Aggregating Partial Rankings with Applications to Peer Grading in Massive Online Open Courses

Ioannis Caragiannis, George A. Krimpas, Alexandros A. Voudouris

The 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015) 675–683.

Speaker: Joseph Chuang-Chieh Lin

Institute of Information Science
Academia Sinica
Taiwan

29 April 2016

Authors

Ioannis Caragiannis

George A. Krimpas

Alexandros A. Voudouris

Outline

MOOCs

- MOOCs: Massive Online Open Courses
  - e.g., Coursera, EdX.
- Outscourcing the grading task to the students.
- They may have incentives to assign LOW scores to everybody else.
  - ▷ Ask each student to grade a SMALL number of her peers' assignments.
  - Then merge individual rankings into a global one (like social choice).

MOOCs

- MOOCs: Massive Online Open Courses
  - e.g., Coursera, EdX.
- Outscourcing the grading task to the students.
- They may have incentives to assign LOW scores to everybody else.
  - ▷ Ask each student to grade a SMALL number of her peers' assignments.
  - Then merge individual rankings into a global one (like social choice).

MOOCs

- MOOCs: Massive Online Open Courses
  - e.g., Coursera, EdX.
- Outscourcing the grading task to the students.
- They may have incentives to assign LOW scores to everybody else.
  - ▷ Ask each student to grade a SMALL number of her peers' assignments.
  - Then merge individual rankings into a global one (like social choice).

MOOCs

- MOOCs: Massive Online Open Courses
  - e.g., Coursera, EdX.
- Outscourcing the grading task to the students.
- They may have incentives to assign LOW scores to everybody else.
  - ▷ Ask each student to grade a SMALL number of her peers' assignments.
  - Then merge individual rankings into a global one (like social choice).

MOOCs

- MOOCs: Massive Online Open Courses
  - e.g., Coursera, EdX.
- Outscourcing the grading task to the students.
- They may have incentives to assign LOW scores to everybody else.
  - ▷ Ask each student to grade a SMALL number of her peers' assignments.
  - Then merge individual rankings into a global one (like social choice).

Terminologies

- $\mathcal{A}$: universe of n elements (students).
- (n, k)-grading scheme:
  a collection $\mathcal{B}$ of size-k subsets (bundles) of $\mathcal{A}$, such that each element of $\mathcal{A}$ belongs to exactly k subsets of $\mathcal{B}$.
- The bundle graph:
  Represent the (n, k)-grading scheme with a bipartite graph.
- $\prec_b$: a ranking of the element b contains (partial order).

The aggregation rule

An aggregation rule:
profile of partial rankings $\mapsto$ complete ranking of all elements.

- Borda:

SPRING FEAST 2016 BALLOT 春酒地點調查

a	金色三麥	5
b	晶湯匙	3
c	北雲餐廳	1
d	西堤牛排	2
e	卡布里喬莎	4

SPRING FEAST 2016 BALLOT 春酒地點調查

a	金色三麥	5
b	晶湯匙	4
c	北雲餐廳	2
d	西堤牛排	1
e	卡布里喬莎	3

SPRING FEAST 2016 BALLOT 春酒地點調查

a	金色三麥	4
b	晶湯匙	5
c	北雲餐廳	1
d	西堤牛排	3
e	卡布里喬莎	2

- a: 14; b: 12; c: 4; d: 6; e: 9.

$$a \prec b \prec e \prec d \prec c.$$

Order-revealing grading scheme

An aggregation rule in peer grading (Borda):

Final Exam Grading

	Curry	2
	Bob	3
	Elvis	1

↓ Alice

Final Exam Grading

	Alice	3
	Curry	2
	David	1

↓ Bob

Final Exam Grading

	Elvis	1
	David	2
	Bob	3

↓ Curry

Final Exam Grading

	Alice	3
	Curry	1
	Bob	2

↓ David

Final Exam Grading

	Alice	3
	Elvis	1
	David	2

↓ Elvis

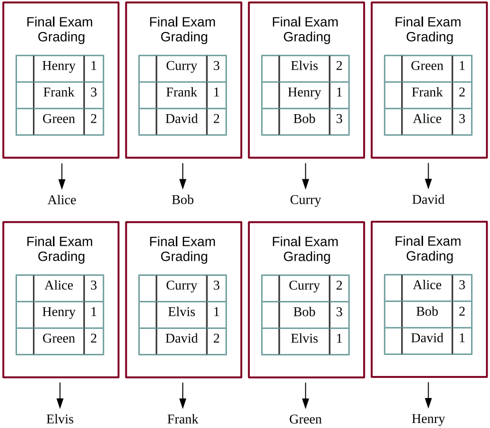
- Alice: 9; Bob: 8; Curry: 5; David: 5; Elvis: 3.

  Alice $\prec$ Bob $\prec$ Curry $\prec$ David $\prec$ Elvis.

Assumption (perfect grading)

Each student grades the assignments in her bundle consistently to the ground truth.

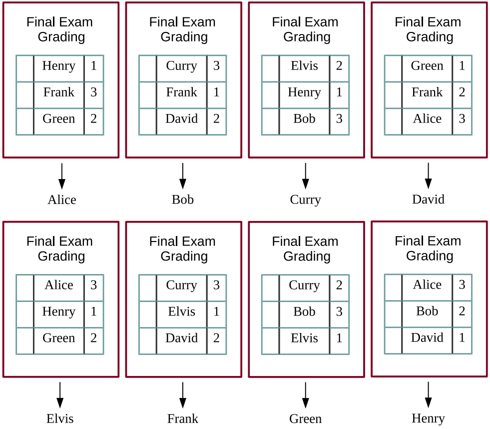
Order-revealing grading scheme (contd.)



- Alice: 9; Bob: 8; Curry: 8; David: 5; Elvis: 4; Frank: 6; Green: 5; Henry: 3.

  Alice $\prec$ Bob $\prec$ Curry $\prec$ Frank $\prec$ David $\prec$ Green $\prec$ Elvis $\prec$ Henry.

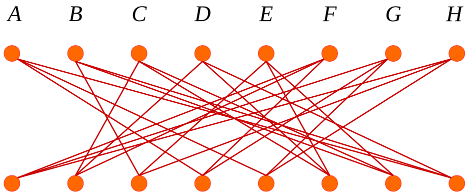
Order-revealing grading scheme (contd.)



- Alice: 9; Bob: 8; Curry: 8; David: 5; Elvis: 4; Frank: 6; Green: 5; Henry: 3.
  Alice $\prec$ Bob $\prec$ Curry $\prec$ Frank $\prec$ David $\prec$ Green $\prec$ Elvis $\prec$ Henry.

The bundle graph

The bundle graph:



- A random k-regular graph:

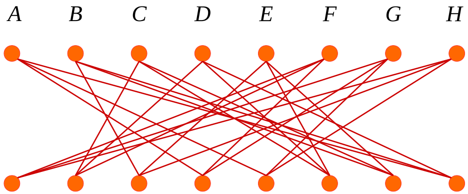
  A complete bipartite $K_{n,n} \mapsto$ removing edges $\{v, v\}$, $\forall v \mapsto$ repeat

  > *"draw a perfect matching uniformly at random among all perfect matchings of the remaining graph"*

  for k times.

The bundle graph

The bundle graph:



- A random k-regular graph:

  A complete bipartite $K_{n,n} \mapsto$ removing edges $\{v, v\}$, $\forall v \mapsto$ repeat

  > *"draw a perfect matching uniformly at random among all perfect matchings of the remaining graph"*

  for k times.

The limitation on the order revealing scheme

- The property of revealing the ground truth for certain:

$$\forall x, y \in \mathcal{A},\ \exists B \in \mathcal{B} \text{ such that } x, y \in B.$$

  - Suppose NO bundle contains both $x, y \in \mathcal{A}$.
  - Let $\prec, \prec'$ be two complete rankings.
    - x, y are in the first two positions in $\prec, \prec'$;
    - $\prec$ and $\prec'$ differs only in the order of x and y.
  - Clearly, partial rankings within the bundles are identical in both cases.
  - No way to identify whether $\prec$ or $\prec'$ is the ground truth.

- To reveal the ground truth with certainty: $k = \Omega(\sqrt{n})$.
  - $n \cdot \binom{k}{2} \geq \binom{n}{2}$.

The limitation on the order revealing scheme

- The property of revealing the ground truth for certain:

$$\forall x, y \in \mathcal{A},\ \exists B \in \mathcal{B} \text{ such that } x, y \in B.$$

  - Suppose NO bundle contains both $x, y \in \mathcal{A}$.
  - Let $\prec, \prec'$ be two complete rankings.
    - x, y are in the first two positions in $\prec, \prec'$;
    - $\prec$ and $\prec'$ differs only in the order of x and y.
  - Clearly, partial rankings within the bundles are identical in both cases.
  - No way to identify whether $\prec$ or $\prec'$ is the ground truth.
- To reveal the ground truth with certainty: $k = \Omega(\sqrt{n})$.
  - $n \cdot \binom{k}{2} \geq \binom{n}{2}$.

The limitation on the order revealing scheme

- The property of revealing the ground truth for certain:

$$\forall x, y \in \mathcal{A}, \exists B \in \mathcal{B} \text{ such that } x, y \in B.$$

  - Suppose NO bundle contains both $x, y \in \mathcal{A}$.
  - Let $\prec, \prec'$ be two complete rankings.
    - x, y are in the first two positions in $\prec, \prec'$;
    - $\prec$ and $\prec'$ differs only in the order of x and y.
  - Clearly, partial rankings within the bundles are identical in both cases.
  - No way to identify whether $\prec$ or $\prec'$ is the ground truth.

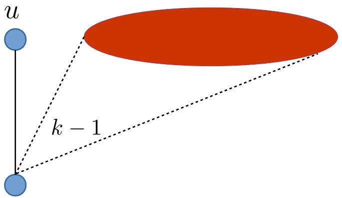
- To reveal the ground truth with certainty: $k = \Omega(\sqrt{n})$.
  - $n \cdot \binom{k}{2} \geq \binom{n}{2}$.

Seeking for approximate order-revealing grading schemes

- Use a bundle graph with a very low degree k (independent of n).
- Randomly permute the elements by $\pi : U \mapsto \mathcal{A}$ before associating them to the nodes of U of the bundle graph.
- Aiming at $\frac{\#\text{correctly recovered pairwise relations}}{\binom{n}{2}}$.

About who grading both u and v

- $\lambda_{u,v} := |N(u) \cap N(v)|$, for $u, v \in U$.
  - $\sum_{v \in U \setminus \{u\}} \lambda_{u,v} = k(k-1)$.



The main result

Theorem 1

When

- Borda is applied as the aggregation rule, and
- all the partial rankings are consistent to the ground truth,

then the expected fraction of correctly recovered pairwise relations is $1 - O(1/\sqrt{k})$.

- $W_{r,q}$: the r.v. denoting $B(a_r) - B(a_q)$ for $r < q$, $a_r, a_q \in \mathcal{A}$.
- $\Gamma^{r,q}_{u,v}$: the event that $\pi(u) = a_r$, $\pi(v) = a_q$.

- $$C := \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \mathbf{E}[\mathbb{1}\{W_{r,q} > 0\}] = \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \Pr[W_{r,q} > 0]$$
$$= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \sum_{u,v \in U} \Pr[W_{r,q} \leq 0 \mid \Gamma^{r,q}_{u,v}] \cdot \Pr[\Gamma^{r,q}_{u,v}]\right)$$
$$= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \frac{1}{n(n-1)} \sum_{u,v \in U} \Pr[W_{r,q} \leq 0 \mid \Gamma^{r,q}_{u,v}]\right)$$

- $W_{r,q}$: the r.v. denoting $B(a_r) - B(a_q)$ for $r < q$, $a_r, a_q \in \mathcal{A}$.
- $\Gamma_{u,v}^{r,q}$: the event that $\pi(u) = a_r$, $\pi(v) = a_q$.

-
$$
\begin{aligned}
C &:= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \mathbf{E}[\mathbb{1}\{W_{r,q} > 0\}] = \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \Pr[W_{r,q} > 0] \\
&= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \sum_{u,v \in U} \Pr[W_{r,q} \le 0 \mid \Gamma_{u,v}^{r,q}] \cdot \Pr[\Gamma_{u,v}^{r,q}]\right) \\
&= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \frac{1}{n(n-1)} \sum_{u,v \in U} \Pr[W_{r,q} \le 0 \mid \Gamma_{u,v}^{r,q}]\right)
\end{aligned}
$$

- $W_{r,q}$: the r.v. denoting $B(a_r) - B(a_q)$ for $r < q$, $a_r, a_q \in \mathcal{A}$.
- $\Gamma^{r,q}_{u,v}$: the event that $\pi(u) = a_r$, $\pi(v) = a_q$.

- $$\begin{aligned} C &:= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \mathbf{E}[\mathbb{1}\{W_{r,q} > 0\}] = \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \Pr[W_{r,q} > 0] \\ &= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \sum_{u,v \in U} \Pr[W_{r,q} \leq 0 \mid \Gamma^{r,q}_{u,v}] \cdot \Pr[\Gamma^{r,q}_{u,v}]\right) \\ &= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \frac{1}{n(n-1)} \sum_{u,v \in U} \Pr[W_{r,q} \leq 0 \mid \Gamma^{r,q}_{u,v}]\right) \end{aligned}$$

- $W_{r,q}$: the r.v. denoting $B(a_r) - B(a_q)$ for $r < q$, $a_r, a_q \in \mathcal{A}$.
- $\Gamma^{r,q}_{u,v}$: the event that $\pi(u) = a_r$, $\pi(v) = a_q$.

- $$\begin{aligned} C &:= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \mathbf{E}[\mathbb{1}\{W_{r,q} > 0\}] = \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \Pr[W_{r,q} > 0] \\ &= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \sum_{u,v \in U} \Pr[W_{r,q} \le 0 \mid \Gamma^{r,q}_{u,v}] \cdot \Pr[\Gamma^{r,q}_{u,v}]\right) \\ &= \sum_{r=1}^{n-1} \sum_{q=r+1}^{n} \left(1 - \frac{1}{n(n-1)} \sum_{u,v \in U} \Pr[W_{r,q} \le 0 \mid \Gamma^{r,q}_{u,v}]\right) \end{aligned}$$

- Given $\Gamma_{u,v}^{r,q}$,
  - the expected Borda score of a_r is $k + (k(k-1) - \lambda_{u,v}) \cdot \frac{n-r-1}{n-2} + \lambda_{u,v}$.
  - the expected Borda score of a_q is $k + (k(k-1) - \lambda_{u,v}) \cdot \frac{n-q}{n-2}$.
- Thus

$$\mathbf{E}[W_{r.q} \mid \Gamma_{u,v}^{r,q}] = (k(k-1) - \lambda_{u,v})\frac{q-r-1}{n-2} + \lambda_{u,v}.$$

Why?

- Given $\Gamma_{u,v}^{r,q}$,
  - the expected Borda score of a_r is $k + (k(k-1) - \lambda_{u,v}) \cdot \frac{n-r-1}{n-2} + \lambda_{u,v}$.
  - the expected Borda score of a_q is $k + (k(k-1) - \lambda_{u,v}) \cdot \frac{n-q}{n-2}$.
- Thus

$$\mathbf{E}[W_{r.q} \mid \Gamma_{u,v}^{r,q}] = (k(k-1) - \lambda_{u,v})\frac{q-r-1}{n-2} + \lambda_{u,v}.$$

Why?

Calculate the Borda score from another point of view

- Element a_r gets one point for each bundle it belongs to;
  - plus one additional point for each appearance of an element with rank higher $> r$ in the bundles a_r belongs to.

- In the bundles of containing a_r:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-r-1}{n-2}$ to have rank higher than r.
- $\mathbf{E}[B(a_r) \mid \Gamma^{r;q}_{u,v}] = k + (k(k-1) - \lambda_{u,v})\frac{n-r-1}{n-2} + \lambda_{u,v}$.
- In the bundles of containing a_q:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-q}{n-2}$ to have rank higher than q.
- $\mathbf{E}[B(a_q) \mid \Gamma^{r;q}_{u,v}] = k + (k(k-1) - \lambda_{u,v})\frac{n-q}{n-2}$.

- In the bundles of containing a_r:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-r-1}{n-2}$ to have rank higher than r.
- $\mathbf{E}[B(a_r) \mid \Gamma^{r,q}_{u,v}] = k + (k(k-1) - \lambda_{u,v})\frac{n-r-1}{n-2} + \lambda_{u,v}$.
- In the bundles of containing a_q:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-q}{n-2}$ to have rank higher than q.
- $\mathbf{E}[B(a_q) \mid \Gamma^{r,q}_{u,v}] = k + (k(k-1) - \lambda_{u,v})\frac{n-q}{n-2}$.

- In the bundles of containing a_r:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-r-1}{n-2}$ to have rank higher than r.
- $\mathbf{E}[B(a_r) \mid \Gamma^{r,q}_{u,v}] = k + (k(k-1) - \lambda_{u,v})\frac{n-r-1}{n-2} + \lambda_{u,v}$.
- In the bundles of containing a_q:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-q}{n-2}$ to have rank higher than q.
- $\mathbf{E}[B(a_q) \mid \Gamma^{r,q}_{u,v}] = k + (k(k-1) - \lambda_{u,v})\frac{n-q}{n-2}$.

- In the bundles of containing a_r:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-r-1}{n-2}$ to have rank higher than r.
- $\mathbf{E}[B(a_r) \mid \Gamma_{u,v}^{r,q}] = k + (k(k-1) - \lambda_{u,v})\frac{n-r-1}{n-2} + \lambda_{u,v}$.
- In the bundles of containing a_q:
  - $\lambda_{u,v}$ appearances of a_q in the bundles of a_r.
  - $k(k-1) - \lambda_{u,v}$ appearances of elements different than a_r, a_q.
    - Each of them has prob. $\frac{n-q}{n-2}$ to have rank higher than q.
- $\mathbf{E}[B(a_q) \mid \Gamma_{u,v}^{r,q}] = k + (k(k-1) - \lambda_{u,v})\frac{n-q}{n-2}$.

Dealing with dependencies

Goal: $\Pr[W_{r,q} \leq 0 \mid \Gamma_{u,v}^{r,q}]$

- Given $\Gamma_{u,v}^{r,q}$, define $S = N(N(u) \cup N(v)) \setminus \{u, v\}$.
- $o : [|S|] \mapsto S$ denotes an arbitrary ordering of nodes of S.
- X_i: the random variable denoting the rank of the element $\pi(o(i))$.
- Define the Doob martingale $Z_0, Z_1, \ldots, Z_{|S|}$ such that
  - $Z_0 = \mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}]$;
  - $Z_i = \mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}, X_1, \ldots, X_i]$.
- Hence, $W_{r,q} \mid \Gamma_{u,v}^{r,q} = Z_{|S|}$.

Martingale

Martingale

A sequence of random variables $Z_0, Z_1, \ldots, Z_m$ is a martingale w.r.t. a sequence of random variables $X_1, X_2, \ldots, X_m$ if $\forall i = 1, \ldots, m$,

$$\mathbf{E}[Z_i \mid X_1, \ldots, X_{i-1}] = Z_{i-1}.$$

Doob martingale (Joseph L. Doob (1910–2004))

- W: a random variable
- $X_1, \ldots, X_m$: a sequence of m random variables.

The sequence $Z_0, Z_1, \ldots, Z_m$ such that

- $Z_0 = \mathbf{E}[W]$;
- $Z_i = \mathbf{E}[W \mid X_1, \ldots, X_i]$, $\forall i = 1, \ldots, m$

is called a Doob martingale.

Azuma-Hoeffding inequality

Azuma-Hoeffding inequality

Let $Z_0, Z_1, \ldots, Z_m$ be a martingale with $Z_i - Z_{i-1} \leq c_i$ for $i = 1, \ldots, m$. Then, for all $t > 0$,

$$\Pr[Z_m - Z_0 \leq -t] \leq \exp\left(-\frac{t^2}{2\sum_{i=1}^{m} c_i^2}\right).$$

Dealing with dependencies (contd.)

- Given $\Gamma_{u,v}^{r,q}$, define $S = N(N(u) \cup N(v)) \setminus \{u, v\}$.
- $o : [|S|] \mapsto S$ denotes an arbitrary ordering of nodes of S.
- X_i: the random variable denoting the rank of the element $\pi(o(i))$.
- Define the Doob martingale $Z_0, Z_1, \ldots, Z_{|S|}$ such that
  - $Z_0 = \mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}]$;
  - $Z_i = \mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}, X_1, \ldots, X_i]$.
- Hence, $W_{r,q} \mid \Gamma_{u,v}^{r,q} = Z_{|S|}$.

Lemma 8

$\forall i \in \{1, 2, \ldots, |S|\}$, it holds that $|Z_i - Z_{i-1}| \le 2(\lambda_{u,o(i)} + \lambda_{v,o(i)})$.

Lemma 3

For every k-regular bipartite graph G,

$\theta_{u,v} = 4 \sum_{z \in (N(u) \cup N(v)) \setminus \{u,v\}} (\lambda_{u,z} + \lambda_{v,z})^2 \le 8k(k-1)(4k-3)$.

- Set $t = \mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}]$ $(= Z_0)$, by the Azuma-Hoeffding inequality:

$$\begin{aligned}
\Pr[Z_{|S|} - Z_0 \le -t] &= \Pr[W_{r,q} \le 0 \mid \Gamma_{u,v}^{r,q}] \\
&\le \exp\left(-\frac{t^2}{2\sum_{i=1}^{m} c_i^2}\right) \\
&= \exp\left(-\frac{\mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}]^2}{2\theta_{u,v}}\right).
\end{aligned}$$

Back to the computation of C

$$
\begin{aligned}
C &= \sum_{r=1}^{n-1}\sum_{q=r+1}^{n}\left(1-\frac{1}{n(n-1)}\sum_{u,v\in U}\Pr[W_{r,q}\le 0 \mid \Gamma_{u,v}^{r,q}]\right)\\
&\ge \sum_{r=1}^{n-1}\sum_{q=r+1}^{n}\left(1-\frac{1}{n(n-1)}\sum_{u,v\in U}\exp\left(-\frac{\mathbf{E}[W_{r,q}\mid \Gamma_{u,v}^{r,q}]^2}{2\theta_{u,v}}\right)\right)\\
&= \sum_{r=1}^{n-1}\sum_{q=r+1}^{n}\left(1-\frac{1}{n(n-1)}\sum_{u,v\in U} e^{-(\beta(u,v)\cdot y(q-r)+\delta(u,v))^2}\right)\\
&= \frac{n(n-1)}{2}-\frac{1}{n(n-1)}\sum_{u,v\in U}\sum_{d=1}^{n-1}(n-d)e^{-(\beta(u,v)\cdot y(d)+\delta(u,v))^2}\\
&\ge \frac{n(n-1)}{2}-\sum_{u,v\in U}\int_0^1(1-y)e^{-(\beta(u,v)\cdot y+\delta(u,v))^2}dy
\end{aligned}
$$

- $\beta(u,v)=\frac{k(k-1)-\lambda_{u,v}}{\sqrt{2\theta_{u,v}}}$;
- $\delta(u,v)=\frac{\lambda_{u,v}}{\sqrt{2\theta_{u,v}}}$;
- $y(t)=\frac{t-1}{n-2}$.

$$
\begin{aligned}
C &\geq \frac{n(n-1)}{2} - \sum_{u,v\in U} \int_0^1 (1-y)e^{-(\beta(u,v)\cdot y+\delta(u,v))^2}dy \\
&\geq \frac{n(n-1)}{2} - \sum_{u,v\in U} \frac{\beta(u,v)+\delta(u,v)}{2\beta(u,v)^2}\sqrt{\pi} \\
&\geq \frac{n(n-1)}{2} - \frac{k-1}{k(k-2)^2}\sqrt{\frac{\pi}{2}} \sum_{u,v\in U} \sqrt{\theta_{u,v}} \\
&\geq \frac{n(n-1)}{2}\left(1 - \frac{48\sqrt{2\pi}}{\sqrt{k}})\right).
\end{aligned}
$$

Claim 9

Let $\beta > 0$, $\delta \geq 0$, then $\int_0^1 (1-y)e^{-(\beta y+\delta)^2}dy \leq \frac{\beta+\delta}{2\beta^2}\sqrt{\pi}$.

- $k \geq 3$ (assumption).

Thank you.

Proof of Lemma 8

Recall:

- Given $\Gamma^{r,q}_{u,v}$, define $S = N(N(u) \cup N(v)) \setminus \{u, v\}$.
- $o : [|S|] \mapsto S$ denotes an arbitrary ordering of nodes of S.
- X_i: the random variable denoting the rank of the element $\pi(o(i))$.
- Define the Doob martingale $Z_0, Z_1, \ldots, Z_{|S|}$ such that
  - $Z_0 = \mathbf{E}[W_{r,q} \mid \Gamma^{r,q}_{u,v}]$;
  - $Z_i = \mathbf{E}[W_{r,q} \mid \Gamma^{r,q}_{u,v}, X_1, \ldots, X_i]$.
- Hence, $W_{r,q} \mid \Gamma^{r,q}_{u,v} = Z_{|S|}$.

Lemma 8

$\forall i \in \{1, 2, \ldots, |S|\}$, it holds that $|Z_i - Z_{i-1}| \leq 2(\lambda_{u,o(i)} + \lambda_{v,o(i)})$.

- $\mu_{u,v,w} = |N(u) \cap N(v) \cap N(w)|$.

Key to the proof of Lemma 8

- The Borda score difference $W_{r,q}$ (conditioned on $\Gamma_{u,v}^{r,q}$):
  - **increases** for each appearance of a_q in the same bundle with a_r;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing a_r but NOT a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing BOTH a_r, a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **decreases** for each appearance of $\pi(o(j))$ in a bundle containing a_q but NOT a_r provided that $\text{rank}(\pi(o(j))) > q$.

$$\begin{aligned} W_{r,q} &= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} - \mu_{u,v,o(j)}) \cdot \mathbb{1}\{X_j > r\} + \sum_{j=1}^{|S|} \mu_{u,v,o(j)} \cdot \mathbb{1}\{r < X_j < q\} \\ &- \sum_{j=1}^{|S|}(\lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}) \\ &= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}). \end{aligned}$$

Key to the proof of Lemma 8

- The Borda score difference $W_{r,q}$ (conditioned on $\Gamma_{u,v}^{r,q}$):
  - **increases** for each appearance of a_q in the same bundle with a_r;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing a_r but NOT a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing BOTH a_r, a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **decreases** for each appearance of $\pi(o(j))$ in a bundle containing a_q but NOT a_r provided that $\text{rank}(\pi(o(j))) > q$.

$$
\begin{aligned}
W_{r,q} &= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} - \mu_{u,v,o(j)}) \cdot \mathbb{1}\{X_j > r\} + \sum_{j=1}^{|S|} \mu_{u,v,o(j)} \cdot \mathbb{1}\{r < X_j < q\} \\
&- \sum_{j=1}^{|S|}(\lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}) \\
&= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}).
\end{aligned}
$$

Key to the proof of Lemma 8

- The Borda score difference $W_{r,q}$ (conditioned on $\Gamma_{u,v}^{r,q}$):
  - **increases** for each appearance of a_q in the same bundle with a_r;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing a_r but NOT a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing BOTH a_r, a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **decreases** for each appearance of $\pi(o(j))$ in a bundle containing a_q but NOT a_r provided that $\text{rank}(\pi(o(j))) > q$.

$$\begin{aligned} W_{r,q} &= \lambda_{u,v} + \sum_{j=1}^{|S|} (\lambda_{u,o(j)} - \mu_{u,v,o(j)}) \cdot \mathbb{1}\{X_j > r\} + \sum_{j=1}^{|S|} \mu_{u,v,o(j)} \cdot \mathbb{1}\{r < X_j < q\} \\ &- \sum_{j=1}^{|S|} (\lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}) \\ &= \lambda_{u,v} + \sum_{j=1}^{|S|} (\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}). \end{aligned}$$

Key to the proof of Lemma 8

- The Borda score difference $W_{r,q}$ (conditioned on $\Gamma^{r,q}_{u,v}$):
  - **increases** for each appearance of a_q in the same bundle with a_r;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing a_r but NOT a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing BOTH a_r, a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **decreases** for each appearance of $\pi(o(j))$ in a bundle containing a_q but NOT a_r provided that $\text{rank}(\pi(o(j))) > q$.

$$W_{r,q} = \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} - \mu_{u,v,o(j)}) \cdot \mathbb{1}\{X_j > r\} + \sum_{j=1}^{|S|} \mu_{u,v,o(j)} \cdot \mathbb{1}\{r < X_j < q\}$$

$$- \sum_{j=1}^{|S|}(\lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\})$$

$$= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}).$$

Key to the proof of Lemma 8

- The Borda score difference $W_{r,q}$ (conditioned on $\Gamma_{u,v}^{r,q}$):
  - **increases** for each appearance of a_q in the same bundle with a_r;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing a_r but NOT a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing BOTH a_r, a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **decreases** for each appearance of $\pi(o(j))$ in a bundle containing a_q but NOT a_r provided that $\text{rank}(\pi(o(j))) > q$.

$$
\begin{aligned}
W_{r,q} &= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} - \mu_{u,v,o(j)}) \cdot \mathbb{1}\{X_j > r\} + \sum_{j=1}^{|S|} \mu_{u,v,o(j)} \cdot \mathbb{1}\{r < X_j < q\} \\
&- \sum_{j=1}^{|S|}(\lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}) \\
&= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}).
\end{aligned}
$$

Key to the proof of Lemma 8

- The Borda score difference $W_{r,q}$ (conditioned on $\Gamma^{r,q}_{u,v}$):
  - **increases** for each appearance of a_q in the same bundle with a_r;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing a_r but NOT a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing BOTH a_r, a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **decreases** for each appearance of $\pi(o(j))$ in a bundle containing a_q but NOT a_r provided that $\text{rank}(\pi(o(j))) > q$.

$$
\begin{aligned}
W_{r,q} &= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} - \mu_{u,v,o(j)}) \cdot \mathbb{1}\{X_j > r\} + \sum_{j=1}^{|S|}\mu_{u,v,o(j)} \cdot \mathbb{1}\{r < X_j < q\} \\
&- \sum_{j=1}^{|S|}(\lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}) \\
&= \lambda_{u,v} + \sum_{j=1}^{|S|}(\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}).
\end{aligned}
$$

Key to the proof of Lemma 8

- The Borda score difference $W_{r,q}$ (conditioned on $\Gamma_{u,v}^{r,q}$):
  - **increases** for each appearance of a_q in the same bundle with a_r;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing a_r but NOT a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **increases** for each appearance of $\pi(o(j))$ in a bundle containing BOTH a_r, a_q provided that $r < \text{rank}(\pi(o(j))) < q$;
  - **decreases** for each appearance of $\pi(o(j))$ in a bundle containing a_q but NOT a_r provided that $\text{rank}(\pi(o(j))) > q$.

$$
\begin{aligned}
W_{r,q} &= \lambda_{u,v} + \sum_{j=1}^{|S|} (\lambda_{u,o(j)} - \mu_{u,v,o(j)}) \cdot \mathbb{1}\{X_j > r\} + \sum_{j=1}^{|S|} \mu_{u,v,o(j)} \cdot \mathbb{1}\{r < X_j < q\} \\
&- \sum_{j=1}^{|S|} (\lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}) \\
&= \lambda_{u,v} + \sum_{j=1}^{|S|} (\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\}).
\end{aligned}
$$

Key to the proof of Lemma 8 (contd.)

$$W_{r,q} = \lambda_{u,v} + \sum_{j=1}^{|S|} (\lambda_{u,o(j)} \cdot \mathbb{1}\{X_j > r\} - \lambda_{v,o(j)} \cdot \mathbb{1}\{X_j > q\})$$

$$\begin{aligned} Z_i - Z_{i-1} &= \mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}, X_1, \ldots, X_i] - \mathbf{E}[W_{r,q} \mid \Gamma_{u,v}^{r,q}, X_1, \ldots, X_{i-1}] \\ &= \sum_{j=i}^{|S|} \lambda_{u,o(j)} \left(\Pr[X_j > r \mid X_1, \ldots, X_i] - \Pr[X_j > r \mid X_1, \ldots, X_{i-1}] \right) - \\ &\quad \sum_{j=i}^{|S|} \lambda_{v,o(j)} \left(\Pr[X_j > q \mid X_1, \ldots, X_i] - \Pr[X_j > q \mid X_1, \ldots, X_{i-1}] \right). \end{aligned}$$

Key to the proof of Lemma 8 (contd.)

$$
\begin{aligned}
Z_i - Z_{i-1} &= \mathbf{E}[W_{r,q} \mid \Gamma^{r,q}_{u,v}, X_1, \ldots, X_i] - \mathbf{E}[W_{r,q} \mid \Gamma^{r,q}_{u,v}, X_1, \ldots, X_{i-1}] \\
&= \sum_{j=i}^{|S|} \lambda_{u,o(j)} \left(\Pr[X_j > r \mid X_1, \ldots, X_i] - \Pr[X_j > r \mid X_1, \ldots, X_{i-1}] \right) - \\
&\quad \sum_{j=i}^{|S|} \lambda_{v,o(j)} \left(\Pr[X_j > q \mid X_1, \ldots, X_i] - \Pr[X_j > q \mid X_1, \ldots, X_{i-1}] \right).
\end{aligned}
$$

Note that:

$$\Pr[X_j > r \mid X_1, \ldots, X_{i-1}] = \tfrac{x+y}{n-i-1}, \qquad \Pr[X_j > q \mid X_1, \ldots, X_{i-1}] = \tfrac{y}{n-i-1} \quad (i \le j \le |S|)$$

$$\Pr[X_j > r \mid X_1, \ldots, X_i] = \tfrac{x+y-\mathbb{1}\{X_i > r\}}{n-i-2}, \ \Pr[X_j > q \mid X_1, \ldots, X_i] = \tfrac{y-\mathbb{1}\{X_i > q\}}{n-i-2} \quad (i+1 \le j \le |S|)$$

- x: # available ranks from $[n] \setminus \{r, q, X_1, \ldots, X_{i-1}\}$ that are between r and q;
- y: # available ranks from $[n] \setminus \{r, q, X_1, \ldots, X_{i-1}\}$ that are $> q$.

Key to the proof of Lemma 8 (contd.)

$$
\begin{aligned}
Z_i - Z_{i-1} \;=\; & \lambda_{u,o(i)}\left(\mathbb{1}\{X_i > r\} - \frac{x+y}{n-i-1}\right) - \lambda_{v,o(j)}\left(\mathbb{1}\{X_i > q\} - \frac{y}{n-i-1}\right) \\
& + \sum_{j=i+1}^{|S|} \lambda_{u,o(j)}\left(\frac{x+y-\mathbb{1}\{X_i > r\}}{n-i-2} - \frac{x+y}{n-i-1}\right) \\
& + \sum_{j=i+1}^{|S|} \lambda_{v,o(j)}\left(\frac{y-\mathbb{1}\{X_i > q\}}{n-i-2} - \frac{y}{n-i-1}\right)
\end{aligned}
$$

Note that:

$$
\Pr[X_j > r \mid X_1, \ldots, X_{i-1}] = \tfrac{x+y}{n-i-1}, \qquad \Pr[X_j > q \mid X_1, \ldots, X_{i-1}] = \tfrac{y}{n-i-1} \qquad (i \le j \le |S|)
$$

$$
\Pr[X_j > r \mid X_1, \ldots, X_i] = \tfrac{x+y-\mathbb{1}\{X_i > r\}}{n-i-2}, \; \Pr[X_j > q \mid X_1, \ldots, X_i] = \tfrac{y-\mathbb{1}\{X_i > q\}}{n-i-2} \quad (i+1 \le j \le |S|)
$$

- x: # available ranks from $[n] \setminus \{r, q, X_1, \ldots, X_{i-1}\}$ that are between r and q;
- y: # available ranks from $[n] \setminus \{r, q, X_1, \ldots, X_{i-1}\}$ that are $> q$.

Key to the proof of Lemma 8 (contd.)

$$\begin{aligned} Z_i - Z_{i-1} &= \lambda_{u,o(i)} \left(\mathbb{1}\{X_i > r\} - \frac{x+y}{n-i-1} \right) - \lambda_{v,o(j)} \left(\mathbb{1}\{X_i > q\} - \frac{y}{n-i-1} \right) \\ &+ \sum_{j=i+1}^{|S|} \lambda_{u,o(j)} \left(\frac{x+y-\mathbb{1}\{X_i > r\}}{n-i-2} - \frac{x+y}{n-i-1} \right) \\ &+ \sum_{j=i+1}^{|S|} \lambda_{v,o(j)} \left(\frac{y-\mathbb{1}\{X_i > q\}}{n-i-2} - \frac{y}{n-i-1} \right) \end{aligned}$$

Note that:

$$\Pr[X_j > r \mid X_1, \ldots, X_{i-1}] = \tfrac{x+y}{n-i-1}, \qquad \Pr[X_j > q \mid X_1, \ldots, X_{i-1}] = \tfrac{y}{n-i-1} \qquad (i \le j \le |S|)$$

$$\Pr[X_j > r \mid X_1, \ldots, X_i] = \tfrac{x+y-\mathbb{1}\{X_i>r\}}{n-i-2}, \ \Pr[X_j > q \mid X_1, \ldots, X_i] = \tfrac{y-\mathbb{1}\{X_i>q\}}{n-i-2} \quad (i+1 \le j \le |S|)$$

- x: # available ranks from $[n] \setminus \{r, q, X_1, \ldots, X_{i-1}\}$ that are between r and q;
- y: # available ranks from $[n] \setminus \{r, q, X_1, \ldots, X_{i-1}\}$ that are $> q$.

Key to the proof of Lemma 8 (contd.)

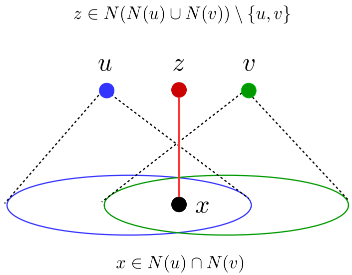
(Assume $n \geq 3k(k-1)+2$; Note: $|S| \leq 2k(k-1)$)

$$\begin{aligned}
Z_i - Z_{i-1} &= \lambda_{u,o(i)}\left(\mathbb{1}\{X_i > r\} - \frac{x+y}{n-i-1}\right) - \lambda_{v,o(j)}\left(\mathbb{1}\{X_i > q\} - \frac{y}{n-i-1}\right) \\
&+ \sum_{j=i+1}^{|S|} \lambda_{u,o(j)}\left(\frac{x+y-\mathbb{1}\{X_i > r\}}{n-i-2} - \frac{x+y}{n-i-1}\right) \\
&+ \sum_{j=i+1}^{|S|} \lambda_{v,o(j)}\left(\frac{y-\mathbb{1}\{X_i > q\}}{n-i-2} - \frac{y}{n-i-1}\right) \\
&= \left(\lambda_{u,o(i)} - \frac{\sum_{j=i+1}^{|S|}\lambda_{u,o(j)}}{n-i-2}\right)\cdot\left(\mathbb{1}\{X_i > r\} - \frac{x+y}{n-i-1}\right) \\
&\quad + \left(\lambda_{v,o(i)} - \frac{\sum_{j=i+1}^{|S|}\lambda_{v,o(j)}}{n-i-2}\right)\cdot\left(\frac{y}{n-i-1} - \mathbb{1}\{X_i > q\}\right).
\end{aligned}$$

Key to the proof of Lemma 3

Lemma 3

For every k-regular bipartite graph G,
$\theta_{u,v} = 4\sum_{z\in(N(u)\cup N(v))\setminus\{u,v\}}(\lambda_{u,z}+\lambda_{v,z})^2 \leq 8k(k-1)(4k-3)$.

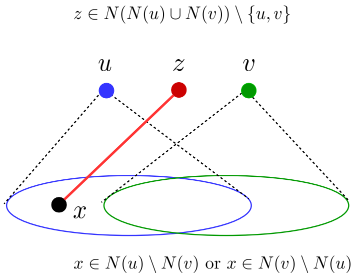


- The edge contributes 2 to the quantity $\lambda_{u,z}+\lambda_{v,z}$.
  - $\lambda_{u,z}+\lambda_{v,z} \leq 2k$.
- The edge contributes $\leq (2k)^2-(2k-2)^2 = 8k-4$ to the quantity $(\lambda_{u,z}+\lambda_{v,z})^2$.
- There are $|N(u)\cap N(v)|(k-2)$ such edges.

Key to the proof of Lemma 3 (contd.)

Lemma 3

For every k-regular bipartite graph G,
$\theta_{u,v} = 4\sum_{z\in(N(u)\cup N(v))\setminus\{u,v\}}(\lambda_{u,z}+\lambda_{v,z})^2 \le 8k(k-1)(4k-3)$.



- The edge contributes 1 to the quantity $\lambda_{u,z} + \lambda_{v,z}$.
  - $\lambda_{u,z} + \lambda_{v,z} \le 2k - 1$.
- The edge contributes $\le (2k-1)^2 - (2k-2)^2 = 4k - 3$ to the quantity $(\lambda_{u,z} + \lambda_{v,z})^2$.
- There are $2(k - |N(u) \cap N(v)|)(k - 1)$ such edges.

Proof of the Gaussian integral (Claim 9)

$$\int_0^1 (1-y)e^{-(\beta y+\delta)^2}dy = \underbrace{\int_0^1 e^{-(\beta y+\delta)^2}dy}_{A} - \underbrace{\int_0^1 ye^{-(\beta y+\delta)^2}dy}_{B}$$

- The error function:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt.$$

- The Gaussian integral:

$$\int_{-\infty}^{\infty} e^{-t^2}dt = \sqrt{\pi}.$$

$$\therefore \quad \int_0^{\infty} e^{-t^2}dt = \frac{\sqrt{\pi}}{2} \quad \text{and} \quad \mathrm{erf}(x) \le \frac{2}{\sqrt{\pi}}\cdot\frac{\sqrt{\pi}}{2} = 1.$$

- $B = \int_0^1 y e^{-(\beta y+\delta)^2} dy$.

  Let $v = \beta y + \delta \quad \therefore \begin{cases} dv = \beta dy \\ y = \frac{v-\delta}{\beta} \end{cases}$.

$$
\begin{aligned}
\therefore B &= \frac{1}{\beta} \int_{\delta}^{\beta+\delta} \frac{v-\delta}{\beta} \cdot e^{-v^2} dv \\
&= \frac{1}{\beta^2} \left(\int_{\delta}^{\beta+\delta} v \cdot e^{-v^2} dv - \delta \cdot \int_{\delta}^{\beta+\delta} e^{-v^2} dv \right) \\
&= \frac{1}{2\beta^2} \int_{\delta}^{\beta+\delta} e^{-v^2} dv^2 - \frac{\delta}{\beta^2} \cdot \frac{\sqrt{\pi}}{2} \cdot \frac{2}{\sqrt{\pi}} \int_{\delta}^{\beta+\delta} e^{-v^2} dv \\
&= \frac{1}{2\beta^2} \left(-e^{-(\beta+\delta)^2} + e^{-\delta^2} \right) - \frac{\delta\sqrt{\pi}}{2\beta^2} \left(\mathrm{erf}(\beta+\delta) - \mathrm{erf}(\delta) \right).
\end{aligned}
$$

- $A = \int_0^1 e^{-(\beta y+\delta)^2} dy$.
  Let $u = \beta y + \delta \quad \therefore du = \beta dy$.

$$\begin{aligned} \therefore A &= \frac{1}{\beta} \int_{\delta}^{\beta+\delta} e^{-u^2} du \\ &= \frac{\sqrt{\pi}}{2\beta} \cdot \frac{2}{\sqrt{\pi}} \left(\int_0^{\beta+\delta} e^{-u^2} du - \int_0^{\delta} e^{-u^2} du \right) \\ &= \frac{\sqrt{\pi}}{2\beta} \left(\mathrm{erf}(\beta+\delta) - \mathrm{erf}(\delta) \right). \end{aligned}$$

- Thus

$$\begin{aligned} A - B &= \frac{\beta+\delta}{2\beta^2} \sqrt{\pi} \left(\mathrm{erf}(\beta+\delta) - \mathrm{erf}(\delta) \right) + \frac{1}{2\beta^2} \left(e^{-(\beta+\delta)^2} - e^{-\delta^2} \right) \\ &\leq \frac{\beta+\delta}{2\beta^2} \sqrt{\pi}. \end{aligned}$$