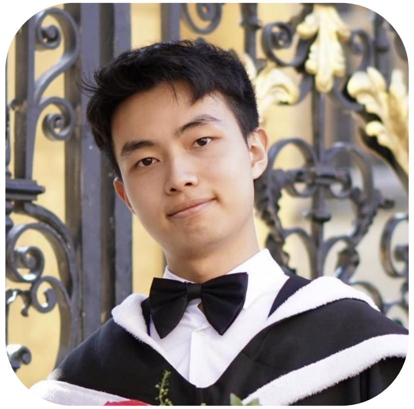
A One-Size-Fits-All Approach to Improving Randomness in Paper Assignment

Yixuan (Even) Xu
Tsinghua

Steven Jecmen
Carnegie Mellon



Zimeng Song
Independent

Fei Fang
Carnegie Mellon

Paper Assignment: Setting

- **Basic setting**
  - n_p **papers**, n_r **reviewers**
  - **Assignment:** $\boldsymbol{x} \in \{0, 1\}^{n_p \times n_r}$
- **Feasibility**
  - Each **paper** gets **exactly** ℓ_p **reviewers**
  - Each **reviewer** gets **at most** ℓ_r **papers**
- **Quality:** A **similarity matrix** $\boldsymbol{S} \in [0, 1]^{n_p \times n_r}$

Paper Assignment (Deterministic)

- Paper assignment as an **integer linear program**

Maximize	**Quality**$(\boldsymbol{x}) = \sum_{p,r} x_{p,r} \cdot S_{p,r}$	(maximize **Quality**)
Subject to	$\sum_{r} x_{p,r} = \ell_p, \forall p$	(**paper** requirement)
	$\sum_{p} x_{p,r} \leq \ell_r, \forall r$	(**reviewer** load)
	$\boldsymbol{x} \in \{0, 1\}^{n_p \times n_r}$	

- Produces the Maximum-**Quality** Assignment

Why Randomness

- **Robustness to malicious behavior**
- **Evaluation of alternative assignments**
- **Reviewer diversity**
- **Reviewer anonymity**

Paper Assignment (Randomized)

- Paper assignment as a **continuous linear program**

Maximize	**Quality**$(\boldsymbol{x}) = \sum_{p,r} x_{p,r} \cdot S_{p,r}$	(maximize **Quality**)
Subject to	$\sum_r x_{p,r} = \ell_p, \forall p$	(**paper** requirement)
	$\sum_p x_{p,r} \leq \ell_r, \forall r$	(**reviewer** load)
	$\boldsymbol{x} \in [\mathbf{0}, \mathbf{1}]^{n_p \times n_r}$	

- Now $x_{p,r}$ denotes the **marginal probability** of assignment
  - Shown by prior work that a **randomized assignment** can be converted into a **distribution of deterministic assignments**

PLRA (Jecmen et al. 2020, Deployed)

- **Probability Limited Randomized Assignment (PLRA)**

Maximize $\text{Quality}(\boldsymbol{x}) = \sum_{p,r} x_{p,r} \cdot S_{p,r}$

Subject to $\sum_r x_{p,r} = \ell_p, \forall p$

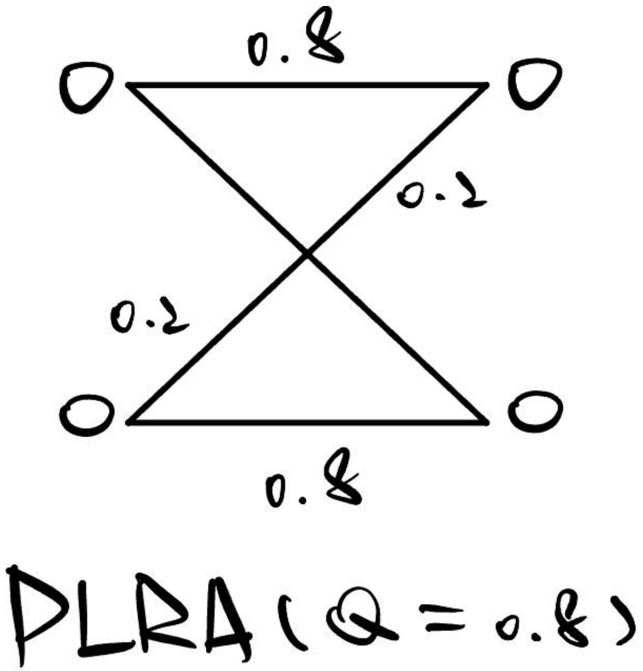
$\sum_p x_{p,r} \leq \ell_r, \forall r$

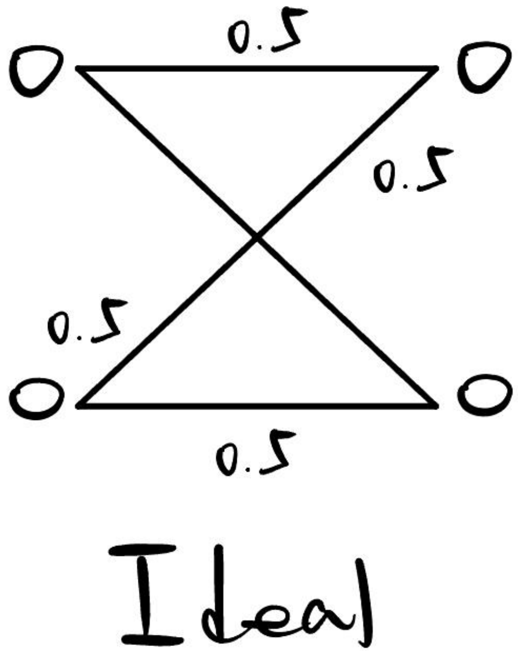
$\boldsymbol{x} \in [0, \boldsymbol{Q}]^{n_p \times n_r}$

- **Hyperparameter Q:**
  - **Guarantees** each **paper-reviewer** pair is matched w.p. $\leq \boldsymbol{Q}$
  - Mainly concerned with **robustness to malicious behavior**

A Problem with PLRA

- The **randomness** of its assignment depends on **Q**
  - Not easy to set, and sometimes **suboptimal with any Q**





Metrics for Randomness

- **Maximum Probability**
  - **Maxprob**$(\boldsymbol{x}) = \max_{p,r}\{x_{p,r}\}$
  - Already used by PLRA as a constraint

- **Our proposed metrics**
  - **Average maximum probability: AvgMax**$(\boldsymbol{x}) = \frac{1}{n_p}\sum_p \max_r\{x_{p,r}\}$
  - **Support size: Support**$(\boldsymbol{x}) = \sum_{p,r} \mathbf{1}[x_{p,r} > 0]$
  - **Entropy: Entropy**$(\boldsymbol{x}) = \sum_{p,r} x_{p,r} \cdot \log(1 / x_{p,r})$
  - **L2 norm: L2Norm**$(\boldsymbol{x}) = \sqrt{\sum_{p,r} x_{p,r}{}^2}$

Perturbed Maximization (PM)

- **Perturbed Maximization (PM)**

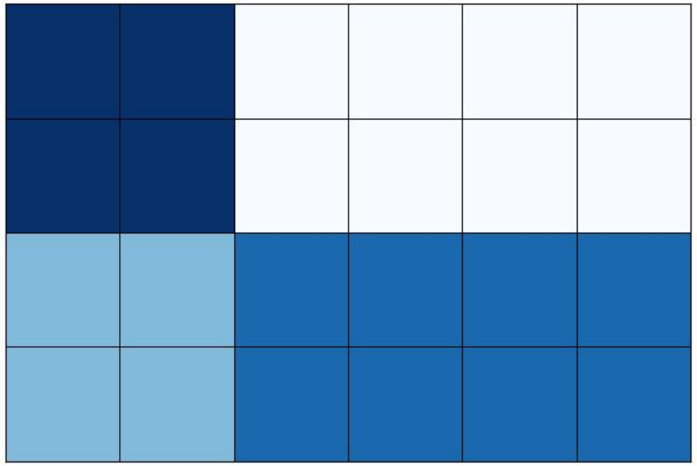
Maximize	**PQuality**$(\boldsymbol{x}) = \sum_{p,r} \boldsymbol{f}(x_{p,r}) \cdot S_{p,r}$
Subject to	$\sum_r x_{p,r} = \ell_p, \forall p$
	$\sum_p x_{p,r} \leq \ell_r, \forall r$
	$\boldsymbol{x} \in [0, Q]^{n_p \times n_r}$

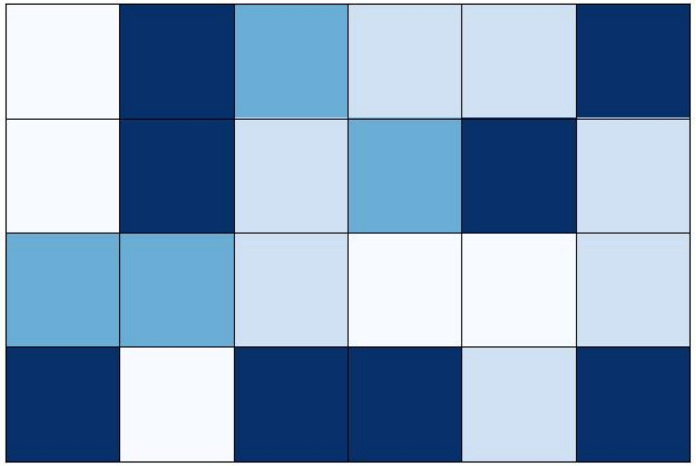
- **Perturbation Function $\boldsymbol{f}(\cdot)$:**
  - A **non-decreasing**, **concave** function from $[0,1] \to [0,1]$
  - Intuition: the higher $x_{p,r}$, the lower gain in **PQuality**

Theoretical Analysis

- **(Informal)** With the same probability limit $\boldsymbol{Q}$ and a **strictly concave** perturbation function $\boldsymbol{f}(\cdot)$, PM outperforms PLRA under **any** of the proposed randomness measures **without loss** in solution **Quality** if the similarity matrix $\boldsymbol{S}$ is:

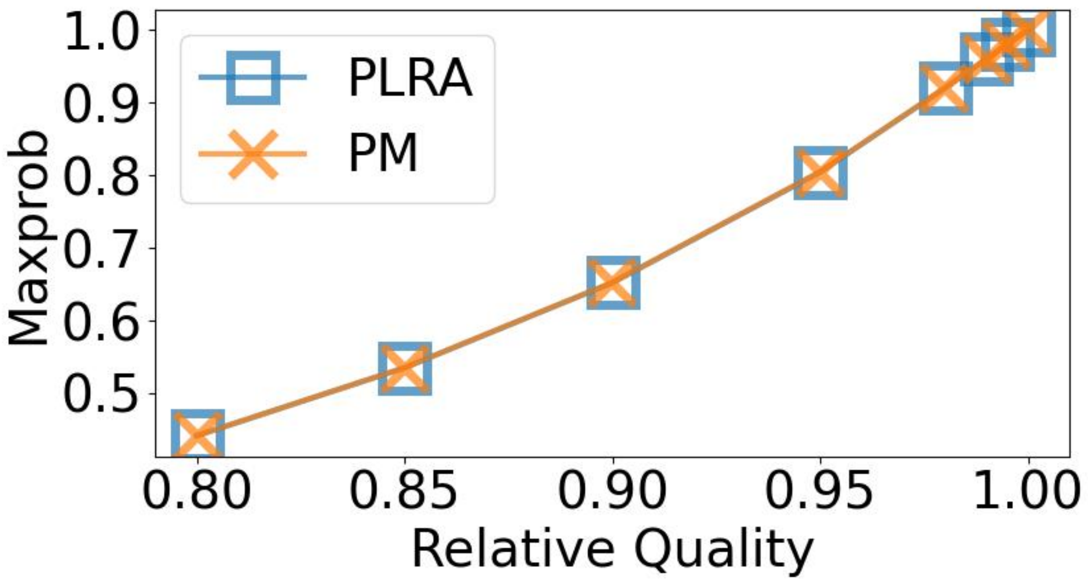
Blockwise Dominant or **Discrete & Random**

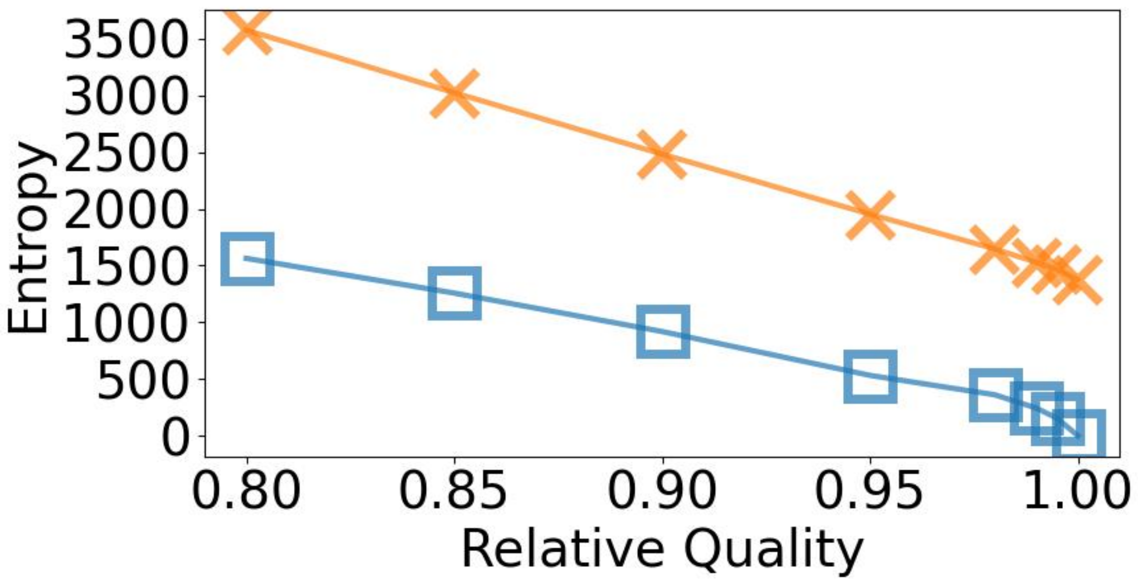




Experiments

- On the bidding data of **AAMAS2015**, PM has exactly the same performance on **Maxprob** with PLRA (where PLRA is **optimal**), and outperforms PLRA on **all** other randomness measures





Our Contributions

- **We define new metrics to measure randomness of randomized paper assignments in peer review**

- **We propose Perturbed Maximization (PM)**
  - Theoretically, PM outperforms prior work on structured matrices
  - Experimentally, PM outperforms prior work on real-world datasets
  - We also study the faster computation of PM (details in the paper)

Acknowledgements: *This work was supported by ONR grant N000142212181, NSF grant IIS-2200410 and partly by Sloan Research Fellowship.*