = i$ ). The permutation  $\sigma_i$  forms the preference list of person  $i$ :  $\sigma_i(k) = j$  if person  $j$  occupies position  $k$  in the preference list of person  $i$ , and each person  $i$  is at the end of their own preference list. Equivalently, the instance can be specified by the ranking list  $R_i$  of each person  $i$ , defined as the inverse permutation of  $\sigma_i$ :  $R_i(j) = k$  if the person  $j$  is the  $k$ -th best for person  $i$ .

---

*Date:* July 2, 2018.

*2010 Mathematics Subject Classification.* 05C30, 05C80, 05C05, 34E05, 60C05.

*Key words and phrases.* stable matchings, partitions, random preferences, asymptotics

For a given roommates instance with  $n$  members, a stable permutation (cyclic partition) is a permutation  $\mathbf{\Pi}$  of  $[n]$  such that:

$$(1.1) \quad \begin{aligned} (1) & \forall i \in [n] : R_i(\mathbf{\Pi}(i)) \leq R_i(\mathbf{\Pi}^{-1}(i)); \\ (2) & \forall 1 \leq i \neq j \leq n : R_i(j) < R_i(\mathbf{\Pi}^{-1}(i)) \implies R_j(i) > R_j(\mathbf{\Pi}^{-1}(j)). \end{aligned}$$

Viewing  $\mathbf{\Pi}$  in terms of its cyclic decomposition, we will refer to  $\mathbf{\Pi}(i)$  and  $\mathbf{\Pi}^{-1}(i)$  as the successor of  $i$  and the predecessor of  $i$  in the permutation  $\mathbf{\Pi}$ . Then condition (1) states no person prefers his predecessor to his successor, and condition (2) states that no two mutually-unaligned members prefer each other to their predecessors. Note that equality in condition (1) is possible iff  $\mathbf{\Pi}^2(i) = i$ , i.e. either  $i$  is a fixed point of  $\mathbf{\Pi}$ , or  $(i, \mathbf{\Pi}(i))$  is a transposition in  $\mathbf{\Pi}$ —in this case we say that  $(i, \mathbf{\Pi}(i))$  forms a pair in the partition  $\mathbf{\Pi}$ . Thus inequality (1) is not vacuous iff  $i$  is in a cycle of length 3 or more, in which case it is strict. Also if  $i$  is a fixed point, then  $R_i(\mathbf{\Pi}^{-1}(i)) = R_i(i) = n$ ; so condition (2) implies that there are no other fixed points, and every  $j \neq i$  prefers his own predecessor to  $i$ . Intuitively, each member  $i$  proposes to  $\mathbf{\Pi}(i)$  and holds a proposal from  $\mathbf{\Pi}^{-1}(i)$ .

Clearly, if a stable partition  $\mathbf{\Pi}$  is such that it has cycles of length 2 only, then  $\mathbf{\Pi}$  is a stable matching. However, while for every even  $n \geq 4$  there are instances without a stable matching, Tan [15], who introduced the notion of a cyclic partition  $\mathbf{\Pi}$ , proved that, for every instance of preferences, (1) there is at least one stable permutation; (2) all stable permutations have the same odd cycles (“parties”); (3) replacing each even cycle  $(i_1, i_2, \dots, i_{2m})$  of a stable permutation by the transpositions  $(i_1, i_2), \dots, (i_{2m-1}, i_{2m})$ , or by the transpositions  $(i_2, i_3), \dots, (i_{2m}, i_1)$  we get another stable, *reduced*, permutation; (4) thus a stable matching exists iff there are no odd cycles.

Suppose that the random problem instance, call it  $I_n$ , is chosen uniformly at random among all  $[(n-1)!]^n$  instances. We showed [14] that the expected number of stable matchings is  $e^{1/2}$  in the limit, implying that the number of stable matchings, if any exist, is bounded in probability. With Robert Irving [7] we proved that the probability that a stable matching exists is at most  $e^{1/2}/2 < 1$  in the limit. In a pleasing contrast, the stable partitions do not have a fixed point (odd party of size 1) with surprisingly high probability  $\geq 1 - O(e^{-\sqrt{n}})$ . So while a stable matching may not exist, stable partitions (that exist always) with high probability have no “pariahs”: every member holds a proposal from another member, while his own proposal is accepted by possibly a different member.

Our task is to analyze asymptotic behavior of a series of leading parameters of the family of stable (reduced) partitions for  $I_n$ , and we focus on those that have no fixed point. Among those parameters are  $\mathcal{S}_n$  and  $\mathcal{O}_n$ , the total number of stable (reduced) partitions and the total number of “parties”, i.

e. odd (common to all those partitions) cycles. We will prove, for instance, that

$$(1.2) \quad \mathbb{E}[\mathcal{S}_n] = (1 + o(1)) \frac{\Gamma(1/4)}{\sqrt{\pi e} 2^{1/4}} n^{1/4},$$

$$(1.3) \quad \mathbb{E}[\mathcal{O}_n] \leq (1 + o(1)) \frac{\Gamma(1/4)}{4\sqrt{\pi e} 2^{1/4}} n^{1/4} \log n.$$

The fact that  $\mathbb{E}[\mathcal{S}_n] \rightarrow \infty$ , but at a moderate rate, can be charitably viewed as supporting the claim that, in probability,  $\mathcal{S}_n \rightarrow \infty$ ; if this is the case then with high probability  $I_n$  has no stable matching. Numerical experiments conducted by Stephan Mertens [11] made him conjecture that solvability probability goes to zero, as fast as  $n^{-1/4}$ . For a rigorous transition from  $\mathbb{E}[\mathcal{S}_n] \rightarrow \infty$  to  $\mathcal{S}_n \rightarrow \infty$ , one would normally want to show that  $\text{Var}(\mathcal{S}_n) \ll \mathbb{E}^2[\mathcal{S}_n]$ . It turns out, however, that  $\text{Var}(\mathcal{S}_n)$  is of order  $n^{3/4}$ , thus exceeding  $\mathbb{E}^2[\mathcal{S}_n]$  by the factor  $n^{1/8}$ , which invalidates this naive two-moment approach. Can the approach be gainfully modified by narrowing the pool of stable partitions?

A key tool for estimating  $\text{Var}(\mathcal{S}_n)$  is an asymptotic formula for the probability that each of two generic (reduced) partitions  $\mathbf{\Pi}_1$  and  $\mathbf{\Pi}_2$  (with the same odd parties, of course) are stable. The symmetric difference of the set of matched pairs in  $\mathbf{\Pi}_1$  and the set of matched pairs in  $\mathbf{\Pi}_2$  is the edge set of the disjoint even cycles of length  $\geq 4$ , whose edges are the matched pairs in  $\mathbf{\Pi}_1$  interlacing the matched pairs in  $\mathbf{\Pi}_2$ . Each such cycle can be viewed as an even *rotation* in both partitions, so that the pair  $(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  gives rise to  $2^\mu$  of stable partitions, with  $\mu$  being the total number of those even cycles. Define a random graph  $G_n = (\mathcal{V}_n, \mathcal{E}_n)$ , where  $\mathcal{V}_n$  is the set of all stable partitions  $\mathbf{\Pi}$ , and  $\mathcal{E}_n$  is the set of pairs  $(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$ , each giving rise to a *single* even cycle. By (5.23),  $\mathbb{E}[\mathcal{V}_n] = \mathbb{E}[\mathcal{S}_n]$  is of order  $n^{1/4}$ . It turns out that  $\mathbb{E}[\mathcal{E}_n]$  is of order  $n^{1/4}$  as well. What, if anything, does this fact tell about the likely range of  $\mathcal{S}_n$ ?

There are two positive results that stem from (5.23)–(5.24). Tan [15], [16] defined a maximum stable matching for an instance  $I$  as a maximum-size matching  $M = M(I)$  which is *internally* stable, i.e. not blocked by any two members from the agent (vertex) set of  $M$ . He proved that  $|M(I)| = (n - \mathcal{O}(I))/2$ . It follows from (5.24) that

$$\mathbb{P}\left(|M(I_n)| \geq \frac{n - \omega(n)n^{1/4} \log n}{2}\right) \geq 1 - O(\omega(n)^{-1}) \rightarrow 1,$$

for  $\omega(n) \rightarrow \infty$ , however slowly. In short, the number of members not in the maximum stable matching is  $O_p(n^{1/4} \log n)$ .

Abraham, Biró and Manlove [1] introduced the alternative notion of a “maximally stable” matching, i.e. a matching  $M$  on  $[n]$  that is blocked by the smallest number of pairs, call it  $B(I)$ , of agents unmatched in  $M$ . They obtained a two-sided bound for  $B(I)$  in terms of preference lists lengths and the odd cycles. A cruder version of the ABM upper bound states that  $B(I) \leq d(I)\mathcal{O}(I)$ , where  $d(I)$  is the length of the longest preference list. Extending our approach, we will show that for the random instance  $I_n$ , with probability  $\geq 1 - \exp(-c(\log n)^{2(1+\delta)})$ , every member’s predecessor is among their best  $n^{1/2}(\log n)^{1+\delta}$  choices. So we can apply the last bound with  $d(I_n) = n^{1/2}(\log n)^{1+\delta}$ . Therefore the bound (5.24) together with the ABM bound imply that with high probability there exists a complete matching which is blocked by  $n^{3/4+o(1)}$  pairs, a strikingly small number relative to the total number  $(\Theta(n^2))$  of potential blocking pairs.

We will also show that with high probability the sum of the ranks of predecessors in every stable partition is asymptotic to  $n^{3/2}$ ; consequently the worst predecessor’s rank in every stable partition is  $n^{1/2}(1 - o(1))$  at least, nearly matching  $n^{1/2}(\log n)^{1+\delta}$ , the likely upper bound.

Here is an application. Suppose we shrink every member’s preference list to their own best  $d$  choices. If the constrained instance has no fixed point then neither does the full-lists instance. Consider an instance  $I_{n,d}$  of the stable partition problem chosen uniformly at random among all instances with some  $d$  acceptable choices for every member. Randomly, and independently, ordering the remaining  $n - 1 - d$  members for every member, we will get the uniformly random (full-lists) instance  $I_n$ . It follows then that if  $d \leq (1 - \varepsilon)n^{1/2}$  ( $d \geq n^{1/2}(\log n)^{1+\delta}$  resp.) then with high probability stable partitions for  $I_{n,d}$  have (do not have resp.) a fixed point.

Finally, we use the analysis of  $\text{Var}(\mathcal{S}_n)$  to show that the expected fraction of members with multiple stable predecessors is of order  $n^{-1/4}$ .

## 2. INTEGRAL FORMULAS FOR STABILITY PROBABILITIES

At the core of our proofs are two integral formulas, one for the probability that a generic cyclic partition is stable, another for the probability that two generic cyclic partitions are stable.

**Lemma 2.1.** *Let  $\Pi$  be a permutation of  $[n]$  with even cycles of length 2 only, and possibly a single fixed point  $h^*$ , i. e.  $\Pi(h^*) = h^*$ . Let  $\text{Odd}(\Pi)$  be the set of all elements from the odd cycles of  $\Pi$  with an exception of the fixed point if it is present. Let  $D(\Pi)$  be the set of unordered pairs  $(i \neq j)$*

such that  $i = \mathbf{\Pi}(j)$  or, not exclusively,  $j = \mathbf{\Pi}(i)$ . Then

$$(2.1) \quad \begin{aligned} \mathbb{P}(\mathbf{\Pi}) &:= \mathbb{P}(\mathbf{\Pi} \text{ is stable}) = \int \cdots \int_{\mathbf{x} \in [0,1]^{n-1}} F(\mathbf{x}) \, d\mathbf{x}, \\ F(\mathbf{x}) &:= \prod_{h \in \text{Odd}(\mathbf{\Pi})} x_h \cdot \prod_{(i,j) \notin D(\mathbf{\Pi})} (1 - x_i x_j) \cdot \prod_{k \neq h^*} (1 - x_k); \end{aligned}$$

if there is no fixed point  $h^*$ , then the third product is replaced by 1, and  $[0, 1]^{n-1}$  by  $[0, 1]^n$ .

If  $\mathbf{\Pi}$  is a matching, we get ([14])

$$(2.2) \quad \mathbb{P}(\mathbf{\Pi} \text{ is stable}) = \int \cdots \int_{\mathbf{x} \in [0,1]^n} \prod_{(i,j) \notin D(\mathbf{\Pi})} (1 - x_i x_j) \, d\mathbf{x}.$$

*Proof.* To generate the random instance  $I_n$ , introduce an array of the independent random variables  $X_{i,j}$  ( $1 \leq i \neq j \leq n$ ), each distributed uniformly on  $[0, 1]$ . Assume that each member  $i \in [n]$  ranks the members  $j \neq i$  in increasing order of the variables  $X_{i,j}$ . Such an ordering is uniform for every  $i$ , and the orderings by different members are independent. Then

$$(2.3) \quad \begin{aligned} \mathbb{P}(\mathbf{\Pi} \text{ is stable} \mid X_{i,\mathbf{\Pi}^{-1}(i)} = x_i, i \in [n]) \\ = \prod_{h \in \text{Odd}(\mathbf{\Pi})} x_h \cdot \prod_{(i,j) \notin D(\mathbf{\Pi})} (1 - x_i x_j) \prod_{k \neq h^*} (1 - x_k) \end{aligned}$$

Indeed, by (1.1),  $\mathbf{\Pi}$  is stable iff

- (1) for every  $h \in \text{Odd}(\mathbf{\Pi})$ :  $X_{h,\mathbf{\Pi}(h)} < X_{h,\mathbf{\Pi}^{-1}(h)}$ ,
- (2) for every  $(i,j) \notin D(\mathbf{\Pi})$ ,  $i, j \neq h^*$ :  $X_{i,j} < X_{i,\mathbf{\Pi}^{-1}(i)} \Rightarrow X_{j,i} > X_{j,\mathbf{\Pi}^{-1}(j)}$ ,
- (3) for every  $i \neq h$ :  $X_{i,\mathbf{\Pi}^{-1}(i)} < X_{i,h^*}$ .

And, conditioned on the event  $\{X_{i,\mathbf{\Pi}^{-1}(i)} = x_i, i \in [n]\}$ , the events above are independent, with (conditional) probabilities  $x_h$ ,  $1 - x_i x_j$  and  $1 - x_k$  respectively. Using Fubini's theorem, we have (2.1).  $\square$

Like analogous formulas in [13], [14] and [7], this is a non-bipartite counterpart of Knuth's formula for stable bipartite matchings, [8]. His derivation was based on the inclusion-exclusion method, coupled with ingenious observation that the resulting sum equals the multidimensional integral of a product-type integrand resembling our  $F(\mathbf{x})$ . Of course, we could get a sum-type formula for  $\mathbb{P}(\mathbf{\Pi} \text{ is stable})$  by expanding the product in (2.3) and integrating the resulting sum term-wise. Moving in the opposite direction, i.e. starting with an inclusion-exclusion formula for  $\mathbb{P}(\mathbf{\Pi} \text{ is stable})$ , finding an integral-type representation of the generic summand, and discerning that the sum of the attendant integrands happens to be an expansion of the

“out-of-the blue” product in (2.3), would have been very problematic. The identity (2.1) is indispensable for asymptotic estimates, thanks to a simple, but powerful, bound

$$(2.4) \quad \prod_{\{i,j\} \notin D(\mathbf{\Pi})} (1 - x_i x_j) \leq \exp\left(-\frac{s^2}{2} + 4.5\right), \quad s := \sum_{i \in [n]} x_i.$$

For instance, this bound and  $\prod_k (1 - x_k) \leq e^{-s}$  will almost immediately yield that the stable partitions have no fixed point with probability  $\geq 1 - e^{-\Theta(n^{1/2})}$ . We will prove a surprisingly simple, yet qualitatively sharp estimate: uniformly for a fixed-point free partitions  $\mathbf{\Pi}$ ,

$$(2.5) \quad P(\mathbf{\Pi} \text{ is stable}) = O\left(\frac{1}{(n + m - 1)!!}\right), \quad m := |\text{Odd}(\mathbf{\Pi})|.$$

We note that Alcalde [2] defined an *exchange stable* matching as a matching  $M$  that, to quote from [10], “admits no *exchange-blocking pair*, which is a pair of members each of whom prefers the other’s partner in  $M$  to their own”. Cechlárová and Manlove [4] proved that, in stark contrast with the classic stable roommates model, the problem of determining whether a given instance admits an exchange-stable matching is NP-complete. The interested reader may wish to check that the formula (2.2) continues to hold for  $P(\mathbf{\Pi} \text{ is exchange-stable})$ . Consequently the expected number of exchange-stable matchings and the expected number of the classic stable matchings are exactly the same, implying that the former is also asymptotic to  $e^{1/2}$ . Let us call a (fixed-point free) partition  $\mathbf{\Pi}$  exchange stable if no two members prefer each other predecessors to their own predecessors under  $\mathbf{\Pi}$ . What about the partitions that are “doubly-stable”, i.e. stable *and* exchange stable? It turns out that

$$P(\mathbf{\Pi}) := P(\mathbf{\Pi} \text{ is doubly stable}) = \int \cdots \int_{\mathbf{x} \in [0,1]^n} F_2(\mathbf{x}) \, d\mathbf{x},$$

$$F_2(\mathbf{x}) := \prod_{h \in \text{Odd}(\mathbf{\Pi})} x_h \cdot \prod_{(i,j) \notin D(\mathbf{\Pi})} (1 - x_i x_j)^2.$$

The counterpart of (2.5) is  $P(\mathbf{\Pi}) = O(2^{-\frac{n+m}{2}}/(n+m-1)!!)$ , implying that the expected number of the doubly stable partitions is of order  $2^{-n/2}$ , way down from  $n^{1/4}$  for the stable partitions.

Continuing, introduce  $\mathcal{R}(\mathbf{\Pi})$ , the sum of the ranks of all predecessors in the preference lists of their successors in a partition  $\mathbf{\Pi}$ . Let  $P_k(\mathbf{\Pi}) := P(\mathbf{\Pi} \text{ is stable and } \mathcal{R}(\mathbf{\Pi}) = k)$ .

**Lemma 2.2.** *Suppose  $\Pi$  is fixed-point free. Then, letting  $m := |\text{Odd}(\Pi)|$ , and  $\bar{x} = 1 - x$ ,*

$$(2.6) \quad \begin{aligned} P_k(\Pi) &= \int \cdots \int_{\mathbf{x} \in [0,1]^n} [z^{k-n-m}] F(\mathbf{x}, z) d\mathbf{x}, \\ F(\mathbf{x}, z) &:= \prod_{h \in \text{Odd}(\Pi)} x_h \cdot \prod_{(i,j) \notin D(\Pi)} (\bar{x}_i \bar{x}_j + zx_i \bar{x}_j + z\bar{x}_i x_j). \end{aligned}$$

*Proof.* First of all, using  $\chi(A)$  to denote the indicator of an even  $A$ , we have

$$P_k(\Pi) = [z^k] \mathbb{E} \left[ z^{\mathcal{R}(\Pi)} \chi(\Pi \text{ is stable}) \right].$$

Here

$$\begin{aligned} \chi(\Pi \text{ is stable}) &= \prod_{(i,j) \notin D(\Pi)} \chi(X_{i,j} > X_{i,\Pi^{-1}(i)} \text{ or } X_{j,i} > X_{j,\Pi^{-1}(j)}) \\ &\quad \times \prod_{h \in \text{Odd}(\Pi)} \chi(X_{h,\Pi(h)} < X_{h,\Pi^{-1}(h)}). \end{aligned}$$

Furthermore

$$\begin{aligned} \mathcal{R}(\Pi) &= \sum_{(i,j) \notin D(\Pi)} [\chi(X_{i,j} < X_{i,\Pi^{-1}(i)}) + \chi(X_{j,i} < X_{j,\Pi^{-1}(j)})] \\ &\quad + \sum_{i \in [n]} 1 + \sum_{h \in \text{Odd}(\Pi)} \chi(X_{h,\Pi(h)} < X_{h,\Pi^{-1}(h)}), \end{aligned}$$

where the second sum accounts for the pairs  $(i, \Pi^{-1}(i))$ ,  $i \in [n]$ . So

$$\begin{aligned} &\mathbb{E} \left[ z^{\mathcal{R}(\Pi)} \chi(\Pi \text{ is stable}) \mid X_{i,\Pi^{-1}(i)} = x_i, i \in [n] \right] \\ &= z^{n+|\text{Odd}(\Pi)|} \prod_{h \in \text{Odd}(\Pi)} x_h \\ &\quad \times \prod_{(i,j) \notin D(\Pi)} \mathbb{E} \left[ z^{\chi(X_{i,j} < x_i) + \chi(X_{j,i} < x_j)} \chi(X_{i,j} > x_i \text{ or } X_{j,i} > x_j) \right] \\ &= z^{n+m} \prod_{h \in \text{Odd}(\Pi)} x_h \cdot \prod_{(i,j) \notin D(\Pi)} (\bar{x}_i \bar{x}_j + zx_i \bar{x}_j + z\bar{x}_i x_j). \end{aligned}$$

So

$$\mathbb{E} \left[ z^{\mathcal{R}(\Pi)} \chi(\Pi \text{ is stable}) \right] = z^{n+m} \int \cdots \int_{\mathbf{x} \in [0,1]^n} F(\mathbf{x}, z) d\mathbf{x},$$

which proves (2.6).  $\square$

Finally, suppose we have a pair of distinct cyclic partitions,  $\Pi_1$  and  $\Pi_2$ . Let  $P(\Pi_1, \Pi_2)$  denote the probability that both  $\Pi_1$  and  $\Pi_2$  are stable. We assume the two partitions have the same odd cycles, since otherwise the probability is zero. Suppose also there is no fixed point. Let

$\text{Odd}_{1,2}$  stand for the vertex set of the family of odd cycles, common to both partitions; so  $\Pi_1(h) = \Pi_2(h)$  for all  $h \in \text{Odd}_{1,2}$ . The cardinality  $|\text{Odd}_{1,2}|$  is even, and  $\Pi_1$  and  $\Pi_2$  induce a pair of perfect matchings  $(M_1, M_2)$  on  $\text{Even}_{1,2} := [n] \setminus \text{Odd}_{1,2}$ . Together,  $M_1$  and  $M_2$  determine a graph  $G(M_1, M_2) = (\text{Even}_{1,2}, E)$ , with the edge set  $E$  formed by the pairs  $(i, j) \in M_1 \cup M_2$ . Each component of  $G(M_1, M_2)$  is either an edge  $e \in M_1 \cap M_2$ , or a circuit of even length at least 4, in which the edges from  $M_1$  and  $M_2$  alternate. The edge set for all these (alternating) circuits is the symmetric difference  $M_1 \Delta M_2$ .

**Lemma 2.3.** *Let  $P(\Pi_1, \Pi_2)$  denote the probability that both  $\Pi_1$  and  $\Pi_2$  are stable. For  $r = 1, 2$ , let  $D_r$  be the set of unordered pairs  $(i \neq j)$  such  $i = \Pi_r(j)$  or, not exclusively,  $j = \Pi_r(i)$ . Then*

$$P(\Pi_1, \Pi_2) = \int \cdots \int_{\mathbf{x}, \mathbf{y} \in [0, 1]^n} F(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} d\mathbf{y},$$

$$F(\mathbf{x}, \mathbf{y}) = \prod_{h \in \text{Odd}_{1,2}} x_h \cdot \prod_{(i,j) \in D_1^c \cup D_2^c} [1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j)];$$

here

$$d\mathbf{x} = \prod_{i \in [n]} dx_i, \quad d\mathbf{y} = \prod_{i: \Pi_1(i) \neq \Pi_2(i)} dy_i,$$

and for every circuit  $\{i_1, \dots, i_\ell\}$  of  $G(M_1, M_2)$ :

$$\begin{aligned} &\text{either } x_{i_1} > y_{i_1}, x_{i_2} < y_{i_2}, \dots, x_{i_\ell} < y_{i_\ell}, \\ &\text{or } x_{i_1} < y_{i_1}, x_{i_2} > y_{i_2}, \dots, x_{i_\ell} > y_{i_\ell}. \end{aligned}$$

We omit the proof since it combines the elements of the proof for  $P(\Pi$  is stable) and of the formula for  $P(\Pi_1, \Pi_2)$  in the case when  $\Pi_1$  and  $\Pi_2$  are matchings, given in [14]. A counterpart of the bound (2.4) is

$$(2.7) \quad \prod_{(i,j) \in D_1^c \cup D_2^c} [1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j)] \leq e^{2^s} \exp\left(-\frac{s_1^2}{2} - \frac{s_2^2}{2} + \frac{s_{1,2}^2}{2}\right);$$

here  $s_1 = \sum_i x_i$ ,  $s_2 = \sum_i y_i$ ,  $s_{1,2} = \sum_i (x_i \wedge y_i)$  and  $i$  runs over  $[n]$ . Never mind enormity of  $e^{2^s}$ ; like (2.4), the bound (2.7) is both simple and instrumental in identifying a relatively small, eminently tractable, part of the integration domain which is “in charge” of the asymptotic behavior of  $P(\Pi_1, \Pi_2)$ .

**Note.** The reader interested in our prior work on stable roommates problem ([13], [14] and [7]) will not find the inequalities (2.4) and (2.7) there. Working on this project, we detected a technical, estimational, glitch (see

the next Section for details) in [13], equally consequential for analysis in [14] and [7]. Luckily the new bounds (2.4)-(2.7) allow to repair this error and, as an unexpected bonus, to simplify the arguments as well. The analysis in this paper can be viewed, in part, as a close template for the correction of that embarrassing oversight. We emphasize that, fortunately, this correction leaves the ultimate asymptotic results in those references intact.

*Proof.* As in the proof of Lemma 2.1, we use the array  $\{X_{i,j} : 1 \leq i \neq j \leq n\}$ . By the definition of stability, we have

$$\{\mathbf{\Pi}_1, \mathbf{\Pi}_2 \text{ are both stable}\} = \bigcap_{h \in \text{Odd}(\mathbf{\Pi}_{1,2})} A_h \bigcap_{(i,j) \in D_1^c \cup D_2^c} (B_{(i,j)})^c.$$

Here  $A_h = \{X_{h, \mathbf{\Pi}_{1,2}(h)} < X_{h, \mathbf{\Pi}_{1,2}^{-1}(h)}\}$ . Furthermore: **(1)** if  $(i, j) \in D_1^c \cap D_2^c$ , then

$$B_{(i,j)} = \left\{ X_{i,j} < X_{i, \mathbf{\Pi}_1^{-1}(i)}; X_{j,i} < X_{j, \mathbf{\Pi}_1^{-1}(j)} \right\} \\ \cup \left\{ X_{i,j} < X_{i, \mathbf{\Pi}_2^{-1}(i)}; X_{j,i} < X_{j, \mathbf{\Pi}_2^{-1}(j)} \right\};$$

**(2)** if  $(i, j) \in D_1^c \cap D_2$ , then necessarily  $(i, j) \in M_1^c \cap M_2$ , and, by stability of  $\mathbf{\Pi}_1$ ,

$$B_{(i,j)} = \left\{ X_{i, \mathbf{\Pi}_2^{-1}(i)} < X_{i, \mathbf{\Pi}_1^{-1}(i)}; X_{j, \mathbf{\Pi}_2^{-1}(j)} < X_{j, \mathbf{\Pi}_1^{-1}(j)} \right\};$$

**(3)** if  $(i, j) \in D_1 \cap D_2^c$ , then necessarily  $(i, j) \in M_1 \cap M_2^c$  and, by stability of  $\mathbf{\Pi}_2$ ,

$$B_{(i,j)} = \left\{ X_{i, \mathbf{\Pi}_1^{-1}(i)} < X_{i, \mathbf{\Pi}_2^{-1}(i)}; X_{j, \mathbf{\Pi}_1^{-1}(j)} < X_{j, \mathbf{\Pi}_2^{-1}(j)} \right\}.$$

Conditioned on the values

$$X_{i, \mathbf{\Pi}_1^{-1}(i)} = x_i, \quad (i \in [n]), \quad X_{i, \mathbf{\Pi}_2^{-1}(i)} = y_i, \quad (i \in [n] : \mathbf{\Pi}_1(i) \neq \mathbf{\Pi}_2(i)),$$

the events  $A_h, B_{(i,j)}$  are all independent. And, denoting the characteristic function of a set  $U \subset [0, 1]^{2n}$  by  $\chi(U)$ , we have  $P(A_h | \cdot) = x_h$ ,

$$P((B_{(i,j)})^c | \cdot) = \begin{cases} 1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j), & \text{Case (1),} \\ \chi(y_i \geq x_i \text{ or } y_j \geq x_j), & \text{Case (2),} \\ \chi(x_i \geq y_i \text{ or } x_j \geq y_j), & \text{Case (3).} \end{cases}$$

Therefore

$$P(\mathbf{\Pi}_1, \mathbf{\Pi}_2 \text{ are both stable} | \cdot) \\ = \prod_{h \in \text{Odd}_{1,2}} x_h \cdot \prod_{(i,j) \in D_1^c \cup D_2^c} [1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j)],$$

provided that  $\forall (i, j) \in M_1^c \cap M_2$ , we have  $y_i \geq x_i$  or  $y_j \geq x_j$  and  $\forall (i, j) \in M_1 \cap M_2^c$ , we have  $x_i \geq y_i$  or  $x_j \geq y_j$ . (The conditional probability is

zero otherwise.) Since the edges from  $M_1 \Delta M_2$  form the disjoint alternating circuits of length  $\geq 4$ , the condition means that for every such circuit  $\{i_1, i_2, \dots, i_\ell, \}$  [with  $(i_1, i_2) \in M_1, (i_2, i_3) \in M_2, \dots, (i_\ell, i_1) \in M_2$ , say], we have

$$\begin{array}{ccc} y_{i_1} \leq x_{i_1} & \text{or} & y_{i_2} \leq x_{i_2}, \\ y_{i_2} \geq x_{i_2} & \text{or} & y_{i_3} \geq x_{i_3}, \\ & & \vdots \\ y_{i_{\ell-1}} \leq x_{i_{\ell-1}} & \text{or} & y_{i_\ell} \leq x_{i_\ell}, \\ y_{i_\ell} \geq x_{i_\ell} & \text{or} & y_{i_1} \geq x_{i_1}. \end{array}$$

We may, of course, assume that all these inequalities are strict. Thus there are only two options on the circuit: either  $x_{i_1} > y_{i_1}, x_{i_2} < y_{i_2}, \dots, x_{i_\ell} < y_{i_\ell}$ , or  $x_{i_1} < y_{i_1}, x_{i_2} > y_{i_2}, \dots, x_{i_\ell} > y_{i_\ell}$ . (In both options, the inequalities alternate.) Application of Fubini's theorem completes the proof.  $\square$

### 3. ESTIMATION TOOLS

To estimate the integrals in Lemma 2.1 and Lemma 2.3 for  $n \rightarrow \infty$ , we will need the following claim, see [12], [14]:

**Lemma 3.1.** *Let  $X_1, \dots, X_\nu$  be independent  $[0, 1]$ -Uniforms. Let  $S = \sum_{i \in [\nu]} X_i$  and  $\mathbf{V} = \{V_i = X_i/S; i \in [\nu]\}$ , so that  $\sum_{i \in [\nu]} V_i = 1$ . Let  $\mathbf{L} = \{L_i; i \in [\nu]\}$  be the set of lengths of the  $\nu$  consecutive subintervals of  $[0, 1]$  obtained by selecting, independently and uniformly at random,  $\nu - 1$  points in  $[0, 1]$ . Then (with  $\chi(A)$  standing for the indicator of an event  $A$ ) the joint density  $f_{S, \mathbf{V}}(s, \mathbf{v})$ , ( $\mathbf{v} = (v_1, \dots, v_{\nu-1})$ ), of  $(S, V)$  is given by*

$$\begin{aligned} f_{S, \mathbf{V}}(s, \mathbf{v}) &= s^{\nu-1} \chi\left(\max_{i \in [\nu]} v_i \leq s^{-1}\right) \chi(v_1 + \dots + v_{\nu-1} \leq 1) \\ (3.1) \quad &\leq \frac{s^{\nu-1}}{(\nu-1)!} f_{\mathbf{L}}(\mathbf{v}), \quad v_\nu := 1 - \sum_{i=1}^{\nu-1} v_i; \end{aligned}$$

here  $f_{\mathbf{L}}(\mathbf{v}) = (\nu-1)! \chi(v_1 + \dots + v_{\nu-1} \leq 1)$  is the density of  $(L_1, \dots, L_{\nu-1})$ .

We will also use the classic identities, Andrews, Askey and Roy [3], Section 1.8:

$$(3.2) \quad \int_{\substack{\mathbf{x} \geq \mathbf{0} \\ x_1 + \dots + x_\nu \leq 1}} \prod_{i \in [\nu]} x_i^{\alpha_i} d\mathbf{x} = \frac{\prod_{i \in [\nu]} \alpha_i!}{(\nu + \alpha)!}, \quad \alpha := \sum_{i \in [\nu]} \alpha_i,$$

$$\int_{\substack{\mathbf{x} \geq \mathbf{0} \\ x_1 + \dots + x_\nu = 1}} \prod_{i \in [\nu]} x_i^{\alpha_i} dx_1 \dots dx_{\nu-1} = \frac{\prod_{i \in [\nu]} \alpha_i!}{(\nu - 1 + \alpha)!}.$$

The identity/bound (3.1) is useful since the random vector  $\mathbf{L}$  had been well studied. It is known, for instance, that

$$(3.3) \quad \mathbf{L} \stackrel{\mathcal{D}}{\equiv} \left\{ \frac{w_i}{\sum_{j \in [\nu]} w_j} \right\}_{i \in [\nu]},$$

where  $w_j$  are independent, exponentially distributed, with the same parameter, 1 say. Here is this property at work.

**Lemma 3.2.** (1) Let  $s \geq 2$ . If  $\varepsilon_\nu \rightarrow 0$ ,  $\varepsilon_\nu \gg \nu^{-\frac{1}{s+1}}$ . Then

$$(3.4) \quad \mathbb{P} \left( \left| \frac{\nu^{s-1}}{s!} \sum_{j \in [\nu]} L_j^s - 1 \right| \geq \varepsilon_\nu \right) = O \left( \exp \left( -c \varepsilon_\nu \nu^{\frac{1}{s+1}} \right) \right).$$

(2) For  $\nu$  even,

$$(3.5) \quad \mathbb{P} \left( \left| 2\nu \sum_{j \in [\nu/2]} L_j L_{j+\nu/2} - 1 \right| \geq \varepsilon_2 \right) = O \left( \exp \left( -c \varepsilon_\nu \nu^{\frac{1}{s+1}} \right) \right).$$

*Proof.* Observe that  $\mathbb{E}[W] = 1$ ,  $\mathbb{E}[W^s] = s!$ . Choose

$$a = \left( 1 + \frac{\varepsilon_\nu}{3} \right) s!, \quad b = 1 - \frac{\varepsilon_\nu}{3s},$$

so that  $a/b^s < (1 + \varepsilon)s!$ , for  $\nu$  sufficiently large. Then, denoting  $\mathcal{W}^{(\ell)} = \sum_j W_j^\ell$ ,

$$\begin{aligned} \mathbb{P} \left( \nu^{s-1} \sum_{j \in [\nu]} L_j^s \geq (1 + \varepsilon_\nu) s! \right) &= \mathbb{P} \left( \frac{\mathcal{W}^{(s)}}{(\mathcal{W}^{(1)})^s} \geq \frac{(1 + \varepsilon_\nu) s!}{\nu} \right) \\ &\leq \mathbb{P} \left( \mathcal{W}^{(s)} \geq a\nu \text{ or } \mathcal{W}^{(1)} < b\nu \right) \leq \mathbb{P} \left( \mathcal{W}^{(s)} \geq a\nu \right) + \mathbb{P} \left( \mathcal{W}^{(1)} < b\nu \right). \end{aligned}$$

Since  $\mathbb{E}[e^{-zW}] < \infty$  for every  $z \geq 0$ , the standard application of Chernoff's method yields

$$(3.6) \quad \mathbb{P} \left( \mathcal{W}^{(1)} < b\nu \right) \leq \exp(-\nu c(b)), \quad c(b) = b - 1 - \log b = \Theta(\varepsilon_\nu^2).$$

Bounding  $\mathbb{P}(\mathcal{W}^{(s)} \geq a\nu)$  is more problematic since  $\mathbb{E}[e^{zW^2}] = \infty$  for  $z > 0$ . Truncation to the rescue! Introduce  $V = \min\{W, n^\alpha\}$ , ( $\alpha < 1$ ); then

$$(3.7) \quad \mathbb{P}(W_j \not\equiv V_j, j \in [\nu]) \leq \nu \mathbb{P}(W \geq \nu^\alpha) = \nu e^{-\nu^\alpha} = e^{-\Theta(\nu^\alpha)}.$$

Further

$$\begin{aligned} \mathbb{E}[e^{n^{-s\alpha}V^s}] &= \int_0^{n^\alpha} e^{(n^{-\alpha}w)^s} e^{-w} dw + e^{1-\nu^\alpha} \\ &\leq 1 + \nu^{-s\alpha} \int_0^\infty w^s e^{-w} dw + O(\nu^{-2s\alpha}) \\ &= 1 + \nu^{-s\alpha} s! + O(\nu^{-2s\alpha}). \end{aligned}$$

So

$$\begin{aligned} \mathbb{P}\left(\sum_{j \in [\nu]} V_j^s \geq a\nu\right) &\leq \frac{(1 + \nu^{-s\alpha} s! + O(\nu^{-2s\alpha}))^\nu}{\exp(\nu(an^{-s\alpha}))} \\ &= \exp\left(-\nu^{1-s\alpha}(a - s!) + O(\nu^{1-2s\alpha})\right). \end{aligned}$$

Combining this bound with (3.7), we select the best  $\alpha = 1/(s+1)$  and obtain

$$(3.8) \quad \mathbb{P}(W^{(s)} \geq a\nu) \leq \exp(-\hat{c}\varepsilon_\nu \nu^{\frac{1}{s+1}}), \quad (\hat{c} > 0).$$

This bound combined with (3.6) prove that

$$\mathbb{P}\left(\nu^{s-1} \sum_{j \in [\nu]} L_j^s \geq (1 + \varepsilon_\nu) s!\right) = O\left(\exp(-c\varepsilon_\nu \nu^{\frac{1}{s+1}})\right).$$

Since  $\mathbb{E}[e^{-zW^s}] < \infty$  for all  $z > 0$ , there is no need for truncation, and we get

$$\mathbb{P}\left(\nu^{s-1} \sum_{j \in [\nu]} L_j^s \leq (1 - \varepsilon_\nu) s!\right) \leq e^{-\Theta(\nu\varepsilon_\nu^2)},$$

So (3.4) follows. The proof of (3.5) is similar, and we omit it.  $\square$

**Note.** In [13] we claimed that the probabilities in Lemma 3.2, for  $\varepsilon$  fixed, could be shown to be exponentially small, and used this claim also in [14] and [7], hoping to apply it again in this study. We have realized though that for the right tail of the sums' distributions we could get only a *sub*-exponential bound, see (3.8). Fortunately, with the new inequalities (2.4)-(2.7) put to use, the sub-exponential bounds (3.4) and (3.5) are all we need. The interested reader may see for themselves that the resulting proof provides a clear recipe for local changes in [13], [14] and [7], which make the thorny issue of exponential bounds go away completely.

In addition to the bounds (3.4), we will need

$$(3.9) \quad \mathbb{P} \left( \max_{j \in [\nu]} L_j^{(\nu)} \geq \frac{1.01 \log^2 \nu}{\nu} \right) \leq e^{-\log^2 \nu},$$

which directly follows from

$$\mathbb{P} \left( \max_{j \in [\nu]} L_j^{(\nu)} \geq x \right) \leq \nu(1-x)^{\nu-1}.$$

#### 4. ESTIMATES OF $\mathbb{E}[\mathcal{S}_n]$ AND $\mathbb{E}[\mathcal{O}_n]$ , RAMIFICATIONS

We need to identify a part of the cube  $[0, 1]^n$  that provides the dominant contribution to the integral in (2.1). This will allow us to estimate, sharply, the expected total length of the odd cycles in the random instance  $I_n$ . Many of the intermediate estimates can be traced back to [13], [14] and [7]. We begin with the pair of two new, instrumental, bounds for the products in the integrands expressing  $\mathbb{P}(\mathbf{\Pi}) := \mathbb{P}(\mathbf{\Pi} \text{ is stable})$  and  $\mathbb{P}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) := \mathbb{P}(\mathbf{\Pi}_1 \text{ and } \mathbf{\Pi}_2 \text{ are both stable})$ .

In the statement below and elsewhere we will write  $A_n \leq_b B_n$  as a shorthand for “ $A_n = O(B_n)$ , uniformly over parameters that determine  $A_n, B_n$ ”, when the expression for  $B_n$  is uncomfortably bulky for an argument of the big O notation. We will also write  $A_n \lesssim B_n$  if  $\limsup A_n/B_n \leq 1$ .

**Lemma 4.1.**

$$\begin{aligned} \prod_{\{i,j\} \notin D(\mathbf{\Pi})} (1 - x_i x_j) &\leq_b \exp\left(-\frac{s^2}{2}\right), \quad s := \sum_{i \in [n]} x_i, \\ \prod_{(i,j) \in D_1^c \cup D_2^c} [1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j)] \\ &\leq_b \exp\left(-\frac{s_1^2}{2} - \frac{s_2^2}{2} + \frac{s_{1,2}^2}{2}\right); \end{aligned}$$

here  $s_1 = \sum_i x_i$ ,  $s_2 = \sum_i y_i$ ,  $s_{1,2} = \sum_i (x_i \wedge y_i)$  and  $i$  runs through  $[n]$ .

*Proof.* (1) Using  $1 - z \leq e^{-z - z^2/2}$ , we have

$$(4.1) \quad \prod_{(i,j) \notin D(\mathbf{\Pi})} (1 - x_i x_j) \leq \exp\left(-\sum_{(i,j) \notin D(\mathbf{\Pi})} \left(x_i x_j + \frac{x_i^2 x_j^2}{2}\right)\right).$$

Here, using  $2ab \leq a^2 + b^2$ ,

$$(4.2) \quad \begin{aligned} \sum_{(i,j) \notin D(\mathbf{\Pi})} x_i x_j &= \frac{s^2}{2} - \frac{1}{2} \sum_{i \in [n]} x_i^2 - \sum_{i \in [n_1/2]} x_i x_{i+n_1/2} - \sum_{h \in \text{Odd}(\mathbf{\Pi})} x_h x_{\mathbf{\Pi}(h)} \\ &\geq \frac{s^2}{2} - \frac{3}{2} \sum_{i \in [n]} x_i^2. \end{aligned}$$

Analogously, and using  $\max_i x_i \leq 1$ ,

$$(4.3) \quad \sum_{(i,j) \notin D(\mathbf{\Pi})} x_i^2 x_j^2 \geq \frac{1}{2} \left( \sum_{i \in [n]} x_i^2 \right)^2 - \frac{3}{2} \sum_{i \in [n]} x_i^4$$

$$(4.4) \quad \geq \frac{1}{2} \left( \sum_{i \in [n]} x_i^2 \right)^2 - \frac{3}{2} \sum_{i \in [n]} x_i^2.$$

Therefore

$$(4.5) \quad \sum_{(i,j) \notin D(\mathbf{\Pi})} \left( x_i x_j + \frac{x_i^2 x_j^2}{2} \right) \geq \frac{s^2}{2} + \frac{1}{2} \left( \sum_{i \in [n]} x_i^2 \right)^2 - 3 \sum_{i \in [n]} x_i^2$$

$$(4.6) \quad \geq \frac{s^2}{2} - 4.5,$$

so that

$$\prod_{(i,j) \notin D(\mathbf{\Pi})} (1 - x_i x_j) \leq \exp \left( -\frac{s^2}{2} + 4.5 \right).$$

**(2)** Let  $M_i$  be the perfect matching on  $\text{Even}_{1,2} = [n] \setminus \text{Odd}_{1,2}$ , induced by  $\mathbf{\Pi}_i$ . Then  $M_1 \cap M_2$  is the set of matched pairs common to  $\mathbf{\Pi}_1$  and  $\mathbf{\Pi}_2$ , and  $M_1 \Delta M_2$  is the edge set of the even circuits, of length 4 at least, formed (in alternating fashion) by the matched pairs in  $M_1$  and  $M_2$ . So  $D_1 \cup D_2$  is the disjoint union of  $M_1 \cap M_2$ ,  $M_1 \Delta M_2$  the set of pairs  $(i, u_{\mathbf{\Pi}_{1,2}})$ .

So, given  $u_i, i \in [n]$ ,

$$(4.7) \quad \begin{aligned} & \sum_{(i \neq j) \in D_1^c \cap D_2^c} u_i u_j = \sum_{(i \neq j)} u_i u_j - \sum_{(i \neq j) \in D_1 \cup D_2} u_i u_j \\ & = \sum_{(i \neq j)} u_i u_j - \sum_{(i,j) \in M_1 \cap M_2} u_i u_j - \sum_{(i,j) \in M_1 \Delta M_2} u_i u_j - \sum_{i \in \text{Odd}_{1,2}} u_i u_{\mathbf{\Pi}_{1,2}(i)} \\ & = \frac{1}{2} \left( \sum_{i \in [n]} u_i \right)^2 - \frac{1}{2} \sum_{i \in [n]} u_i^2 - \sum_{(i,j) \in M_1 \cap M_2} u_i u_j - \sum_{(i,j) \in E_{1,2}} u_i u_j; \end{aligned}$$

here  $E_{1,2}$  is the edge set of the odd cycles and the even circuits, formed by  $\mathbf{\Pi}_1$  and  $\mathbf{\Pi}_2$ . This exact formula certainly implies that

$$\frac{1}{2} \left( \sum_{i \in [n]} u_i \right)^2 - 3 \sum_{i \in [n]} u_i^2 \leq \sum_{(i \neq j) \in D_1^c \cap D_2^c} u_i u_j \leq \frac{1}{2} \left( \sum_{i \in [n]} u_i \right)^2.$$

Therefore we bound

$$\begin{aligned}
& \sum_{(i \neq j) \in D_1^c \cap D_2^c} [x_i x_j + y_i y_j - (x_i \wedge y_i)(x_j \wedge y_j)] \\
& \geq \sum_{(i \neq j)} [x_i x_j + y_i y_j - (x_i \wedge y_i)(x_j \wedge y_j)] - 3 \sum_{i \in [n]} (x_i^2 + y_i^2) \\
& \geq \frac{s_1^2}{2} + \frac{s_2^2}{2} - \frac{s_{1,2}^2}{2} - 3 \sum_{i \in [n]} (x_i^2 + y_i^2).
\end{aligned}$$

Furthermore

$$\begin{aligned}
[x_i x_j + y_i y_j - (x_i \wedge y_i)(x_j \wedge y_j)]^2 & \geq [x_i x_j + y_i y_j - (x_i x_j \wedge y_i y_j)]^2 \\
& \geq \left( \frac{x_i x_j + y_i y_j}{2} \right)^2 \geq \frac{1}{8} (x_i^2 x_j^2 + y_i^2 y_j^2).
\end{aligned}$$

So

$$\begin{aligned}
& \sum_{(i \neq j) \in D_1^c \cap D_2^c} [x_i x_j + y_i y_j - (x_i \wedge y_i)(x_j \wedge y_j)]^2 \\
& \geq \frac{1}{8} \sum_{(i \neq j) \in D_1^c \cap D_2^c} (x_i^2 x_j^2 + y_i^2 y_j^2) \\
& \geq \frac{1}{16} \left( \sum_{i \in [n]} x_i^2 \right)^2 + \frac{1}{16} \left( \sum_{i \in [n]} y_i^2 \right)^2 - \sum_{i \in [n]} (x_i^4 + y_i^4).
\end{aligned}$$

As

$$\sum_{i \in [n]} (x_i^4 + y_i^4) \leq \sum_{i \in [n]} (x_i^2 + y_i^2),$$

we obtain

$$\begin{aligned}
& \prod_{(i \neq j) \in D_1^c \cap D_2^c} [1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j)] \\
& \leq \exp \left( -\frac{s_1^2}{2} - \frac{s_2^2}{2} + \frac{s_{1,2}^2}{2} \right) \\
& \times \exp \left[ -\frac{1}{32} \left( \sum_{i \in [n]} x_i^2 \right)^2 - \frac{1}{32} \left( \sum_{i \in [n]} y_i^2 \right)^2 + 4 \sum_{i \in [n]} (x_i^2 + y_i^2) \right].
\end{aligned}$$

It remains to observe that  $-\frac{z^2}{32} + 4z \leq 128$ .  $\square$

**4.1. Bounds for  $\mathbf{P}(\Pi)$ , probability of a fixed point, and the likely  $|\mathbf{Odd}(\Pi)|$ .** Here are our first applications of Lemma 4.1.

**Lemma 4.2.** Denoting  $m = |\text{Odd}(\mathbf{\Pi})|$ ,

$$P(\mathbf{\Pi}) \leq_b \begin{cases} \frac{e^{-n^{1/2}}}{(n+m-2)!!}, & \mathbf{\Pi} \text{ has a fixed point,} \\ \frac{1}{(n+m-1)!!}, & \mathbf{\Pi} \text{ has no fixed point.} \end{cases}$$

*Proof.* For the second case, by (2.1) and Lemma 4.1,

$$P(\mathbf{\Pi}) \leq_b \text{Int}(m) := \int_{\mathbf{x} \in [0,1]^n} \overbrace{\cdots}^n e^{-\frac{s^2}{2}} \prod_{h=n-m+1}^n x_h d\mathbf{x}.$$

Disregarding the constraint  $\max_i x_i \leq 1$ , and using (3.2), we obtain

$$(4.8) \quad \begin{aligned} \text{Int}(m) &\leq \frac{1}{(n+m-1)!} \int_0^\infty e^{-\frac{s^2}{2}} s^{n+m-1} ds \\ &= \frac{(n+m-2)!!}{(n+m-1)!} = \frac{1}{(n+m-1)!!}. \end{aligned}$$

If  $\mathbf{\Pi}$  has a fixed point  $h^*$ , then using

$$\prod_{k \neq h^*} (1 - x_k) \leq e^{-s}, \quad s := \sum_{k \neq h^*} x_k,$$

we obtain that

$$P(\mathbf{\Pi}) \leq_b \frac{1}{(n+m-2)!} \int_0^\infty e^{-s-\frac{s^2}{2}} s^{n+m-2} ds.$$

A quick glance at the integrand shows that the dominant contribution to the integral comes from  $s$  within, say,  $\log n$  distance from the integrand's maximum point

$$s^* = (n+m-2)^{1/2} - \frac{1}{2} + O(n^{-1/2}),$$

$(n+m-2)^{1/2}$  being the maximum point of  $e^{-s^2/2} s^{n+m-2}$ . So the above integral is of order

$$e^{-n^{1/2}} \int_0^\infty e^{-\frac{s^2}{2}} s^{n+m-2} ds = e^{-n^{1/2}} (n+m-3)!!,$$

whence

$$P(\mathbf{\Pi} \text{ is stable}) \leq_b \frac{e^{-n^{1/2}}}{(n+m-2)!!}.$$

□

Now the total number of permutations  $\mathbf{\Pi}$  of  $[n]$  with a fixed point and  $|\text{Odd}(\mathbf{\Pi})| = m$  is at most

$$n \binom{n-1}{m} m! (n-m-2)!! = \frac{n!}{(n-m-1)!!}.$$

**Corollary 4.3.**

$$P(\text{stable } \mathbf{\Pi}'\text{s have a fixed point}) = O(n^2 e^{-\sqrt{n}}) \rightarrow 0.$$

*Proof.* By Lemma 4.1 and the union bound, the probability in question is of order

$$\begin{aligned} & e^{-n^{1/2}} \sum_{m \geq 3} \frac{n!}{(n-m-1)!!(n+m-2)!!} \\ & \leq e^{-n^{1/2}} n \frac{n!}{(n-m-1)!!(n+m-2)!!} \Big|_{m=3} = O(n^2 e^{-n^{1/2}}). \end{aligned}$$

□

Our original proof in [14] was considerably more involved, and reliant on the problematic existence of the exponential bounds, the issue we touched upon in the previous sections, and will stop bringing up in the sequel.

From now on we focus on stable partitions without a fixed point. Here is another low hanging fruit.

**Corollary 4.4.** *Denoting by  $\text{Odd}(\mathbf{\Pi})$  the set of members in the odd cycles of stable partitions,*

$$P(|\text{Odd}(\mathbf{\Pi})| \geq n^{1/2} \log n) \leq_b \exp(-\log^2 n/3),$$

*i.e. with super-polynomially high probability (quite surely in terminology of [9]) the total length of all odd cycles is below  $n^{1/2} \log n$ .*

*Proof.* Denote  $m_n = \lceil n^{1/2} \log n \rceil$ . The total number of potential stable partitions with an even  $|\text{Odd}(\mathbf{\Pi})| = m \geq 4$  is at most

$$\binom{n}{m} m!(n-m-1)!! = \frac{n!}{(n-m)!!}.$$

So, by Lemma 4.1, Stirling formula, and the inequality

$$(1+x) \log(1+x) + (1-x) \log(1-x) \geq x^2,$$

the probability in question is of order

$$\begin{aligned} & \sum_{m=m_n}^n \frac{n!}{(n-m)!!(n+m-1)!!} \leq_b \sum_{m=m_n}^n \frac{n^n}{(n-m)^{\frac{n-m}{2}}(n+m)^{\frac{n+m}{2}}} \\ & \leq \sum_{m=m_n}^n \exp\left(-\frac{m^2}{2n}\right) \leq_b n^{1/2} \int_{x \geq m_n/n^{1/2}} e^{-\frac{x^2}{2}} dx \ll e^{-\log^2 n/3}. \end{aligned}$$

□

Focusing on the likely stable partitions, we may and will consider only the permutations  $\mathbf{\Pi}$  without a fixed point and with  $|\text{Odd}(\mathbf{\Pi})| \leq m_n$ .

**4.2. Sharp estimate of  $\mathbf{P}(\mathbf{\Pi})$ .** In steps, we will chop off the peripheral parts of the integration cube  $[0, 1]^n$  till we get to its part narrow enough to allow us to approximate the integrand in the formula (2.1) within  $1 + o(1)$  factor, so that the accumulative error cost is of order  $e^{-\Theta(\log^2 n)}$ .

**Step 1.** For the first reduction, we set  $s_n = n^{1/2} + 3 \log n$ , and define

$$(4.9) \quad \mathbf{P}_1(\mathbf{\Pi}) := \int_{\mathbf{x} \in C_1} \overbrace{\cdots}^n F(\mathbf{x}) d\mathbf{x}, \quad C_1 := \{\mathbf{x} \in [0, 1]^n : s \leq s_n\},$$

**Lemma 4.5.**

$$(4.10) \quad \mathbf{P}(\mathbf{\Pi}) - \mathbf{P}_1(\mathbf{\Pi}) \leq_b \frac{e^{-3 \log^2 n}}{(n+m-1)!!}.$$

*Proof.* By Lemma 4.1,

$$(4.11) \quad \mathbf{P}(\mathbf{\Pi}) - \mathbf{P}_1(\mathbf{\Pi}) \leq_b \frac{1}{(n+m-1)!} \int_{s \geq s_n} \exp\left(-\frac{s^2}{2}\right) s^{n+m-1} ds.$$

The integrand, write it as  $e^{h(s)}$ , attains its maximum at  $s_{n,m} = (n+m-1)^{1/2}$ , and

$$\begin{aligned} e^{h(s(n,m))} &= \exp\left(-\frac{n+m-1}{2}\right) (n+m-1)^{\frac{n+m-1}{2}} \\ &\leq_b n^{1/2} (n+m-2)!!. \end{aligned}$$

Further

$$\begin{aligned} h(s_n) &= h(s(n,m)) + (1+o(1)) \frac{h''(s(n,m))}{2} (s_n - s(n,m))^2 \\ &\leq h(s(n,m)) - 4 \log^2 n, \\ h'(s_n) &= -s_n + \frac{n+m-1}{s_n} \leq -5 \log n. \end{aligned}$$

Now, since  $h(s)$  is concave, we have

$$\int_{s \geq s_n} e^{h(s)} ds \leq e^{h(s_n)} \int_{s \geq s_n} \exp(h'(s_n)(s - s_n)) = \frac{e^{h(s_n)}}{-h'(s_n)}.$$

Therefore

$$\mathbf{P}(\mathbf{\Pi}) - \mathbf{P}_1(\mathbf{\Pi}) \leq_b \frac{n^{1/2} e^{-4 \log^2 n} (n+m-2)!!}{\log n (n+m-1)!} \leq \frac{e^{-3 \log^2 n}}{(n+m-1)!!}.$$

□

Next, motivated by the inequalities (4.2) and (4.3), we will derive sharp asymptotics, on progressively smaller  $C_j \subset C_1$ , for the leading sums  $\sum_{i \in [n_1]} x_i^2$ ,  $\sum_{i \in [n_1/2]} x_i x_{i+n_1/2}$ , ( $n_1 := n - m$ ), and obtain sufficiently strong

upper bounds for the secondary sums  $\sum_{h \in \text{Odd}(\mathbf{\Pi})} x_h^2$ ,  $\sum_{i \in [n]} x_i^4$  and  $\sum_{i \in [n]} x_i^6$ . We will end up with a rather sharp asymptotic formula for  $\prod_{(i,j)} (1 - x_i x_j)$  on the terminal dominant subset of  $C_1$ .

**Step 2.** With  $s := \sum_{i \in [n]} x_i$ , define  $\mathbf{u} = \{u_i = x_i/s : i \in [n]\}$ . Introduce  $t_1(\mathbf{u}) = \max_{i \in [n]} u_i$ . Define  $C_2 = \left\{ \mathbf{x} \in C_1 : t_1(\mathbf{u}) \leq 1.01 \frac{\log^2 n}{n} \right\}$ , and let  $P_j(\mathbf{\Pi})$  be the integral of  $F(\mathbf{x})$  over  $C_j$ . Introduce  $L_1, \dots, L_n$ , the lengths of the  $n$  consecutive subintervals of  $[0, 1]$  obtained by choosing, at random,  $n - 1$  points in  $[0, 1]$ . Applying Lemma 3.1, the identity (3.2) and Lemma 3.2 (1) with  $\nu = n$ , we have

$$\begin{aligned} P_1(\mathbf{\Pi}) - P_2(\mathbf{\Pi}) &\leq \int_{\mathbf{x} \geq \mathbf{0}} \cdots \int_{\mathbf{x} \geq \mathbf{0}}^n s^m e^{-\frac{s^2}{2}} \chi \left\{ t_1(\mathbf{u}) \geq 1.01 \frac{\log^2 n}{n} \right\} \prod_{h \in \text{Odd}(\mathbf{\Pi})} u_h \, d\mathbf{x} \\ &\leq \frac{\mathbb{E} \left[ \chi \left\{ \max_{i \in [n]} L_i \geq 1.01 \frac{\log^2 n}{n} \right\} \prod_{h=1}^m L_h \right]}{(n-1)!} \int_0^\infty e^{-\frac{s^2}{2}} s^{m+n-1} \, ds. \end{aligned}$$

By the union bound, the expected value is below

$$\begin{aligned} (4.12) \quad n \mathbb{E} \left[ \chi \left\{ L_n \geq 1.01 \frac{\log^2 n}{n} \right\} \prod_{h=1}^{m-1} L_h \right] &\leq n! \int_{\substack{z_1 + \dots + z_{n-1} \\ \leq 1 - 1.01 \frac{\log^2 n}{n}}}^{n-1} \prod_{h=1}^{m-1} z_h \, d\mathbf{z} \\ &= n! \left( 1 - 1.01 \frac{\log^2 n}{n} \right)^{n+m-2} \int_{z_1 + \dots + z_{n-1} \leq 1}^{n-1} \prod_{h=1}^{m-1} z_h \, d\mathbf{z} \leq \frac{n! e^{-1.01 \log^2 n}}{(n+m-2)!}. \end{aligned}$$

**Lemma 4.6.**

$$(4.13) \quad P_1(\mathbf{\Pi}) - P_2(\mathbf{\Pi}) \leq \frac{e^{-\log^2 n}}{(n+m-1)!}.$$

In addition, since  $m \leq m_n = \lceil n^{1/2} \log n \rceil$  and  $s \leq s_n = n^{1/2} + 3 \log n$ , it follows from (4.13) that on  $C_2$

$$(4.14) \quad \begin{aligned} \sum_{h \in \text{Odd}(\mathbf{\Pi})} x_h &\leq_b s n^{-1} \log^3 n, & \sum_{h \in \text{Odd}(\mathbf{\Pi})} x_h^2 &\leq_b n^{-1/2} \log^5 n, \\ \sum_{i=1}^n x_i^4 &\leq_b n^{-1} \log^8 n, & \sum_{i=1}^n x_i^6 &\leq_b n^{-2} \log^{12} n. \end{aligned}$$

**Step 3.** With  $\xi := \sum_{i \in [n_1]} x_i$ , define  $\mathbf{v} = \{v_i = x_i/\xi : i \in [n_1]\}$ . Introduce  $t_2(\mathbf{v}) = \sum_{i \in [n_1]} v_i^2$ . Define

$$C_3 = \left\{ \mathbf{x} \in C_2 : \left| \frac{n_1}{2} t_2(\mathbf{v}) - 1 \right| \leq n^{-\sigma} \right\}, \quad \sigma < 1/3.$$

Introduce  $\mathcal{L}_1, \dots, \mathcal{L}_{n_1}$ , the lengths of the  $n_1$  consecutive subintervals of  $[0, 1]$  in the partition of  $[0, 1]$  by the random  $n_1 - 1$  points. Analogously to **Step 2**, we have

$$\begin{aligned} P_2(\mathbf{\Pi}) - P_3(\mathbf{\Pi}) &\leq \int_{\mathbf{x} \geq \mathbf{0}} \overbrace{\dots}^n \int e^{-\frac{s^2}{2}} \chi \left\{ \left| \frac{n_1}{2} t_2(\mathbf{v}) - 1 \right| \geq n^{-\sigma} \right\} \prod_{h \in \text{Odd}(\mathbf{\Pi})} u_h \, d\mathbf{x} \\ &\leq \frac{\mathbb{P} \left( \left| \frac{n_1}{2} t_2(\mathcal{L}) - 1 \right| \geq n^{-\sigma} \right)}{(n_1 - 1)!(2m - 1)!} \iint_{\eta, \xi \geq 0} e^{-\frac{(\xi + \eta)^2}{2}} \xi^{n_1 - 1} \eta^{2m - 1} \, d\xi \, d\eta \\ &\leq \frac{e^{-\Theta(n^{1/3 - \sigma})}}{(n + m - 1)!} \int_0^\infty e^{-\frac{s^2}{2}} s^{n + m - 1} \, ds = \frac{e^{-\Theta(n^{1/3 - \sigma})}}{(n + m - 1)!!}. \end{aligned}$$

**Lemma 4.7.**

$$(4.15) \quad P_2(\mathbf{\Pi}) - P_3(\mathbf{\Pi}) \leq \frac{e^{-\Theta(n^{1/3 - \sigma})}}{(n + m - 1)!!}.$$

Similarly, with  $t_3(\mathbf{v}) := \sum_{i \in [n_1/2]} v_i v_{i+n_1/2}$  and

$$C_4 := \left\{ \mathbf{x} \in C_3 : |2n_1 t_3(\mathbf{v}) - 1| \leq n^{-\sigma} \right\},$$

**Lemma 4.8.**

$$(4.16) \quad P_3(\mathbf{\Pi}) - P_4(\mathbf{\Pi}) \leq \frac{e^{-\Theta(n^{1/3 - \sigma})}}{(n + m - 1)!!}.$$

Combining the estimates (4.10), (4.13), (4.15), we have

**Lemma 4.9.** *Let  $\mathbf{\Pi}$  be such that  $m = |\text{Odd}(\mathbf{\Pi})| \leq m_n$ . Then*

$$P(\mathbf{\Pi}) - P_4(\mathbf{\Pi}) \leq \frac{e^{-\Theta(\log^2 n)}}{(n + m - 1)!!},$$

where  $P_4(\mathbf{\Pi})$  is the integral of  $F(\mathbf{x})$  over  $C_4 \subset [0, 1]^n$  defined by the additional constraints: with  $s := \sum_{i \in [n]} x_i$ ,  $s_n = n^{1/2} + 3 \log n$ ,  $\xi = \sum_{i \in [n_1]} x_i$ ,

$$(4.17) \quad s \leq s_n, \quad \max_{i \in [n]} x_i \leq 1.02 \frac{s \log^2 n}{n};$$

$$(4.18) \quad \left| \frac{n_1 \sum_{i \in [n_1]} x_i^2}{2\xi^2} - 1 \right| \leq n^{-\sigma}, \quad \left| \frac{2n_1 \sum_{i \in [n_1/2]} x_i x_{i+n_1/2}}{\xi^2} - 1 \right| \leq n^{-\sigma}.$$

The constraint (4.18) involves only  $\{x_i\}_{i \in [n_1]}$ . Furthermore, given  $s$ , the constraint (4.17) imposes the uniform upper bound for the *individual* components  $x_i$ ,  $i \in [n]$ : no mixing the components  $x_i$ ,  $i \in [n_1]$ , and  $x_h$ ,  $h \in \text{Odd}(\mathbf{\Pi})$ , either. Also, this constraint *implies* that

$$(4.19) \quad \max_{i \in [n]} x_i \leq 4n^{-1/2} \log^2 n = o(1),$$

meaning that the constraint  $\max_i x_i \leq 1$  is superfluous. Moreover, the inequality (4.19) yields the *equality*

$$\prod_{(i,j) \notin D(\mathbf{\Pi})} (1 - x_i x_j) = \exp\left(- \sum_{(i,j) \notin D(\mathbf{\Pi})} \left(x_i x_j + \frac{x_i^2 x_j^2}{2}\right) + O\left(\sum_{i \in [n]} x_i^6\right)\right),$$

that holds uniformly for  $\mathbf{x} \in C_4$ , with the remainder term  $\ll n^{-1}$ , see (4.14). It is the matter of simple algebra to obtain from the constraints on  $C_4$ :

**Lemma 4.10.** *Uniformly for  $m \leq m_n$  and  $\mathbf{x} \in C_4$ ,*

$$(4.20) \quad F(\mathbf{x}) = \exp\left(-\frac{s^2}{2} \left(1 - \frac{3}{n}\right) - \frac{s^4}{n^2} + O(n^{-\sigma})\right) \prod_{h \in \text{Odd}(\mathbf{\Pi})} x_h.$$

Thus, introducing  $\eta = \sum_{h \in \text{Odd}(\mathbf{\Pi})} x_h$ , so that as  $s = \xi + \eta$ , within the factor  $1 + O(n^{-\sigma})$  the integrand depends on  $(\xi, \eta)$  and  $\prod_h x_h$ . Observe also that, on  $C_4$ ,

$$\max_{i \in [n_1]} \frac{x_i}{\xi} \sim \max_{i \in [n_1]} \frac{x_i}{s} \leq 1.02 \frac{\log^2 n}{n} \ll \frac{1}{\xi}.$$

So denoting  $\psi_n(s) = \frac{s^2}{2} \left(1 - \frac{3}{n}\right) + \frac{s^4}{n^2}$ , and applying Lemma 3.1, (3.1), we have:  $P_4(\mathbf{\Pi})$ , the integral of  $F(\mathbf{x})$  over  $C_4$ , is given by

$$(4.21) \quad \begin{aligned} P_4(\mathbf{\Pi}) &= (1 + O(n^{-\sigma})) \int_{\mathbf{x} \in C_4} \overbrace{\dots}^n e^{-\psi_n(\xi+\eta)} \prod_{h \in \text{Odd}(\mathbf{\Pi})} x_h d\mathbf{x} \\ &= (1 + O(n^{-\sigma})) \iint_{\xi+\eta \leq s_n} e^{-\psi_n(\eta+\xi)} \frac{\xi^{n_1-1}}{(n_1-1)!} \cdot \frac{\eta^{2m-1}}{(m-1)!} d\eta d\xi \\ &\quad \cdot \text{P}\left(\left|\frac{n_1}{2} \sum_{i \in [n_1]} \mathcal{L}_i^2 - 1\right| \leq n^{-\sigma}, \left|2n_1 \sum_{i \in [n_1/2]} \mathcal{L}_i \mathcal{L}_{i+n_1/2} - 1\right| \leq n^{-\sigma}\right). \end{aligned}$$

From the step (4) we know that the probability factor is at least  $1 - e^{-\log^2 n}$ . The double integral, denote it  $I_{n,m}$ , is given by

$$I_{n,m} = \frac{(2m-1)!}{(m-1)!(n+m-1)!} \int_{s \leq s_n} e^{-\psi_n(s)} s^{n+m-1} ds.$$

The integrand attains its maximum at  $\hat{s} = (n + m - 1)^{1/2} - \Theta(n^{-1/2})$ , so that  $s_n - \hat{s} \geq 2 \log n$ . and it is easy to show that

$$\int_{|s-\hat{s}| \geq \log n} e^{-\psi_n(s)} s^{n+m-1} ds \leq e^{-\Theta(\log^2 n)} \int_{|s-s^*| \leq \log n} e^{-\psi_n(a)} s^{n+m-1} ds.$$

Besides,  $s^4/n^2 = 1 + O(m/n)$  for  $|s - \hat{s}| \leq \log n$ . Therefore

$$\begin{aligned} I_{n,m} &= e^{-1+O(m/n)} \frac{(2m-1)!}{(m-1)!(n+m-1)!} \int_{s \geq 0} e^{-\frac{s^2(1-3/n)}{2}} s^{n+m-1} ds \\ &= e^{-1+O(m/n)} \frac{(2m-1)!}{(m-1)!(n+m-1)!} \cdot \frac{(n+m-2)!!}{(1-3/n)^{(n+m)/2}} \\ &= e^{1/2+O(m/n)} \frac{(2m-1)!}{(m-1)!(n+m-1)!!}. \end{aligned}$$

Since  $m/n = O(n^{-1/2} \log n)$ , and  $\sigma < 1/3$  in (4.21), we have proved

**Lemma 4.11.** *Uniformly for even  $m \leq m_n$  and  $\mathbf{\Pi}$  with  $|\text{Odd}(\mathbf{\Pi})| = m$ ,*

$$P_4(\mathbf{\Pi}) = (1 + O(n^{-\sigma})) \frac{e^{1/2}}{(n+m-1)!!}.$$

Consequently, by Lemma 4.9 ,

$$(4.22) \quad P(\mathbf{\Pi}) = (1 + O(n^{-\sigma})) \frac{e^{1/2}}{(n+m-1)!!}.$$

**Note.** The formula (4.22) works for  $m = 0$  as well, meaning that

$$P(\text{matching } \mathbf{\Pi} \text{ is stable}) = (1 + O(n^{-\sigma})) \frac{e^{1/2}}{(n-1)!!}.$$

So the expected number of stable *matchings* tends to  $e^{1/2}$  as  $n \rightarrow \infty$ , [13].

### 4.3. The expectations of the numbers of stable partitions and odd parties.

**Theorem 4.12.** *Let  $\mathcal{S}_n$  and  $\mathcal{O}_n$  denote the total number of odd stable partitions  $\mathbf{\Pi}$ , and the total number of odd cycles. Then*

$$(4.23) \quad \mathbb{E}[\mathcal{S}_n] = (1 + O(n^{-1/4})) \frac{\Gamma(1/4)}{\sqrt{\pi e} 2^{1/4}} n^{1/4},$$

$$(4.24) \quad \mathbb{E}[\mathcal{O}_n] \lesssim \frac{\Gamma(1/4)}{4\sqrt{\pi e} 2^{1/4}} n^{1/4} \log n.$$

*Proof.* For even  $m$ , let  $f(m)$  denote the total number of permutations of  $m$ , having only odd cycles, each of length 3 at least. For even  $k$ , let  $f(m, k)$  denote the total number of permutations of  $[m]$  having only  $k$  odd cycles, each of length 3, at least; so  $f(m) = \sum_k f(m, k)$ . Then the total number of permutations of  $[n]$  with  $k$  odd cycles, each of length 3 at least, with  $m$

elements overall, and even cycles of length 2 only, is  $\binom{n}{m} f(m, k)(n - m - 1)!!$ . So, by Lemma 4.11, we have

$$(4.25) \quad \mathbb{E}[\mathcal{S}_n] = (e^{1/2} + O(n^{-\sigma})) \sum_{m \leq m_n} \binom{n}{m} \frac{f(m)(n - m - 1)!!}{(n + m - 1)!!}.$$

A standard argument from permutation enumeration shows that

$$(4.26) \quad \sum_{m \geq 4} \frac{f(m)}{m!} x^m = \exp\left(\sum_{\text{odd } j \geq 3} \frac{x^j}{j}\right) = e^{-x} \sqrt{\frac{1+x}{1-x}}, \quad (|x| < 1).$$

So, using the saddle-point method (Flajolet and Sedgewick [5]),

$$(4.27) \quad f(m) = (e^{-1}\sqrt{2} + O(m^{-1})) \frac{(2m-1)!!}{2^m}.$$

With a bit of work, based on Stirling formula, it follows that

$$\begin{aligned} & (e^{1/2} + O(n^{-\sigma})) \binom{n}{m} \frac{f(m)(n - m - 1)!!}{(n + m - 1)!!} \\ &= (1 + O(n^{-\sigma} + m^{-1})) \sqrt{\frac{2}{\pi e}} \cdot m^{-1/2} \exp\left(-\frac{m^2}{2n}\right). \end{aligned}$$

Combining this formula with (4.25), and choosing  $\sigma = 1/4$ , we complete the proof of (4.23).

A bivariate extension of (4.26) is

$$\sum_{m \geq 4} \frac{x^m}{m!} \sum_{k \geq 2} y^k f(m, k) = \exp\left(y \sum_{\text{odd } j \geq 3} \frac{x^j}{j}\right).$$

Differentiating this identity at  $y = 1$ , we obtain

$$(4.28) \quad \begin{aligned} \sum_{m \geq 4} \frac{x^m}{m!} \sum_{k \geq 2} k f(m, k) &= \sum_{\text{odd } j \geq 3} \frac{x^j}{j} \exp\left(\sum_{\text{odd } j \geq 3} \frac{x^j}{j}\right) \\ &= \left(\frac{1}{2} \log \frac{1}{1-x} + \frac{1}{2} \log(1+x) - 1\right) e^{-x} \sqrt{\frac{1+x}{1-x}}. \end{aligned}$$

So, analogously to (4.27), we obtain

$$\sum_{k \geq 2} k f(m, k) = (1 + O(m^{-1})) \frac{e^{-1}\sqrt{2} \log m}{2} \cdot \frac{(2m-1)!!}{2^m}.$$

Combining this formula with the counterpart of (4.25), i.e. with

$$\mathbb{E}[\mathcal{O}_n] \leq (e^{1/2} + O(n^{-\sigma})) \sum_{m \leq m_n} \binom{n}{m} \frac{(n - m)!!}{(n + m - 1)!!} \sum_{k \geq 2} k f(m, k),$$

we have (4.24). □

Tan [15], [16] defined a maximum stable matching for an instance  $I$  as a matching  $M$  of maximum size (number of matched pairs) such that no pair of members, both having a partner in  $M$ , prefer each other to their partners. In short, no two members assigned in  $M$ , but not to each other, block  $M$ . He proved that a maximum stable matching has size  $(n - \mathcal{O})/2$ , (see also Manlove [10]).

**Corollary 4.13.** *Let  $\mathcal{M}_n$  denote the size of the maximum stable matching for the random instance  $I_n$ . Then*

$$\begin{aligned} \mathbb{E}[\mathcal{M}_n] &\geq \frac{n - cn^{1/4} \log n}{2}, \quad c = \frac{\Gamma(1/4)}{3\sqrt{\pi e} 2^{1/4}}, \\ \mathbb{P}\left(\mathcal{M}_n \geq \frac{n - n^{1/4} \log^2 n}{2}\right) &\geq 1 - O(\log^{-1} n), \end{aligned}$$

so that the number of members unassigned in the maximum stable matching is likely to be of order  $O(n^{1/4} \log^2 n)$ .

#### 4.4. A “maximally stable” matching in the random instance $I_n$ .

For a given set of preferences, Abraham, Biró and Manlove [1] (see also [10]) defined a “maximally stable” matching as a perfect matching  $M$  on  $[n]$  that is blocked by the smallest number of pairs,  $B(I_n)$ , of members not matched with each other in  $M$ . (Two members block  $M$  if they prefer each other to their partners in  $M$ .) A weaker corollary of the bound in [1] states that  $B(I_n) \leq d(I_n)\mathcal{O}(I_n)$ , where  $\mathcal{O}(I_n)$  is the number of odd parties (common to all stable partitions for  $I_n$ ) and  $d(I_n)$  is the length of the longest preference list.

Once we estimate  $R_{\max}$ , defined as the largest rank of a predecessor in the uniformly random instance  $I_n$ , we will be able to apply the ABM inequality via replacing  $d(I_n)$  with  $R_{\max}$ .

For a stable  $\mathbf{\Pi}$  (without a fixed point), introduce  $X(\mathbf{\Pi}) := \max_i X_{i, \mathbf{\Pi}^{-1}(i)}$ . Intuitively,  $\max_{\mathbf{\Pi}} X(\mathbf{\Pi})$  controls the worst predecessor’s rank. From Lemma 4.9, and the proof of Theorem 4.12, it follows that

$$\mathbb{P}\left(\max_{\mathbf{\Pi}} X(\mathbf{\Pi}) \geq \frac{\log^2 n}{n^{1/2}}\right) \leq e^{-\Theta(\log^2 n)}.$$

A bit more generally, for every  $\delta > 0$ ,

$$(4.29) \quad \mathbb{P}_\delta := \mathbb{P}\left(\max_{\mathbf{\Pi}} X(\mathbf{\Pi}) \geq \frac{\log^{1+\delta} n}{n^{1/2}}\right) \leq e^{-\Theta(\log^{1+\delta} n)}.$$

Denoting  $x_n = \frac{\log^{1+\delta} n}{n^{1/2}}$ , let  $R_i := |\{j \neq i : X_{i,j} \leq x_n\}|$ . Since  $X_{i,j}$  are independent  $[0, 1]$ -Uniforms, we have  $R_i \stackrel{\mathcal{D}}{\equiv} \text{Bin}(n-1, p = x_n)$ . Let  $c > 1$ ;

by the classic (Chernoff) bound for the tails of the binomial distribution,

$$\mathbb{P}(R_i \geq c(n-1)x_n) \leq \exp(-f(c)(n-1)x_n), \quad f(c) := 1 + c(\log c - 1).$$

So,  $\mathbb{P}(R_i \geq 2nx_n) \leq e^{-n^{1/2}/3}$  if  $n$  is large enough. Invoking (4.29), we have then

$$\begin{aligned} \mathbb{P}\left(R_{\max} \geq 2n^{1/2} \log^{1+\delta} n\right) &\leq P_\delta + \sum_{i \in [n]} \mathbb{P}\left(R_i \geq 2n^{1/2} \log^{1+\delta} n\right) \\ &\leq e^{-\Theta(\log^{1+\delta} n)} + ne^{-n^{1/2}/3}. \end{aligned}$$

Thus

**Lemma 4.14.** *For  $\delta > 0$  arbitrarily small, quite surely  $R_{\max}$  is of order  $n^{1/2} \log^{1+\delta} n$ .*

Combining Lemma 4.14 with (4.24) in Theorem 4.12, we have proved

**Corollary 4.15.** *With high probability, there exists a perfect matching which is blocked by at most  $n^{3/4}(\log n)^{2+\delta}$  unmatched pairs.*

#### 4.5. Likely range of $\mathcal{R}(\Pi)$ in a stable, fixed-point free, partition $\Pi$ .

In Lemma 2.6 we proved that  $\mathbb{P}_k(\Pi)$  the probability that  $\Pi$  is stable and the total rank of all  $n$  predecessors  $\mathcal{R}(\Pi)$  equals  $k$ , necessarily exceeding  $n + |\text{Odd}(\Pi)|$ , is given by

$$(4.30) \quad \begin{aligned} \mathbb{P}_k(\Pi) &= \int \cdots \int_{\mathbf{x} \in [0,1]^n} \left[ z^{\bar{k}} \right] F(\mathbf{x}, z) d\mathbf{x}, \\ F(\mathbf{x}, z) &:= \prod_{h \in \text{Odd}(\Pi)} x_h \cdot \prod_{(i,j) \notin D(\Pi)} (\bar{x}_i \bar{x}_j + zx_i \bar{x}_j + z\bar{x}_i x_j), \end{aligned}$$

where  $m := |\text{Odd}(\Pi)|$  and  $\bar{k} := k - (n + m)$ .

**Theorem 4.16.** *For  $\varepsilon \in (0, 1)$ ,*

$$\mathbb{P}\left(\max_{\Pi} \left| \frac{\mathcal{R}(\Pi)}{n^{3/2}} - 1 \right| \geq \varepsilon\right) \leq e^{-\Theta(\log^2 n)}.$$

*Proof.* Predictably, we will prove the claim via the union bound, i.e. summing the bounds of the respective probabilities for the individual partitions. It suffices then to consider the partitions  $\Pi$  with  $m \leq m_n = \lceil n^{1/2} \log n \rceil$ .

First of all, since  $F(\mathbf{x}, z)$  in (4.30) is a polynomial of  $z$  with non-negative coefficients, we have a Chernoff-type bound: for  $k := \lceil (1 + \varepsilon)n^{3/2} \rceil$ ,

$$\mathbb{P}(\mathcal{R}(\Pi) \geq k) \leq I(\Pi, k) := \int \cdots \int_{\mathbf{x} \in [0,1]^n} \inf_{z \geq 1} \left[ z^{-\bar{k}} F(\mathbf{x}, z) \right] d\mathbf{x}.$$

The integrand is, at most,

$$F(\mathbf{x}, 1) = F(\mathbf{x}) \leq_b e^{-\frac{s^2}{2}} \prod_{h \in \text{Odd}(\mathbf{\Pi})} x_h,$$

( $s = \sum_{i \in [n]} x_i$ ), see Lemma 4.1. Therefore the proof of Lemma 4.9 delivers, with only notational modification, that

$$(4.31) \quad I(\mathbf{\Pi}, k) - I_4(\mathbf{\Pi}, k) \leq \frac{e^{-\Theta(\log^2 n)}}{(n+m-1)!}.$$

Here  $I_4(\mathbf{\Pi}, k)$  is the integral of  $\inf_{z \geq 1} [z^{-\bar{k}} F(\mathbf{x}, z)]$  over  $C_4 \subset [0, 1]^n$ , defined by the additional constraints: with  $\xi = \sum_{i \in [n_1]} x_i$ , ( $n_1 := n - m$ ),

$$(4.32) \quad s \leq s_n := n^{1/2} + 3 \log n, \quad \max_{i \in [n]} x_i \leq 1.02 \frac{s \log^2 n}{n};$$

$$(4.33) \quad \left| \frac{n_1 \sum_{i \in [n_1]} x_i^2}{2\xi^2} - 1 \right| \leq n^{-\sigma}, \quad \left| \frac{2n_1 \sum_{i \in [n_1/2]} x_i x_{i+n_1/2}}{\xi^2} - 1 \right| \leq n^{-\sigma}.$$

Instead of looking for the best  $z = z(\mathbf{x}) \geq 1$  where  $z^{-\bar{k}} F(\mathbf{x}, z)$  attains, or is close to, its infimum, we confine ourselves to a sub-optimal  $z = z(s) \geq 1$  (i.e. dependent on  $s$  only), which makes  $z^{-\bar{k}} F(\mathbf{x}, z)$  suitably small for all  $\mathbf{x} \in C_4$ . Consider  $z \leq \frac{2\bar{k}}{sn}$ ; as we shall see shortly, the minimum point of an auxiliary bound for the integrand does satisfy this constraint.

Using  $1 + x \leq e^x$ , the constraints (4.32), (4.33) and  $z \leq \frac{2\bar{k}}{sn}$ , we have

$$\begin{aligned} \prod_{(i,j) \notin D(\mathbf{\Pi})} (\bar{x}_i \bar{x}_j + z x_i \bar{x}_j + z \bar{x}_i x_j) &= \prod_{(i,j) \notin D(\mathbf{\Pi})} (1 + (1 - 2z)x_i x_j + (z - 1)(x_i + x_j)) \\ &\leq \exp\left( \sum_{(i,j) \notin D(\mathbf{\Pi})} [(1 - 2z)x_i x_j + (z - 1)(x_i + x_j)] \right) \\ &\leq_b \exp\left( (1 - 2z) \frac{s^2}{2} + n(z - 1)s \right); \end{aligned}$$

therefore

$$z^{-\bar{k}} F(\mathbf{x}, z) \leq_b \exp\left( (1 - 2z) \frac{s^2}{2} + n(z - 1)s - \bar{k} \log z \right) \prod_{h \in \text{Odd}(\mathbf{\Pi})} x_h.$$

So, applying the identity (3.2),

$$(4.34) \quad I_4(\mathbf{\Pi}, k) \leq_b \frac{1}{(n+m-1)!} \int_0^{s_n} \exp(H(z, s)) ds,$$

$$H(z, s) := (1 - 2z) \frac{s^2}{2} + n(z - 1)s - \bar{k} \log z + (n + m - 1) \log s.$$

Let us use (4.34) to prove that

$$(4.35) \quad I_4(\mathbf{\Pi}, k) \leq \frac{e^{-\Theta(\varepsilon^2 n)}}{(n+m-1)!!}.$$

As a function of  $z$ ,  $H(z, s)$  is convex, and has its absolute minimum at

$$\bar{z} = \bar{z}(s) := \frac{\bar{k}}{(n-s)s} \sim \frac{\bar{k}}{ns} < \frac{2\bar{k}}{ns}.$$

This *decreasing* function of  $s$  is ‘‘Chernoff-admissible’’ when  $s$  is such that  $z(s) \geq 1$ . Let  $s_1$  be the smaller root of the (quadratic) equation  $\bar{z}(s) = 1$ :

$$s_1 > \frac{\bar{k}}{n}, \quad s_1 = \frac{\bar{k}}{n} + O(1) = (1 + \varepsilon)n^{1/2} + O(1).$$

Thus our best hope is a function

$$z(s) = \begin{cases} \bar{z}(s), & \text{if } s \leq s_1, \\ 1, & \text{if } s > s_1. \end{cases}$$

(i)  $s > s_1$ . Here  $s_1 \sim \frac{\bar{k}}{n} > (n+m-1)^{1/2}$ , the maximum point of

$$h(s) := H(1, s) = -\frac{s^2}{2} + (n+m-1) \log s.$$

So, arguing as in the proof of Lemma 4.5,

$$(4.36) \quad \begin{aligned} & \frac{1}{(n+m-1)!} \int_{s_1}^{s_n} \exp(H(1, s)) ds \leq b \frac{e^{h(s_1)}}{(-h'(s_1))(n+m-1)!} \\ & \leq \frac{e^{-\Theta(\varepsilon^2 n)}(n+m-2)!!}{(n+m-1)!} = \frac{e^{-\Theta(\varepsilon^2 n)}}{(n+m-1)!!}. \end{aligned}$$

(ii)  $s < s_1$ . Let  $\bar{h}(s) := H(z(s), s)$ . Since  $H_z(z(s), s) = 0$ , we have

$$\begin{aligned} \bar{h}'(s) &= H_s(z, s)|_{z=z(s)} = \left(s - \frac{\bar{k}}{n-s}\right) + \left(\frac{\bar{k}}{s} - n\right) + \frac{n+m-1}{s}, \\ \bar{h}''(s) &= 1 - \frac{\bar{k}}{(n-s)^2} - \frac{k-1}{s^2}. \end{aligned}$$

By the second formula, we have  $\bar{h}''(s) < 0$  for  $s \leq s_n$ , i.e.  $\bar{h}(s)$  is concave.

By the first formula, we have

$$\begin{aligned} \bar{h}'\left(\frac{\bar{k}}{n}\right) &\geq -(1+o(1))\frac{\bar{k}^2}{n^3} + \frac{n^2}{\bar{k}} \sim \frac{n^{1/2}}{1+\varepsilon} \rightarrow \infty, \\ \bar{h}'(s_1) &= \frac{(n+m-1) - s_1^2}{s_1} \sim -\frac{(2\varepsilon + \varepsilon^2)n^{1/2}}{1+\varepsilon} \rightarrow -\infty. \end{aligned}$$

Thus  $\max\{\bar{h}(s) : s \leq s_n\}$  is attained at a unique point  $s_2 \in [\bar{k}/n, s_1]$ ; in particular,  $s_1 - s_2 = O(1)$ . Since  $|\bar{h}''(s)| = O(n^{1/2})$ , it follows—via Taylor’s approximation of  $\bar{h}(s_1)$  (=  $h(s_1)$ ) at  $s_2$ —that  $\bar{h}(s_2) = h(s_1) + O(n^{1/2})$ .

Therefore, similarly to (4.36), we obtain

$$\frac{1}{(n+m-1)!} \int_0^{s_1} \exp(H(z(s), s)) ds \leq \frac{e^{-\Theta(\varepsilon^2 n)}}{(n+m-1)!!}.$$

This bound together with (4.36) imply (4.35), which in combination with (4.31) deliver

$$P(\mathcal{R}(\mathbf{\Pi}) \geq k) \leq \frac{e^{-\Theta(\log^2 n)}}{(n+m-1)!!}.$$

As in the proof of Corollary 4.4, it follows that

$$P(\exists \mathbf{\Pi} : \mathcal{R}(\mathbf{\Pi}) \geq (1+\varepsilon)n^{3/2}) \leq e^{-\Theta(\log^2 n)}.$$

Similarly

$$P(\exists \mathbf{\Pi} : \mathcal{R}(\mathbf{\Pi}) \leq (1-\varepsilon)n^{3/2}) \leq e^{-\Theta(\log^2 n)}.$$

□

## 5. $E[\mathcal{S}_n^2]$ AND THE EXPECTED NUMBER OF MEMBERS WITH MULTIPLE STABLE PREDECESSORS

First of all

$$E[(\mathcal{S}_n)_2] = \sum_{\mathbf{\Pi}_1 \neq \mathbf{\Pi}_2} P(\mathbf{\Pi}_1, \mathbf{\Pi}_2),$$

where  $P(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  is the probability that  $\mathbf{\Pi}_1$  and  $\mathbf{\Pi}_2$  are both stable. By Lemma 2.3,

$$(5.1) \quad \begin{aligned} P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) &= \int \cdots \int_{\mathbf{x}, \mathbf{y} \in [0,1]^n} F(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y}, \\ F(\mathbf{x}, \mathbf{y}) &:= \prod_h x_h \cdot \prod_{(i \neq j)} [1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j)], \\ d\mathbf{x} &= \prod_{i \in [n]} dx_i, \quad d\mathbf{y} = \prod_{i \in [n]: \mathbf{\Pi}_1(i) \neq \mathbf{\Pi}_2(i)} dy_i, \end{aligned}$$

where  $h \in \text{Odd}(\mathbf{\Pi}_{1,2})$ ,  $(i \neq j) \in D_1^c \cap D_2^c$ ,  $(D_t = D(\mathbf{\Pi}_t))$ ,  $y_i = x_i$  if  $\mathbf{\Pi}_1(i) = \mathbf{\Pi}_2(i)$ , and for every circuit  $\{i_1, \dots, i_\ell\}$ ,  $(\ell \geq 4)$ , formed by alternating pairs matched either in  $\mathbf{\Pi}_1$  or  $\mathbf{\Pi}_2$ , we have :

$$(5.2) \quad \begin{aligned} &\text{either } x_{i_1} > y_{i_1}, x_{i_2} < y_{i_2}, \dots, x_{i_\ell} < y_{i_\ell}, \\ &\text{or } x_{i_1} < y_{i_1}, x_{i_2} > y_{i_2}, \dots, x_{i_\ell} > y_{i_\ell}. \end{aligned}$$

Let  $\mu = \mu(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  be the total number of these circuits, and  $2\nu = 2\nu(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$ , be their total length. Obviously, there are  $2^\mu$  ways to select one of two ‘‘alternation’’ sequences described in (5.2) for each of the  $\mu$  circuits. Whatever the choice, there are exactly  $\nu$  vertices  $i$ , on those circuits, where  $y_i > x_i$  and

$\nu$  vertices where  $y_i < x_i$ . Let  $A$  and  $B$  denote the correspondent subsets,  $|A| = |B| = \nu$ . So

$$(5.3) \quad \begin{aligned} y_i &> x_i, & \text{if } i \in A; & \quad y_i < x_i, & \text{if } i \in B, \\ y_i &= x_i, & \text{if } i \in [n] \setminus (A \cup B) = \text{Odd}_{1,2} \cup (\text{Even}_{1,2} \setminus (A \cup B)), \end{aligned}$$

$\text{Even}_{1,2} := [n] \setminus \text{Odd}_{1,2}$ . The individual contributions of these  $2^\mu$  choices of the inequalities along the circuits to the integral in (5.1) are all the same. This means that  $P(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  equals the RHS integral in (5.1), with inequalities (5.3) instead of (5.2), *times*  $2^\mu$ .

As in the previous section, we need first to identify the subrange of  $(\mathbf{x}, \mathbf{y})$  that provides an asymptotically dominant contribution to the integral, and second to find a sharp approximation for that contribution. Like Theorem 4.12, the key instrument is the bound for the double-indexed product in the definition of  $F(\mathbf{x}, \mathbf{y})$  proved in Lemma 4.1:

$$(5.4) \quad \begin{aligned} F(\mathbf{x}, \mathbf{y}) &\leq_b \exp\left(-\frac{s_1^2}{2} - \frac{s_2^2}{2} + \frac{s_{1,2}^2}{2}\right) \prod_h x_h, \\ s_1 &:= \sum_{i \in [n]} x_i, \quad s_2 := \sum_{i \in [n]} y_i, \quad s_{1,2} := \sum_{i \in [n]} (x_i \wedge y_i). \end{aligned}$$

Here  $(\mathbf{x}, \mathbf{y})$  are subject to the constraints (5.3). To make use of this bound, we change the variables of integration:

$$(5.5) \quad x'_i = \begin{cases} x_i - y_i, & i \in B, \\ x_i, & i \notin B, \end{cases} \quad y'_i = \begin{cases} y_i - x_i, & i \in A, \\ y_i, & i \notin A. \end{cases}$$

Here  $\mathbf{x}', \mathbf{y}' \in [0, 1]^n$ , such that  $x'_i = y'_i$  if  $\mathbf{\Pi}_1(i) = \mathbf{\Pi}_2(i)$ , and the Jacobian  $\partial(\mathbf{x}, \mathbf{y})/\partial(\mathbf{x}', \mathbf{y}')$  equals 1. Furthermore, switching to  $(\mathbf{x}', \mathbf{y}')$  and introducing

$$(5.6) \quad \xi_1 = \sum_{i \in [n] \setminus B} x'_i + \sum_{i \in B} y'_i, \quad \xi_2 = \sum_{i \in B} x'_i, \quad \xi_3 = \sum_{i \in A} y'_i,$$

we obtain

$$(5.7) \quad \begin{aligned} &-\frac{s_1^2}{2} - \frac{s_2^2}{2} + \frac{s_{1,2}^2}{2} = -\frac{1}{2} \left( \sum_{i \in [n] \setminus B} x'_i + \sum_{i \in B} y'_i + \sum_{i \in B} x'_i \right)^2 \\ &-\frac{1}{2} \left( \sum_{i \in [n] \setminus B} x'_i + \sum_{i \in B} y'_i + \sum_{i \in A} y'_i \right)^2 + \frac{1}{2} \left( \sum_{i \in [n] \setminus B} x'_i + \sum_{i \in B} y'_i \right)^2 \\ &= -\frac{1}{2} (\xi_1 + \xi_2)^2 - \frac{1}{2} (\xi_1 + \xi_3)^2 + \frac{1}{2} \xi_1^2 = -\frac{1}{2} (\xi_1 + \xi_2 + \xi_3)^2 + \xi_2 \xi_3. \end{aligned}$$

Notice that

$$(5.8) \quad \begin{aligned} \xi_1 + \xi_2 + \xi_3 &= \sum_{i \in [n]} x'_i + \sum_{i \in A \cup B} y'_i = \sum_{i \in [n]} (x_i \vee y_i), \\ \xi_2 + \xi_3 &= \sum_{i \in B} x'_i + \sum_{i \in A} y'_i = \sum_{i \in [n]} |x_i - y_i|. \end{aligned}$$

In full analogy with the case of  $\mathbb{E}[\mathcal{O}_n]$ , the bound (5.4) and the identity (5.7) will allow us to shrink, in several steps, the range of  $(\mathbf{x}, \mathbf{y})$  to a core range, on which the integrand  $F(\mathbf{x}, \mathbf{y})$  can be sharply approximated.

(1) Recall that we consider the partitions  $\mathbf{\Pi}$  with the total length of all odd cycles  $m = m(\mathbf{\Pi}) \leq m_n = \lceil n^{1/2} \log n \rceil$ . Our first step is to dispense with the pairs  $(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  of the partitions such that  $2\nu = 2\nu(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \geq 2m_n$ .

**Lemma 5.1.**

$$\begin{aligned} \mathbb{E}[(\mathcal{S}_n)_2] - \mathbb{E}_1[(\mathcal{S}_n)_2] &\leq e^{-\Theta(\log^2 n)}, \\ \mathbb{E}_1[(\mathcal{S}_n)_2] &:= \sum_{\nu(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq m_n} \mathbb{P}(\mathbf{\Pi}_1, \mathbf{\Pi}_2). \end{aligned}$$

*Proof.* By the equations (5.4) and (5.7), and the identity (3.2), we have

$$(5.9) \quad \begin{aligned} \mathbb{P}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) &\leq_b 2^\mu \int \cdots \int_{\mathbf{x}', \mathbf{y}' \geq \mathbf{0}} \exp\left(-\frac{1}{2} \left(\sum_j \xi_j\right)^2 + \xi_2 \xi_3\right) \\ &\quad \times \left( \prod_{h \in \text{Odd}_{1,2}} x'_h \right) \prod_{i \in [n]} dx'_i \prod_{j \in A \cup B} dy'_j \\ &= 2^\mu \iiint_{\xi_j \geq \mathbf{0}} \exp\left(-\frac{1}{2} \left(\sum_j \xi_j\right)^2 + \xi_2 \xi_3\right) \frac{\xi_1^{n+m-1}}{(n+m-1)!} \cdot \frac{(\xi_2 \xi_3)^{\nu-1}}{[(\nu-1)!]^2} d\xi. \end{aligned}$$

Expanding  $\exp(\xi_2 \xi_3) = \sum_{k \geq 0} \xi_2^k \xi_3^k / k!$  and using again, term-wise, (3.2), we obtain

$$(5.10) \quad \begin{aligned} \mathbb{P}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) &\leq_b 2^\mu \sum_{k \geq 0} s(n+m, \nu, k), \\ s(n+m, \nu, k) &:= \frac{[(\nu-1+k)!]^2}{[(\nu-1)!]^2 k! (n+m+2(\nu+k)-1)!}. \end{aligned}$$

For  $m=0$  this sum was estimated in [14]. For our case the estimate from [14] becomes

$$(5.11) \quad \sum_{k \geq 0} s(n+m, \nu, k) \leq_b n \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} (n+m)^{-\nu}.$$

Furthermore, the number of ordered pairs  $(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  with parameters  $m$ ,  $\nu$  and  $\mu$  is

$$(5.12) \quad \binom{n}{m} f(m) \binom{n-m}{2\nu} (n-m-2\nu-1)!! \cdot 2^\mu f(2\nu, \mu);$$

here, as we recall,  $f(m)$  is the total number of permutations of  $[m]$  with only odd cycles of length 3 or more, and  $f(2\nu, \mu)$  is the total number of circuit partitions of  $[2\nu]$  with  $\mu$  circuits, each of even length 4 at least. The factor  $2^\mu$  counts the total number of ways to assign, in the alternating fashion, the edges of the circuits to the matching sets of  $\mathbf{\Pi}_1$  and  $\mathbf{\Pi}_2$ . Clearly then,  $2^\mu f(2\nu, \mu)$  is the total number of permutations of  $[2\nu]$  with only even cycles, of length 4 at least. We add that  $(n-m-2\nu-1)!!$  is the total number of ways to form the  $(n-m-2\nu)/2$  matched pairs out of  $n-m-2\nu$  elements outside the circuits, i.e. the pairs *common* to  $\mathbf{\Pi}_1$  and  $\mathbf{\Pi}_2$ .

So, by (5.10), (5.11) and (5.12), we obtain

$$(5.13) \quad \sum_{\nu(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \geq m_n} P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq_b n \sum_{m \leq m_n} \binom{n}{m} f(m) \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} \\ \times \sum_{\nu \geq m_n} \binom{n-m}{2\nu} (n-m-2\nu-1)!! (n+m)^{-\nu} \sum_{\mu} 2^{2\mu} f(2\nu, \mu).$$

Now  $f(m) \leq m!$ , and from [14] (Appendix) it follows that

$$(5.14) \quad \sum_{\mu} 2^{2\mu} f(2\nu, \mu) = e^{-1+O(\nu^{-1})} (2\nu)! = O((2\nu)!).$$

Also

$$\frac{(n-m-2\nu-1)!!}{(n-m-2\nu)!} = \left[ 2^{\frac{n-m-2\nu}{2}} \left( \frac{n-m-2\nu}{2} \right)! \right]^{-1}.$$

So the bound (5.13) yields

$$\sum_{\nu(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \geq m_n} P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq_b n! n \cdot \sum_{m \leq m_n} \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} \\ \times \sum_{\nu \geq m_n} \left[ (n+m)^\nu 2^{\frac{n-m-2\nu}{2}} \left( \frac{n-m-2\nu}{2} \right)! \right]^{-1} \\ \leq n! n^2 \cdot \sum_{m \leq m_n} \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} \left[ (n+m)^{m_n} 2^{\frac{n-m-2m_n}{2}} \left( \frac{n-m-2m_n}{2} \right)! \right]^{-1},$$

since in the  $\nu$ -sum the terms decrease with  $\nu$ . Applying Stirling formula for the two factorials and using  $m \leq m_n \ll n^{2/3}$  in the expansions of  $\log(1+z)$ ,

( $z = m/n$ ,  $z = -\frac{m+2m_n}{n}$ ), we transform this bound into

$$(5.15) \quad \sum_{\nu(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \geq m_n} P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq n^2 \cdot \sum_{m \leq m_n} \exp\left(-\frac{(m+2m_n)^2}{4n}\right) \leq e^{-0.99 \log^2 n}.$$

□

From now on we will consider only *admissible* pairs  $(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$ , i.e. those satisfying  $m(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq m_n$  and  $\nu(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq m_n$ .

(2) For the admissible pairs  $(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$ , we can discard large parts of the  $(\mathbf{x}, \mathbf{y})$ 's range, like we did for individual partitions  $\mathbf{\Pi}$  in the case of  $E[\mathcal{S}_n]$ .

For a generic set  $\mathcal{C}$  of  $(\mathbf{x}, \mathbf{y})$  with  $\mathbf{x}, \mathbf{y} \in [0, 1]^n$ , we define

$$P_{\mathcal{C}}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) = \int \cdots \int_{(\mathbf{x}, \mathbf{y}) \in \mathcal{C}} F(\mathbf{x}, \mathbf{y}) \, d\mathbf{x} d\mathbf{y}, \quad E_{\mathcal{C}}[(\mathcal{S}_n)_2] = \sum_{\mathbf{\Pi}_1, \mathbf{\Pi}_2} P_{\mathcal{C}}(\mathbf{\Pi}_1, \mathbf{\Pi}_2).$$

**Lemma 5.2.** *Introducing  $s_n = n^{1/2} + 6 \log n$ , and  $\mathcal{C}_1 = \left\{ \mathbf{x}, \mathbf{y} : \sum_{i \in [n]} (x_i \vee y_i) \leq s_n \right\}$ , we have*

$$E_1[(\mathcal{S}_n)_2] - E_{\mathcal{C}_1}[(\mathcal{S}_n)_2] \leq e^{-\Theta(\log^2 n)}.$$

*Proof.* We already observed, (5.8), that  $\sum_i (x_i \vee y_i) = \sum_j \xi_j$ . So, similarly to (5.9)-(5.10) we have:

$$(5.16) \quad \begin{aligned} & P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_1}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \\ & \leq_b 2^\mu \iiint_{\xi_1 + \xi_2 + \xi_3 \geq s_n} \exp\left(-\frac{1}{2} \left(\sum_j \xi_j\right)^2 + \xi_2 \xi_3\right) \frac{\xi_1^{n+m-1}}{(n+m-1)!} \cdot \frac{(\xi_2 \xi_3)^{\nu-1}}{[(\nu-1)!]^2} \, d\xi \\ & = 2^\mu \sum_{k \geq 0} \frac{[(\nu-1+k)!]^2}{[(\nu-1)!]^2 k! (n+m+2(\nu+k)-1)!} \\ & \quad \times \int_{s \geq s_n} \exp\left(-\frac{s^2}{2}\right) s^{n+m+2(\nu+k)-1} \, ds. \end{aligned}$$

(Relaxing the constraint on  $s$  to  $s \geq 0$  we get back to (5.10).) The last integrand attains its maximum at

$$s_{\max} = (n+m+2(\nu+k)-1)^{1/2},$$

which is below  $s_n - 3 \log n$  if  $k \leq m_n$ . Let  $S_{\leq m_n}$  and  $S_{> m_n}$  denote the sub-sums of the sum above, for  $k \leq m_n$  and  $k > m_n$  respectively. Then,

expanding integration to  $[0, \infty)$ , we obtain

$$\begin{aligned} S_{>m_n} &\leq \sum_{k>m_n} \frac{[(\nu-1+k)!]^2}{[(\nu-1)!]^2 k! (n+m+2(\nu+k)-1)!!} \\ &\leq_b \frac{[(\nu+m_n)!]^2}{[(\nu-1)!]^2 m_n! (n+m+2(\nu+m_n)+1)!!}; \end{aligned}$$

since  $\nu \leq m_n$ , the ratio of the consecutive terms in the sum is below  $2/3$ . Dropping  $[(\nu-1)!]^2$  in the denominator and using the Stirling formula for the other factorials, we simplify the bound to

$$(5.17) \quad S_{>m_n} \leq_b \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} (n+m)^{-(\nu+m_n)}.$$

The bound is smaller than the bound (5.11) for the full sum of  $s(n+m, \nu, k)$  by the factor  $(n+m)^{m_n}$ . Turn to  $S_{\leq m_n}$ . This time the bottom integral over  $s \geq s_n$  in (5.16) is small, compared to the integral over all  $s \geq 0$ , because for  $k \leq m_n$  the maximum point of the integrand is at distance  $3 \log n$ , at least, from the interval  $[s_n, \infty)$ . More precisely, using the argument in the proof of Lemma 4.5, we have

$$\int_{s \geq s_n} \exp\left(-\frac{s^2}{2}\right) s^{n+m+2(\nu+k)-1} ds \leq_b e^{-8 \log^2 n} (n+m+2(\nu+k)-2)!!.$$

Therefore

$$\begin{aligned} (5.18) \quad S_{\leq m_n} &\leq_b e^{-8 \log^2 n} \sum_{k \leq m_n} \frac{[(\nu-1+k)!]^2}{[(\nu-1)!]^2 k! (n+m+2(\nu+k)-1)!!} \\ &\leq e^{-8 \log^2 n} \sum_{k \geq 0} \frac{[(\nu-1+k)!]^2}{[(\nu-1)!]^2 k! (n+m+2(\nu+k)-1)!!} \\ &= e^{-8 \log^2 n} \sum_{k \geq 0} s(n+m, \nu, k). \end{aligned}$$

Combining (5.17), (5.18) and (5.11) we transform the inequality (5.16) into

$$(5.19) \quad P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_1}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq e^{-\Theta(\log^2 n)} 2^\mu \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} (n+m)^{-\nu}.$$

So, like the part **(1)** in the proof of Lemma 5.1,

$$\begin{aligned} (5.20) \quad &\sum_{\mathbf{\Pi}_1, \mathbf{\Pi}_2} [P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_1}(\mathbf{\Pi}_1, \mathbf{\Pi}_2)] \leq e^{-\Theta(\log^2 n)} n! m_n \\ &\times \sum_{m \leq m_n} \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} \left[ (n+m)^0 2^{\frac{n-m-2 \cdot 0}{2}} \left( \frac{n-m-2 \cdot 0}{2} \right)! \right]^{-1} \\ &= e^{-\Theta(\log^2 n)}. \end{aligned}$$

□

We need some additional reduction of the last range  $\mathcal{C}_2$ . The bound (5.4) will continue to be the key tool, until the resulting range is narrow enough to permit a sufficiently sharp bound of the double product

$$G(\mathbf{x}, \mathbf{y}) = \prod_{(i \neq j) \in D_1^c \cap D_2^c} [1 - x_i x_j - y_i y_j + (x_i \wedge y_i)(x_j \wedge y_j)]$$

in (5.1). Define  $\mathcal{N} = \mathcal{N}(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  and  $\mathcal{M} = \mathcal{M}(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  as the vertex set of all odd cycles and even cycles, of length 4 or more, and the vertex set of the edges common to both partitions, respectively. So  $|\mathcal{N}| = m + 2\nu$ , and  $|\mathcal{M}| = n - (m + 2\nu)$ . Arguing as in the proof of Lemma 4.1, but retaining more terms, we have

$$(5.21) \quad G(\mathbf{x}, \mathbf{y}) \leq \exp\left(-\frac{s_1^2}{2} - \frac{s_2^2}{2} + \frac{s_{1,2}^2}{2} + \frac{1}{2} \sum_{i \in \mathcal{M}} x_i^2 + \sum_{(i \neq j) \in M_1 \cap M_2} x_i x_j\right) - \frac{1}{4} \left( \sum_{i \in [n]} (x_i \wedge y_i)^2 \right)^2 + O\left( \sum_{i \in \mathcal{N}} (x_i^2 + y_i^2) \right) + O\left( \sum_{i \in [n]} x_i^4 \right).$$

Thus we have to find sharp approximations of the three explicit sums and to establish the  $o(1)$  bounds of the remainders for almost all  $(\mathbf{x}, \mathbf{y}) \in \mathcal{C}_1$ . With those approximations at hand we will obtain an explicit upper bound for  $\mathbb{E}[(\mathcal{S}_n)_2]$ . For brevity will not present a proof of a matching lower bound.

(3) By (5.5) and (5.6),  $s := \xi_1 + \xi_2 + \xi_3 = \sum_{i \in [n]} x'_i + \sum_{i \in A \cup B} y'_i$ .

**Lemma 5.3.** *Define  $\mathbf{u}' = \{u'_i\}_{i \in [n]}$ , where  $u'_i = x'_i/s$ , for  $i \in [n]$ , and  $u'_i = y'_i/s$  for  $i \in A \cup B$ . Define  $T_1(\mathbf{u}') = \max_i u'_i$ . For*

$$\mathcal{C}_2 := \left\{ (\mathbf{x}, \mathbf{y}) \in \mathcal{C}_1 : T_1(\mathbf{u}') \leq 1.01 \frac{\log^2 n}{n} \right\},$$

we have

$$P_{\mathcal{C}_1}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_2}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq 2^\mu e^{-\Theta(\log^2 n)} \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} n^{-\nu}.$$

*Proof.* Introduce  $L'_1, \dots, L'_{n+2\nu}$ , the intervals lengths in the random partition of  $[0, 1]$  by the  $n+2\nu-1$  random points. Analogously to (5.9), but using

the sharper inequality in Lemma 3.1, (3.1), we have: with  $s := \xi_1 + \xi_2 + \xi_3$ ,

$$\begin{aligned}
& P_{\mathcal{C}_1}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_2}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \\
& \leq_b 2^\mu \int \cdots \int_{\substack{\mathbf{x}', \mathbf{y}' \geq \mathbf{0} \\ T_1(\mathbf{u}') > 1.01 \frac{\log^2 n}{n}}} e^{-\frac{s^2}{2} + \xi_2 \xi_3} \prod_{h \in \text{Odd}_{1,2}} x_h \prod_{i \in [n]} dx'_i \prod_{j \in A \cup B} dy'_j \\
& \leq \frac{2^\mu}{(n + 2\nu - 1)!} \mathbb{E} \left[ \chi \left( T_1(\mathbf{L}') \geq 1.01 \frac{\log^2 n}{n} \right) \prod_{h \in \text{Odd}_{1,2}} L'_h \right] \\
& \quad \times \int \cdots \int_{\mathbf{x}', \mathbf{y}' \geq \mathbf{0}} e^{-\frac{s^2}{2} + \xi_2 \xi_3} s^m \prod_{i \in [n]} dx'_i \prod_{j \in A \cup B} dy'_j.
\end{aligned}$$

Arguing as in (4.12), the expectation factor is less than

$$e^{-1.01 \log^2 n} \frac{(n + 2\nu)!}{(n + m + 2\nu - 2)!}.$$

The integral is less than

$$\begin{aligned}
I_n(m, \nu) & := \iiint_{\xi_j \geq 0} e^{-\frac{s^2}{2} + \xi_2 \xi_3} s^m \frac{\xi_1^{n-1}}{(n-1)!} \cdot \frac{(\xi_2 \xi_3)^{\nu-1}}{[(\nu-1)!]^2} d\xi \\
& = \sum_{k \geq 0} \frac{[(\nu-1+k)!]^2}{(n-1)! [(\nu-1)!]^2 k! (2(\nu+k)-1)!} \\
& \quad \times \iint_{\xi_1, \xi_4 \geq 0} e^{-\frac{(\xi_1 + \xi_4)^2}{2}} (\xi_1 + \xi_4)^m \xi_1^{n-1} \xi_4^{2(\nu+k)-1} d\xi_1 d\xi_4.
\end{aligned}$$

Here the double integral equals

$$\frac{(n-1)! (2(\nu+k)-1)! (n+m+2(\nu+k)-2)!!}{(n+2(\nu+k)-1)!}.$$

So

$$I_n(m, \nu) = \sum_{k \geq 0} \frac{[(\nu-1+k)!]^2 (n+m+2(\nu+k)-2)!!}{[(\nu-1)!]^2 k! (n+2(\nu+k)-1)!}$$

Therefore

$$P_{\mathcal{C}_1}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_2}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq_b 2^\mu e^{-\log^2 n} \sum_{k \geq 0} s'(n, m, \nu, k),$$

$$s'(n, m, \nu, k) := \frac{[(\nu-1+k)!]^2 (n+m+2(\nu+k)-2)!! (n+2\nu)!}{(n+m+2\nu-2)! [(\nu-1)!]^2 k! (n+2(\nu+k)-1)!}.$$

The summand  $s'(n, m, \nu, k)$  is similar to the summand  $s(n, m, \nu, k)$  defined in (5.10). Closely following the derivation of the bound for  $\sum_{k \geq 0} s(n, \nu, k)$

in [14], we obtain

$$\sum_{k \geq 0} s'(n, m, \nu, k) \leq_b n^2 \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} n^{-\nu},$$

compare to (5.11). The last two bounds complete the proof.  $\square$

On  $\mathcal{C}_1 \supset \mathcal{C}_2$  we have

$$s = \sum_{i \in [n]} x'_i + \sum_{i \in A \cup B} y'_i = \sum_{i \in [n]} (x_i \vee y_i) \leq s_n = n^{1/2} + 6 \log n.$$

and on  $\mathcal{C}_2$

$$\max \left\{ \max_i \frac{x'_i}{s}, \max_{j \in A \cup B} \frac{y'_j}{s} \leq 1.01 \frac{\log^2 n}{n} \right\}.$$

Since  $m, \nu \leq n^{1/2} \log n$ , we have then the counterparts of the bounds in (4.14). Namely, on  $\mathcal{C}_2$ ,

$$(5.22) \quad \begin{aligned} \sum_{i \in \mathcal{N}} (x'_i + y'_i) &\leq_b s n^{-1} \log^3 n, & \sum_{i \in \mathcal{N}} (x'_i + y'_i)^2 &\leq_b n^{-1/2} \log^5 n, \\ \sum_{i \in \mathcal{N}} (x'_i + y'_i)^4 &\leq_b n^{-1} \log^8 n, \end{aligned}$$

(4) With  $\xi := \sum_{i \in \mathcal{M}} x'_i (= \sum_{i \in \mathcal{M}} x_i)$ , define  $v'_i = x'_i / \xi$  for  $i \in \mathcal{M}$ ,  $|\mathcal{M}| = n - m - 2\nu$ . Introduce  $T_2(\mathbf{v}') = \sum_{i \in \mathcal{M}} (v'_i)^2$ , and  $V'$  the set of all  $\mathbf{v}'$  such that

$$\left| \frac{n - m - 2\nu}{2} T_2(\mathbf{v}') - 1 \right| \leq n^{-\sigma}.$$

**Lemma 5.4.** *For  $\sigma < 1/3$ , let  $\mathcal{C}_3 = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{C}_2 : \mathbf{v}' \in V'\}$ . Then*

$$P_{\mathcal{C}_2}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_3}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq_b 2^\mu e^{-\Theta(n^{1/3-\sigma})} \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} n^{-\nu}.$$

*Proof.* Introduce

$$\xi_4 := \xi_1 - \xi = \sum_{i \in B^c \cap \mathcal{M}^c} x'_i + \sum_{i \in B} y'_i = \sum_{i \in \text{Odd}_{1,2}} x'_i + \sum_{i \in A} x'_i + \sum_{i \in B} y'_i.$$

Then with  $s := \xi + \xi_4 + \xi_2 + \xi_3$ ,

$$\begin{aligned} &P_{\mathcal{C}_2}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_3}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \\ &\leq_b 2^\mu \int \cdots \int_{\mathbf{x}', \mathbf{y}': \mathbf{v}' \in V'} e^{-\frac{s^2}{2} + \xi_2 \xi_3} \prod_{i \in M} dx'_i \prod_{j \in \text{Odd}_{1,2}} x_j dx_j \\ &\quad \times \prod_{k \in A} dx'_k \prod_{\ell \in B} dy'_\ell \prod_{b \in B} dx'_b \prod_{a \in A} dy'_a. \end{aligned}$$

Now the integrand depends on  $\{x'_i\}_{i \in \mathcal{M}}$  only through  $\xi = \sum_{i \in \mathcal{M}} x'_i$ . So, introducing the random intervals  $\mathcal{L}'_1, \dots, \mathcal{L}'_{n-m-2\nu}$  forming the partition of  $[0, 1]$ , we obtain

$$\begin{aligned} & P_{\mathcal{C}_2}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_3}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \\ & \leq_b 2^\mu \mathbb{P} \left( \left| \frac{n-m-2\nu}{2} T_2(\mathcal{L}') - 1 \right| \geq n^{-\sigma} \right) \iiint_{\xi, \xi_j \geq 0} e^{-\frac{s^2}{2} + \xi_2 \xi_3} \frac{\xi^{|\mathcal{M}|-1} d\xi}{(|\mathcal{M}|-1)!} \\ & \quad \times \frac{\xi_4^{2m+2\nu-1} d\xi_4}{(2m+2\nu-1)!} \cdot \frac{\xi_2^{\nu-1} d\xi_2}{(\nu-1)!} \cdot \frac{\xi_3^{\nu-1} d\xi_3}{(\nu-1)!}. \end{aligned}$$

The probability is of order  $e^{-\Theta(n^{1/3-\sigma})}$ , and the integral equals the bottom 3-dimensional integral in (5.9). Jointly with (5.10) and (5.11) this proves the claim.  $\square$

Finally, introduce  $T_3(\mathbf{v}') = \sum_{(i,j) \in \mathcal{M}_1 \cap \mathcal{M}_2} v'_i v'_j$ ; (here, of course,  $i, j \in \mathcal{M}$ ).

**Lemma 5.5.** *For  $\sigma < 1/3$ , let*

$$\mathcal{C}_4 = \{(\mathbf{x}, \mathbf{y}) \in \mathcal{C}_3 : |2(n-m-2\nu)T_3(\mathbf{v}') - 1| \leq n^{-\sigma}\};$$

*Then*

$$P_{\mathcal{C}_3}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_4}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq 2^\mu e^{-\Theta(n^{1/3-\sigma})} \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} n^{-\nu}.$$

The proof is a copy of the previous argument.

The Lemmas 5.2, 5.3, 5.4 and 5.5 imply

**Lemma 5.6.** *For every admissible pair  $\mathbf{\Pi}_1, \mathbf{\Pi}_2$ ,*

$$P(\mathbf{\Pi}_1, \mathbf{\Pi}_2) - P_{\mathcal{C}_4}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \leq_b e^{-\Theta(\log^2 n)} 2^\mu \left( \frac{e}{n+m} \right)^{\frac{n+m}{2}} n^{-\nu}.$$

*Here  $P_{\mathcal{C}_4}(\mathbf{\Pi}_1, \mathbf{\Pi}_2)$  is the integral of  $F(\mathbf{x}, \mathbf{y})$  over*

$$\mathcal{C}_4 \subset \{\mathbf{x}, \mathbf{y} \in [0, 1]^n : x_i = y_i \text{ if } \mathbf{\Pi}_1(i) = \mathbf{\Pi}_2(i)\},$$

*defined by the additional constraints: denoting  $\xi := \sum_{i \in \mathcal{M}} x'_i (= \sum_{i \in \mathcal{M}} x_i)$ ,*

$$(5.23) \quad s := \sum_{i \in [n]} x'_i + \sum_{j \in A \cup B} y'_j (= \xi_1 + \xi_2 + \xi_3) \leq s_n (= n^{1/2} + 6 \log n),$$

$$(5.24) \quad \max \left\{ \max_{i \in [n]} x'_i, \max_{j \in A \cup B} y'_j \right\} \leq 1.01 \frac{s \log^2 n}{n},$$

$$(5.25) \quad \left| \frac{|\mathcal{M}|}{2\xi^2} \sum_{i \in \mathcal{M}} x_i^2 - 1 \right| \leq n^{-\sigma}, \quad \left| \frac{2|\mathcal{M}|}{\xi^2} \sum_{(i,j) \in \mathcal{M}_1 \cap \mathcal{M}_2} x_i x_j - 1 \right| \leq n^{-\sigma}.$$

The constraint (5.25) involves only  $\{x_i\}_{i \in \mathcal{M}}$ , and the constraint (5.24) imposes the bound for the individual components  $x'_i$  and  $y'_j$ . Since  $s_n \leq 2n^{1/2}$ , the latter implies that

$$(5.26) \quad \max \left\{ \max_{i \in [n]} x'_i, \max_{j \in A \cup B} y'_j \right\} \leq 3n^{-1/2} \log^2 n,$$

obviating the constraint  $x'_i \leq 1, y'_j \leq 1$ . On  $\mathcal{C}_4$  the inequality (5.21) can be drastically simplified. First of all, the bottom part of the bound (5.21) is

$$-\frac{1}{4} \left( \sum_{i \in \mathcal{M}} x_i^2 \right)^2 + O(n^{-1/2} \log^5 n).$$

Second,

$$\begin{aligned} \sum_{i \in \mathcal{M}} x_i^2 &= (1 + O(n^{-\sigma})) \frac{2\xi^2}{|\mathcal{M}|} = \frac{2\xi^2}{|\mathcal{M}|} + O(n^{-\sigma}), \\ \sum_{(i,j) \in M_1 \cap M_2} x_i x_j &= (1 + O(n^{-\sigma})) \frac{\xi^2}{2|\mathcal{M}|} = \frac{\xi^2}{2|\mathcal{M}|} + O(n^{-\sigma}), \end{aligned}$$

and  $\xi = \xi_1(1 + O(n^{-1} \log^2 n))$ . In addition,  $|\mathcal{M}| = n(1 + O(n^{-1/2} \log n))$ . Therefore (5.21) becomes

$$G(\mathbf{x}, \mathbf{y}) \leq (1 + O(n^{-\sigma})) \exp[H(\boldsymbol{\xi})], \quad H(\boldsymbol{\xi}) = -\frac{s^2}{2} + \xi_2 \xi_3 + \frac{3\xi_1^2}{2n} - \frac{\xi_1^4}{n^2}.$$

**Lemma 5.7.**

$$\begin{aligned} P_{\mathcal{C}_4}(\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2) &\leq \frac{2^\mu (1 + O(n^{-\sigma}))}{(n+m-1)! [(\nu-1)!]^2} \cdot \mathcal{I}(n+m, \nu), \\ \mathcal{I}(n+m, \nu) &:= \iiint_{(\xi_1, \xi_2, \xi_3) \in \mathbf{R}} \exp[H(\boldsymbol{\xi})] \xi_1^{n+m-1} \cdot (\xi_2 \xi_3)^{\nu-1} d\boldsymbol{\xi}, \\ \mathbf{R} &:= \{ \boldsymbol{\xi} \geq \mathbf{0} : \xi_1 \leq n^{1/2} + 6 \log n; \xi_2, \xi_3 \leq 2 \log^3 n \}. \end{aligned}$$

The proof, of course, is based on the description of  $\mathcal{C}_4$ , and it runs along the familiar lines of our preceding proofs; in particular, see the proof of Lemma 5.4. We omit the details. Furthermore, by the asymptotic formula for  $\mathcal{I}(n, \nu)$  from [14] (3.60), we have

$$\mathcal{I}(n+m, \nu) = (1 + o(1)) \left( \frac{\pi e}{n+m} \right)^{1/2} \left( \frac{n+m}{e} \right)^{\frac{n+m}{2}} (n+m)^{-\nu} [(\nu-1)!]^2.$$

So, by Lemma 5.7,

$$P_{\mathcal{C}_4}(\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2) \leq \frac{2^\mu (1 + O(n^{-\sigma}))}{(n+m-1)!} \left( \frac{\pi e}{n+m} \right)^{1/2} \left( \frac{n+m}{e} \right)^{\frac{n+m}{2}} (n+m)^{-\nu}.$$

Therefore, within the factor  $1 + O(n^{-\sigma})$ ,

$$\begin{aligned}
& \sum_{\mathbf{\Pi}_1, \mathbf{\Pi}_2} P_{\mathcal{C}_4}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) \\
& \leq \sum_{m \leq m_n} \binom{n}{m} \frac{f(m)}{(n+m-1)!} \left( \frac{\pi e}{n+m} \right)^{1/2} \left( \frac{n+m}{e} \right)^{\frac{n+m}{2}} \\
& \quad \times \sum_{\nu \leq m_n} \binom{n-m}{2\nu} (n-m-2\nu-1)!! (n+m)^{-\nu} \cdot \sum_{\mu} 2^{2\mu} f(2\nu, \mu) \\
& = n! \sum_{m \leq m_n} \frac{f(m)}{m! (n+m-1)!} \left( \frac{\pi e}{n+m} \right)^{1/2} \left( \frac{n+m}{e} \right)^{\frac{n+m}{2}} \\
& \quad \times \sum_{\nu \leq m_n} \frac{(n-m-2\nu-1)!!}{(n-m-2\nu)! (2\nu)!} (n+m)^{-\nu} \cdot \sum_{\mu} 2^{2\mu} f(2\nu, \mu).
\end{aligned}$$

cf. (5.13). The sum over  $\mu$  is  $e^{-1}(2\nu)!(1 + O(1/\nu))$ , see (5.14). So the sum over  $\nu$  is asymptotic to

$$\frac{e^{-1}}{(n-m)!!} \sum_{\nu \leq m_n} (n+m)^{-\nu} \prod_{j=0}^{\nu-1} (n-m-2j) \sim \frac{e^{-1}}{(n-m)!!} \sum_{\nu \leq m_n} e^{-\frac{\nu^2}{n} - \frac{2\nu m}{n}}.$$

Thus, since  $f(m)/m! = e^{-1} \sqrt{\frac{2}{\pi m}} (1 + O(m^{-1}))$ , the  $m$ -term in the resulting sum is (within a factor  $1 + O(m^{-1})$ )

$$\begin{aligned}
& e^{-2} \sqrt{\frac{2}{\pi m}} \cdot \frac{n!}{(n-m)!! (n+m-1)!} \left( \frac{e}{n+m} \right)^{1/2} \left( \frac{n+m}{e} \right)^{\frac{n+m}{2}} \\
& \quad \times \sum_{\nu \leq m_n} e^{-\frac{\nu^2}{n} - \frac{2\nu m}{n}} \\
& \sim e^{-3/2} \sqrt{\frac{2}{\pi^2 m}} \cdot e^{-\frac{m^2}{2n}} \sum_{\nu \leq m_n} e^{-\frac{\nu^2}{n} - \frac{2\nu m}{n}}.
\end{aligned}$$

So

$$\begin{aligned}
\sum_{\mathbf{\Pi}_1, \mathbf{\Pi}_2} P_{\mathcal{C}_4}(\mathbf{\Pi}_1, \mathbf{\Pi}_2) & \lesssim e^{-3/2} \sqrt{\frac{2}{\pi^2}} \sum_{m, \nu \leq m_n} \frac{1 + O(m^{-1})}{m^{1/2}} e^{-\frac{m^2}{2n} - \frac{2\nu m}{n} - \frac{\nu^2}{n}} \sim cn^{3/4}, \\
c & := e^{-3/2} \sqrt{\frac{2}{\pi^2}} \iint_{x, y \geq 0} x^{-1/2} e^{-\frac{x^2}{2} - 2xy - y^2} dx dy \approx 0.617.
\end{aligned}$$

Thus, since  $E[\mathcal{S}_n]$  is of order  $n^{1/4}$ , we have

**Theorem 5.8.**  $E[\mathcal{S}_n^2] \lesssim cn^{3/4}$ .

With extra work, we could have proved that  $E[\mathcal{S}_n^2] \gtrsim cn^{3/4}$ , as well. Since  $E[\mathcal{S}_n^2] \gg E^2[\mathcal{S}_n]$ , we cannot deduce that  $\mathcal{S}_n \rightarrow \infty$  in probability, even though  $E[\mathcal{S}_n] \rightarrow \infty$ . We firmly believe that the argument itself may help to define a subset of stable partitions for which the two-moments approach will work just fine. For now we are content to use the techniques above to prove a result that would have been out of reach if not for the analysis of  $E[\mathcal{S}_n^2]$ .

**Theorem 5.9.** *Let  $q_n$  denote the fraction of members that have more than one stable predecessor. Then  $E[q_n] \lesssim 2ecn^{-1/4}$ , so that with high probability almost all members have a unique stable predecessor.*

*Proof.* It suffices to consider the members outside the odd cycles. If any such member has some two stable partners, it belongs to a cycle of even length  $\geq 4$  formed by the alternating pairs matched in the corresponding stable partitions  $\Pi_1$  and  $\Pi_2$ . Notice that selecting every other edge of those cycles, we get a stable partition. Therefore, without loss of generality we can assume that  $\Pi_1$  and  $\Pi_2$  form a unique cycle of even length  $2\nu \geq 4$ . It follows that  $Q_n$ , the total number of members with at least two stable partners, is below the total length of the single cycles formed by these special pairs of stable partitions  $\Pi_1$  and  $\Pi_2$ . The bound does look crude, but it works.

To bound the total expected length of those cycles, we need to estimate  $\sum 2\nu(\Pi_1, \Pi_2)P_{\mathcal{C}_4}(\Pi_1, \Pi_2)$ . For those pairs we have  $\mu := \mu(\Pi_1, \Pi_2) = 1$ , and  $\sum_{\mu} 2^{2\mu} f(2\nu, \mu) = 2(2\nu - 1)!$ . Therefore

$$E[Q_n] \lesssim \sum_{\Pi_1, \Pi_2} 2\nu(\Pi_1, \Pi_2)P_{\mathcal{C}_4}(\Pi_1, \Pi_2) \lesssim 2ecn^{3/4}.$$

□

**Acknowledgment.** Almost thirty years ago Don Knuth introduced me to his ground-breaking work on random stable marriages. Don's ideas and techniques have been a source of inspiration for me ever since. The masterful book by Dan Gusfield and Rob Irving encouraged me to continue working on stable matchings. I am very grateful to Rob for the chance to work with him on the stable roommates problem back in 1994, and for his encouragement these last months. Itai Ashlagi, Peter Biró, Jennifer Chayes, Gil Kalai, Yash Kanoria, Jacob Leshno and the recent monograph by David Manlove made me aware of a significant progress in theory and applications of two/one-sided stable matchings. I thank the organizers for a valuable opportunity to participate in MATCH-UP 2017 Conference.

## REFERENCES

- [1] D. J. Abraham, P. Biró and D. F. Manlove, “Almost stable” matchings in the roommates problem, Proceedings of WAOA '05: the 3rd Workshop on Approximation and Online Algorithms, Lecture Notes in Computer Science, **3879** (2006) 1–14.
- [2] J. Alcalde, *Exchange-proofness or divorce proofness? Stability in one-sided matching markets*, Economic Design **1** (1995) 275–287.
- [3] G. E. Andrews, R. Askey and R. Roy, *Special Functions*, Cambridge University Press (1999).
- [4] K. Cechlárová and D. F. Manlove, *The exchange-stable marriage problem*, Disc. Appl. Math. **152** (2005) 109–122.
- [5] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press (2009).
- [6] D. Gusfield and R. W. Irving, *The Stable Marriage Problem, Structure and Algorithms*, The MIT Press (1989).
- [7] R. W. Irving and B. Pittel, *An upper bound for the solvability probability of a random stable roommates instance*, **5** (1994) 465–486.
- [8] D. E. Knuth, *Stable marriage and its relation to other combinatorial problems: an introduction to the mathematical analysis of algorithms*, CRM Proceedings and Lecture notes (1996).
- [9] D. E. Knuth, R. Motwani and B. Pittel, *Stable husbands*, Random Struct Algorithms **1** (1991) 1–14.
- [10] D. F. Manlove, *Algorithmics of Matching under Preferences*, World Scientific (2013).
- [11] S. Mertens, *Random stable matchings*, J. Statist. Mechanics: Theory and Experiment **10** (2005).
- [12] B. Pittel, *The average number of stable matchings*, SIAM J Disc Math **2** (1989) 530–549.
- [13] B. Pittel, *On a random instance of a “stable roommates” problem: Likely behavior of the proposal algorithm*, Comb Probab Comput **2** (1993) 53–92.
- [14] B. Pittel, *The “stable roommates” problem with random preferences*, Ann Probab **21** (1993) 1441–1477.
- [15] J. J. M. Tan, *A necessary and sufficient condition for the existence of a complete stable matching*, J. Algorithms **12** (1991) 1–25.
- [16] J. J. M. Tan, *Stable matchings and stable partitions*, International J. Computer Math. **39** (1991) 11–20.

DEPARTMENT OF MATHEMATICS, THE OHIO STATE UNIVERSITY, COLUMBUS, OHIO 43210, USA

*E-mail address:* `bgp@math.ohio-state.edu`