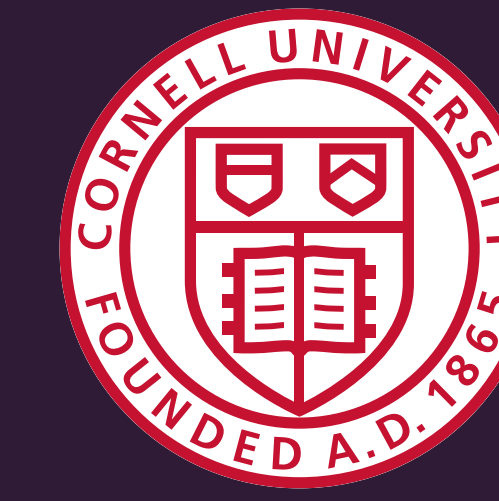


CPII: An Improved Parallel Stable Matching Algorithm

Will Bradley¹, Jeffrey Xu², Dr. Enyue Lu³

¹Cornell University, ²Rutgers University, ³Salisbury University



Background

Consider n men and n women. Each man gives a unique preference rank from 1 to n to each of the women, and likewise for the women to the men. The **stable matching problem** concerns the finding of a bijection between the women and men so that no man and woman form an **unstable pair**, i.e. they would prefer to be with each other over their assigned partners. In 1962, Gale and Shapley showed that a stable matching always exists and presented an algorithm to do so in $O(n^2)$ worst-case time [2]. Further results due to Knuth showed that the average-case time is $O(n \log n)$ [3].

It is a natural extension to try to parallelize the problem, but naïve parallelization of the **Gale-Shapley algorithm** (GS) is inefficient— $O(n^2 \log n)$ worst-case time with n processors [1]. Lu’s **Parallel Iterative Improvement algorithm** (PII) empirically improves efficiency by using n^2 processors but fails to terminate around 14% of the time for large n [5]. There have been multiple improvements made to PII, including the **Convergent PII algorithm** (CPII) due to Clark, which is simpler and provably terminates in $O(n^2)$ iterations [1]. However, Clark’s original CPII paper leaves the *average*-case runtime as an open problem. Our research empirically determines and then proves that CPII takes $O(n/\log n)$ iterations on average, which under a completely-connected network topology implies an average runtime of $O(n)$.

Throughout our work, we assume all preference orders are uniformly and independently distributed.

CPII

CPII solves the stable matching problem by beginning with an empty **partial matching** $K_1 = \emptyset$ and iteratively adding man-woman pairs that are (in a manner more precisely defined below) most unstable, until a stable matching is reached. Specifically, in iteration i , CPII calculates the following sets of pairs, taking the convention that every person prefers having a partner to being single:

- A_i , the set of all unstable pairs under the current matching
- B_i , the set of each man’s most preferred unstable pair, if such a pair exists
- C_i , the set of each woman’s most preferred pair in B_i , if such a pair exists
- R_i , the set of all pairs in K_i that share a member with a pair in C_i

Then, CPII constructs a new partial matching K_{i+1} that includes all pairs in C_i , as described by the below pseudocode:

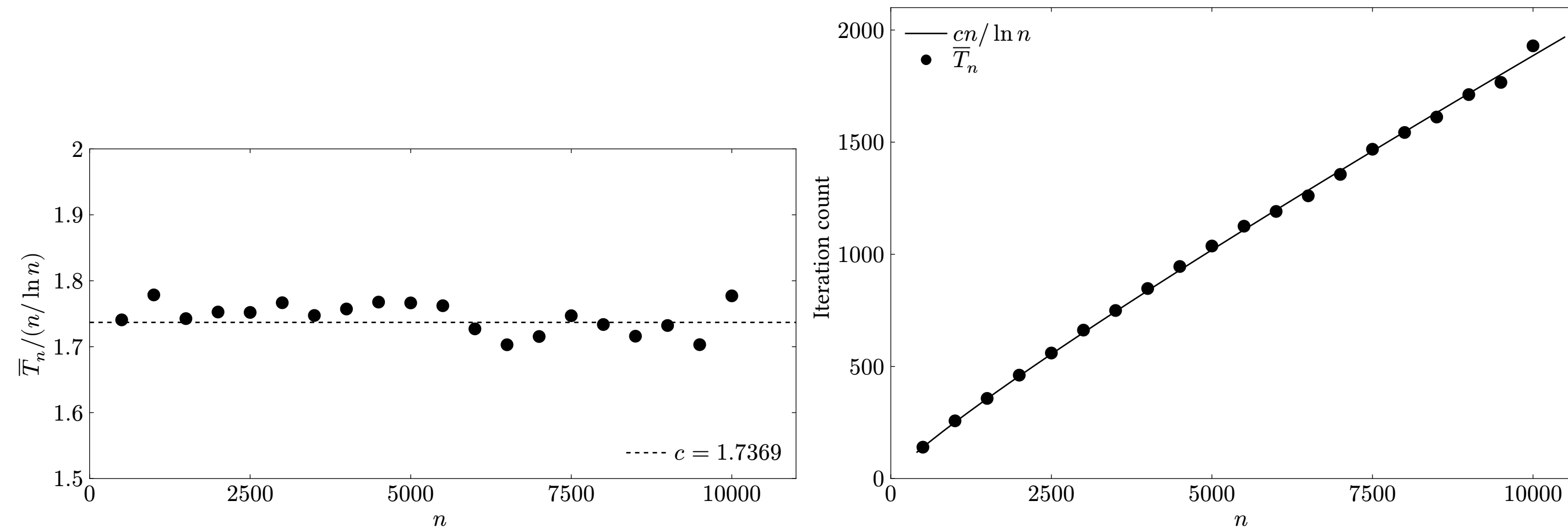
```
CPII():
1  $i \leftarrow 1$ ;  $K_1 \leftarrow \emptyset$ 
2 loop:
3   calculate  $A_i, B_i, C_i, R_i$  from  $K_i$ 
4   if  $A_i = \emptyset$ :
5     return  $K_i$ 
6   else:
7      $K_{i+1} \leftarrow (K_i \setminus R_i) \cup C_i$ 
8    $i \leftarrow i + 1$ 
```

Since the algorithm terminates exactly when there are no unstable pairs left and $A_i = \emptyset$ implies $|K_i| = n$, CPII returns a (full) stable matching. Clark additionally showed that CPII always terminates (specifically, within n^2 iterations), so it is correct [1].

Empirical Results

Clark showed empirically that CPII often took substantially fewer iterations to converge than Gale-Shapley and a naïve parallel Gale-Shapley implementation [1]. Our tests suggested that CPII takes an average of $O(n/\ln n)$ iterations to converge and therefore has an average running time of $O(n)$.

Plots showing the mean number of iterations for different problem sizes, from intervals of 500 up to 10000, are shown below.



Theoretical Results

We show the following results related to CPII convergence:

Theorem 1

$$\mathbb{E}[|B_{i+1}|] / \mathbb{E}[|B_i|] \leq \mathbb{E}_{|A_i|>0, |B_i|>0} \left[(1 - |B_i|/|A_i|)^{|B_i|} \right].$$

Lemma 2

For any constant α and $b \in O((\ln n)^2)$ such that $\mathbb{P}[|B_i| = b]$ is $\Omega(n^{-\alpha})$, then $\mathbb{E}[|A_i| \mid |B_i| = b] \leq b^2 + O(bn/\ln n)$, uniformly in i .

It turns out that $\mathbb{E}[|A_i| \mid |B_i| = b]$ is closely related to the sum of the ranks women assign to their partners under the partial matching K_i . Therefore, we use the following extension of a result due to Pittel [4] in order to prove Lemma 2.

Lemma 3

For size $n \times n$ preference matrices M, W whose rows are permutations of $[n]$, let $\mathcal{R}(M, W)$ be the maximum sum of women’s ranks of their partners over all stable matchings.

If the rows of M and W are uniformly and independently distributed random preferences for each man and woman, then

$$\mathbb{P}_{M,W}[\mathcal{R}(M, W) > C_\rho(n)] \leq O(n^{-\delta}),$$

for all positive $\rho \in O(\ln \ln n)$ and $\delta \in (0, \kappa e^\rho]$, where $\kappa > 0$ is some absolute constant and $C_\rho(n) = (n^2/\ln n)(1 + (\ln \ln n + \rho)/(\ln n))$.

That is, we have a probabilistic bound on how much the maximum sum of women’s ranks—and therefore of *any* such sum—exceeds $n^2/\ln n$.

We use Lemma 2 to obtain the following recurrent bound on $\mathbb{E}[|B_i|]$:

Lemma 4

For any size n and iteration i , we have that

$$\mathbb{E}[|B_{i+1}|] \leq \mathbb{E}[|B_i|] - \left(1 - \frac{1}{e}\right) \frac{\ln n}{c_0 n} \left(\mathbb{E}[|B_i|] - ne^{-(i-1)(\ln n)^2/n} - \frac{1}{n} \right)_+^2 + \frac{(\ln n)^4}{n^3}.$$

where c_0 is some absolute constant and the subscript $+$ operator is defined by $y_+ := \max(0, y)$.

Theoretical Results

Now it is straightforward to get the following result:

Theorem 5

There exists some index $i(n) \in O(n/\ln n)$ such that $\mathbb{E}[|B_{i(n)}|]$ is $O(1)$.

We can use Lemma 2 once again to show that unless we are already conditioning on a rare event, $\mathbb{E}[|B_i|]$ decreases quickly enough to prove the desired runtime:

Theorem 6

For constant $\alpha \geq 0$ and i such that $\mathbb{E}[|B_i|] \geq n^{-\alpha}$, $1 - \mathbb{E}[|B_{i+1}|] / \mathbb{E}[|B_i|]$ is $\Omega(\ln n/n)$, uniformly in i .

Theorem 7

Let T_n denote the number of iterations taken by CPII in a stable matching problem of size n . Then,

$$\mathbb{E}[T_n] \in O(n/\ln n).$$

To prove Theorem 7, write $\mathbb{E}[T_n] = \sum_{i=1}^{n^2} \mathbb{P}[B_i \neq \emptyset]$. By Theorem 5, we may split the sum at some index $i_0 \in O(n/\ln n)$ such that $\mathbb{E}[|B_{i_0}|] \in O(1)$. The part of the sum corresponding to early iterations $i < i_0$ is $O(n/\ln n)$, as each probability is at most 1. For iterations $i \geq i_0$, we use Markov’s inequality to bound each summand by $\mathbb{P}[|B_i| \neq \emptyset] = \mathbb{P}[|B_i| \geq 1] \leq \mathbb{E}[|B_i|]$; then by Theorem 6, the total is bounded by a geometric sequence with first term $O(1)$ and common ratio $1 - \Omega(\ln n/n)$. Summing both parts gives the desired result.

The above image provides a visualization of this general proof strategy.

Conclusions and Future Work

Our results show that under independent and uniform random preferences, CPII converges in an average of $O(n/\ln n)$ iterations and therefore has an average-case runtime of $O(n)$, a significant improvement on existing algorithms such as Gale-Shapley. Empirical testing of CPII has been shown to corroborate these results.

One potential future direction is to derive full distributions for the number of iterations in a CPII problem instance, or the number of Gale-Shapley proposals carried out by a single CPII iteration. We could additionally look at which CPII can also be made more practical by investigating whether the assumption of uniform preferences can be weakened, or by implementing CPII on simpler network topologies.

References

- [1] Caleb Clark and Enyue Lu. 2018. Improved Parallel Stable Matching. (2018).
- [2] D. Gale and L. S. Shapley. 1962. College Admissions and the Stability of Marriage. *The American Mathematical Monthly* 69, 1 (1962), 9–15. Retrieved from <http://www.jstor.org/stable/2312726>
- [3] Donald E. Knuth. 1997. *Stable Marriage and Its Relation to Other Combinatorial Problems*. American Mathematical Society.
- [4] Boris Pittel. 1989. The Average Number of Stable Matchings. *SIAM Journal on Discrete Mathematics* 2, 4 (November 1989), 530–549. <https://doi.org/10.1137/0402048>
- [5] Colin White and Enyue Lu. 2013. An Improved Parallel Iterative Algorithm for Stable Matching. 2013. SuperComputing.

Acknowledgements

This project was generously funded by NSF 2447041 under the Research Experiences for Undergraduates program. We thank the REU principal investigator and inventor of PII Dr. Enyue Lu, who was additionally our research mentor, as well as Caleb Clark for his helpful correspondence. Lastly, we thank our friends and colleagues over the summer: Alea Atkinson, Naisha Bhandari, Matthew Goldman, Caroline Griem, Ella Kennedy, Juniper Pasternak, JJ Salter, and Joyce Zheng.