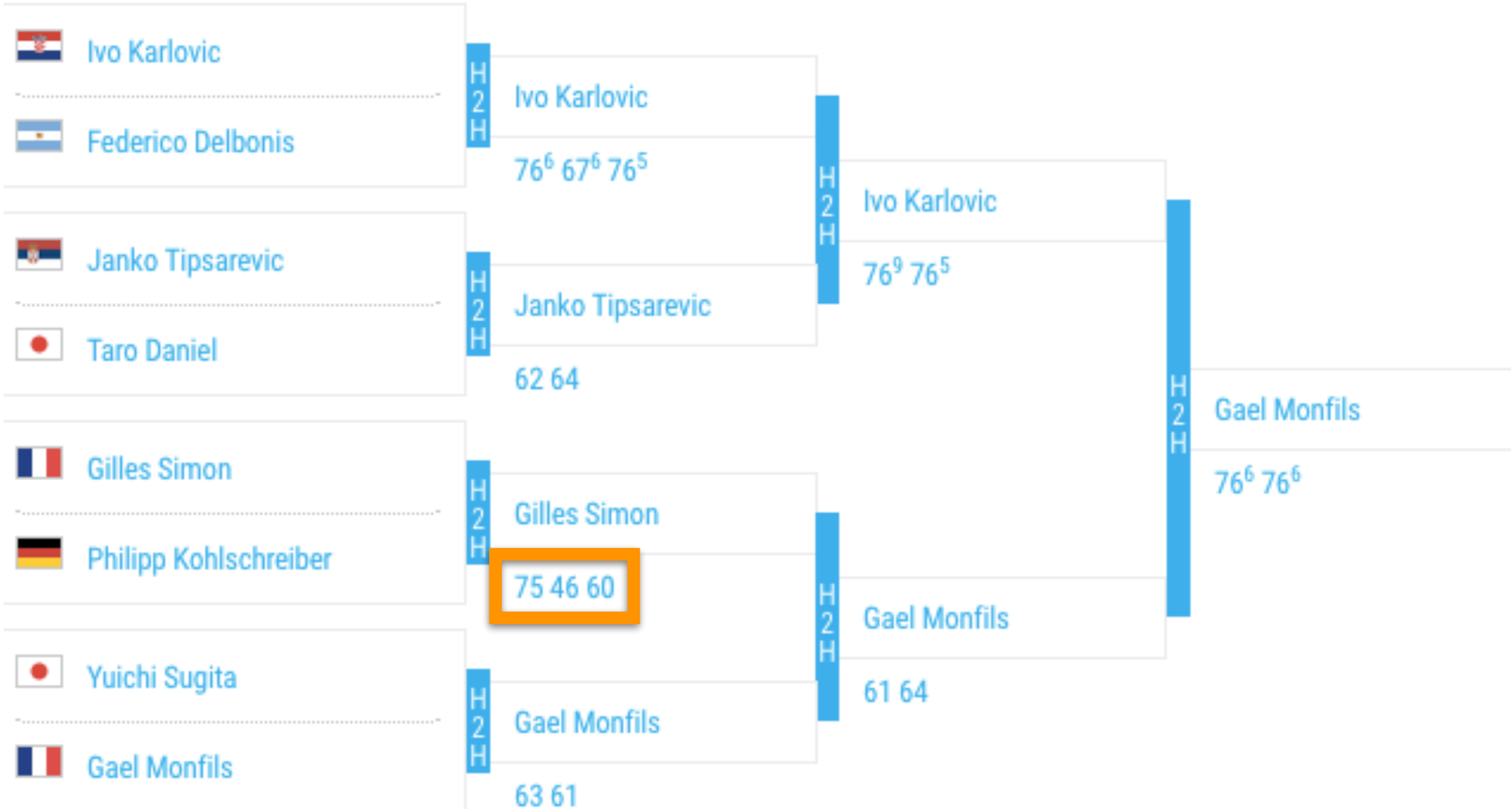
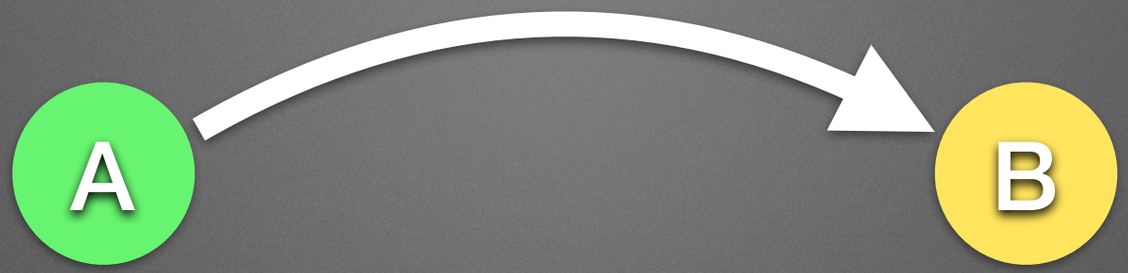
Tennis Players Network

Chia-Hung Yang

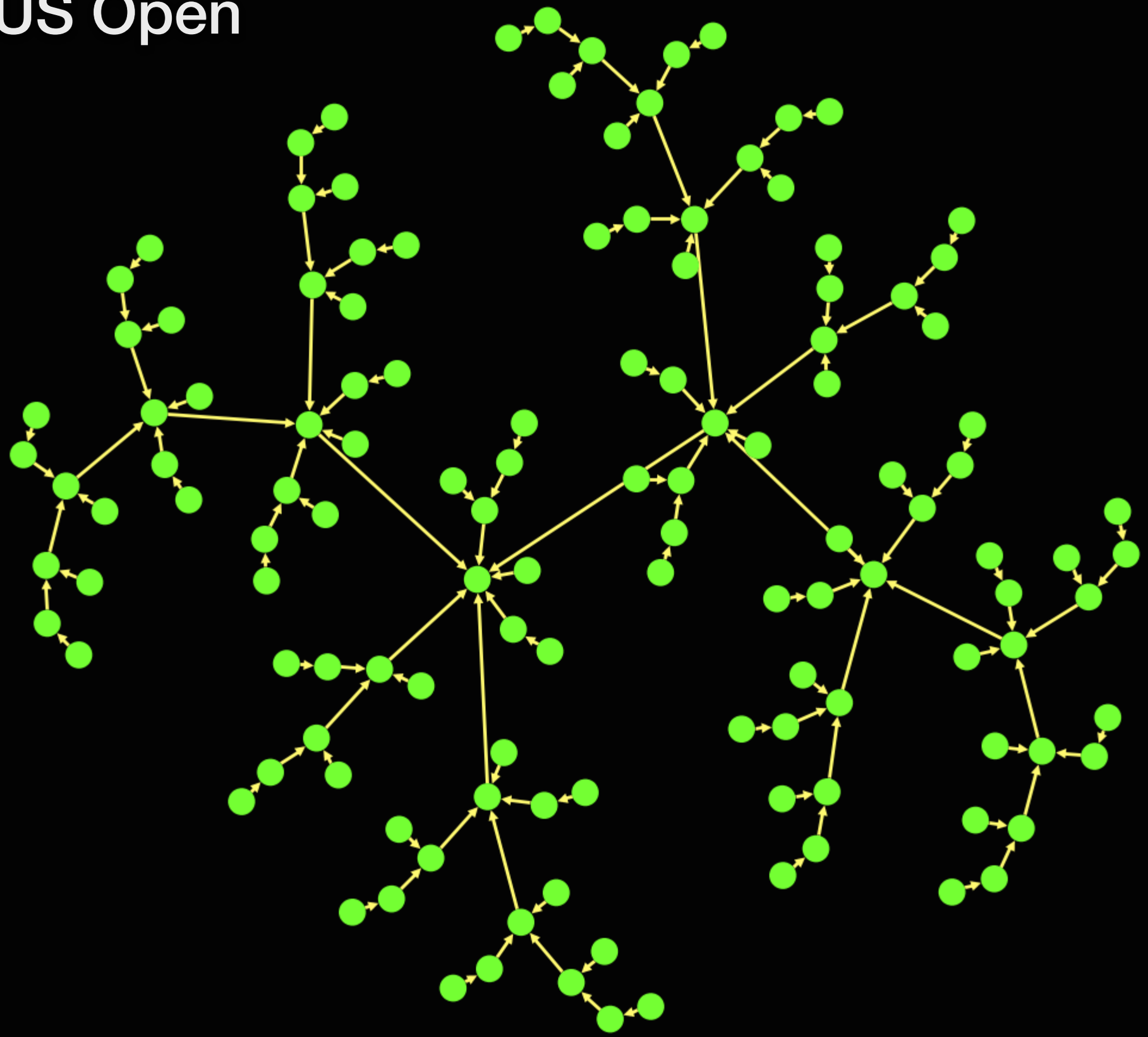


# of sets B won over A

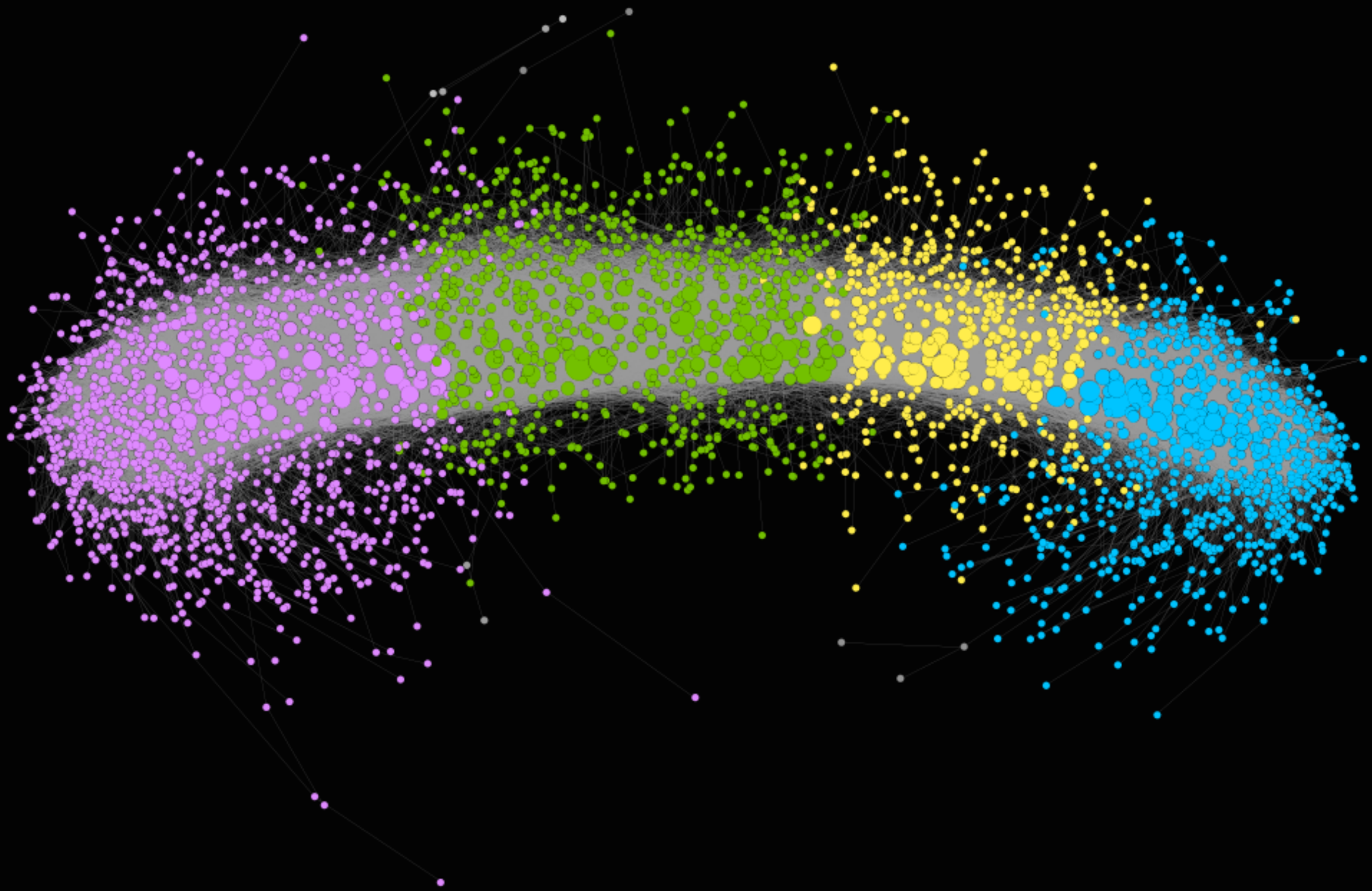


- Directly related to players’ performance
- Weighted and directed
- Collect data from 1976 to 2015
- N = 3230; L = 106774

2015 US Open



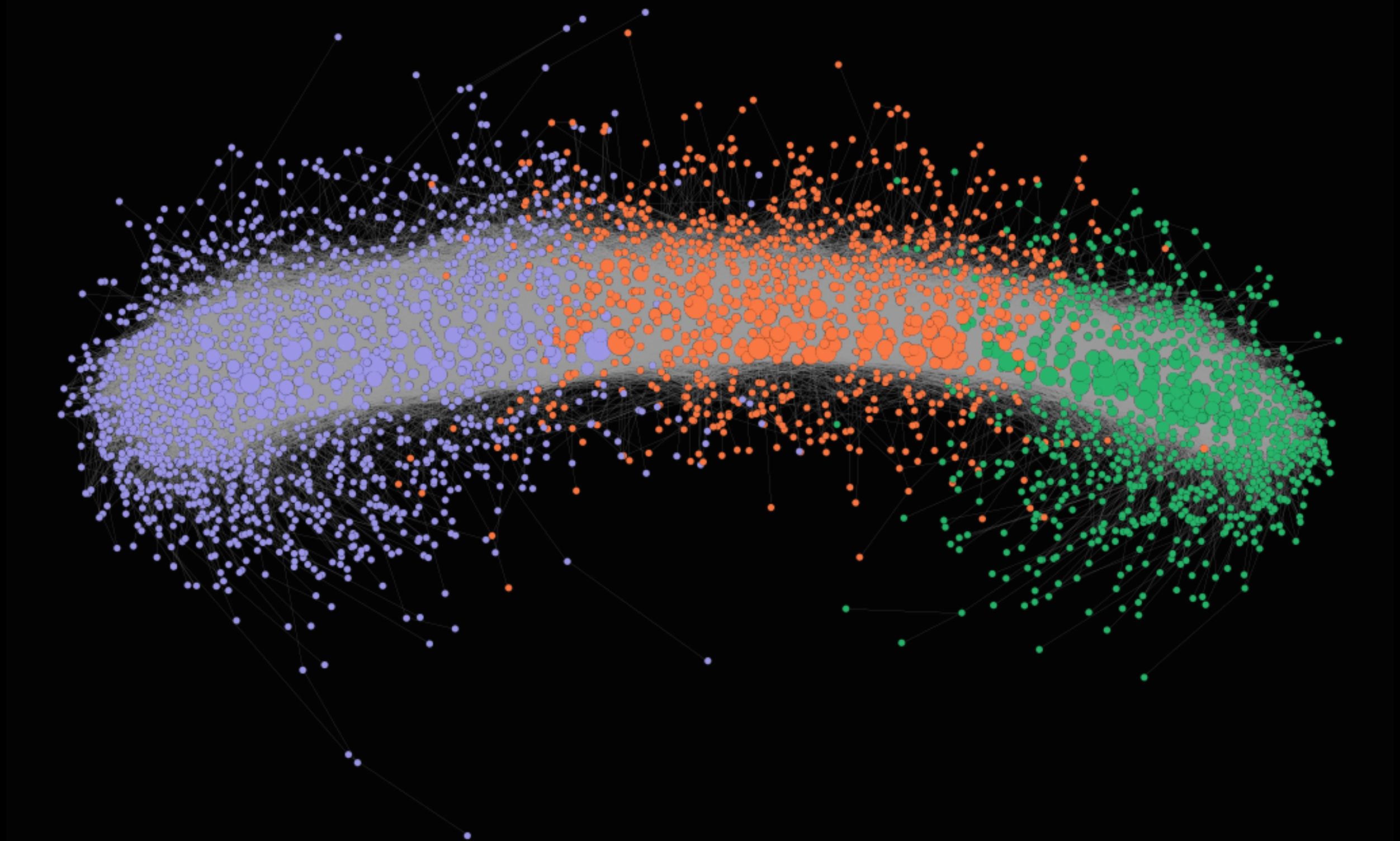
1976 ~ 2015



1976~1989

1990~2002

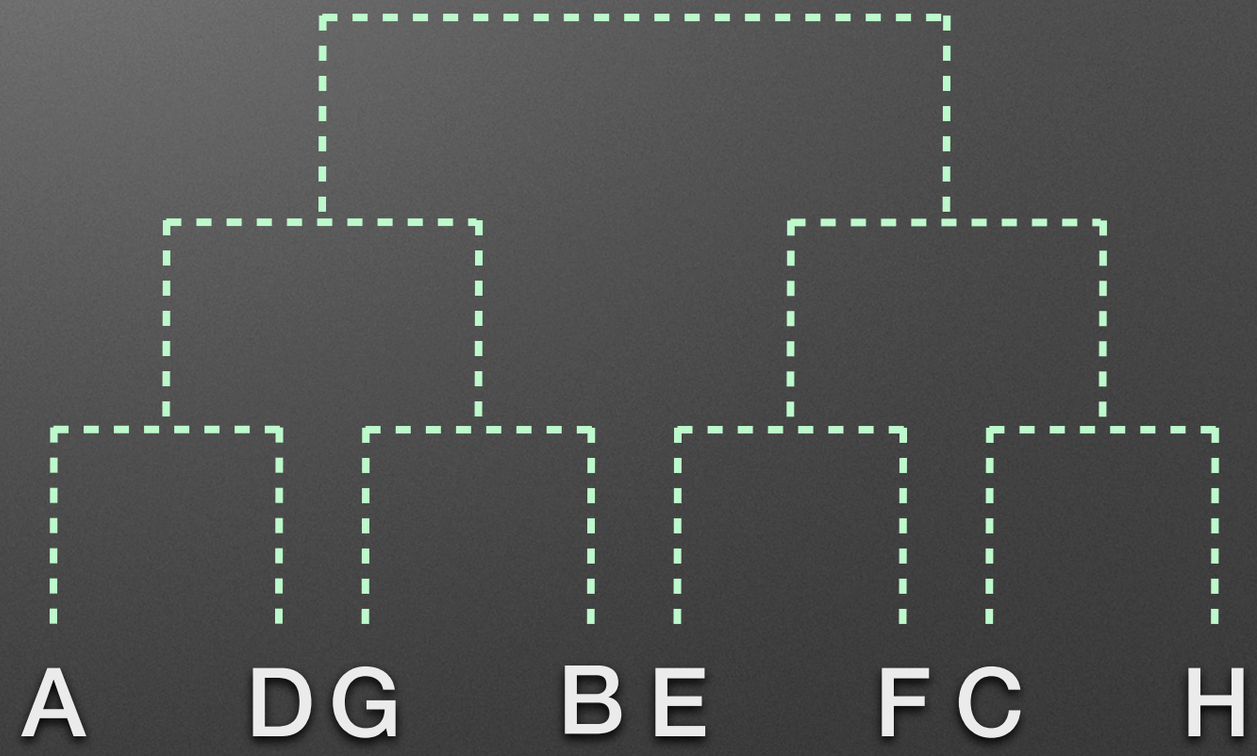
2003~2015



Null Model

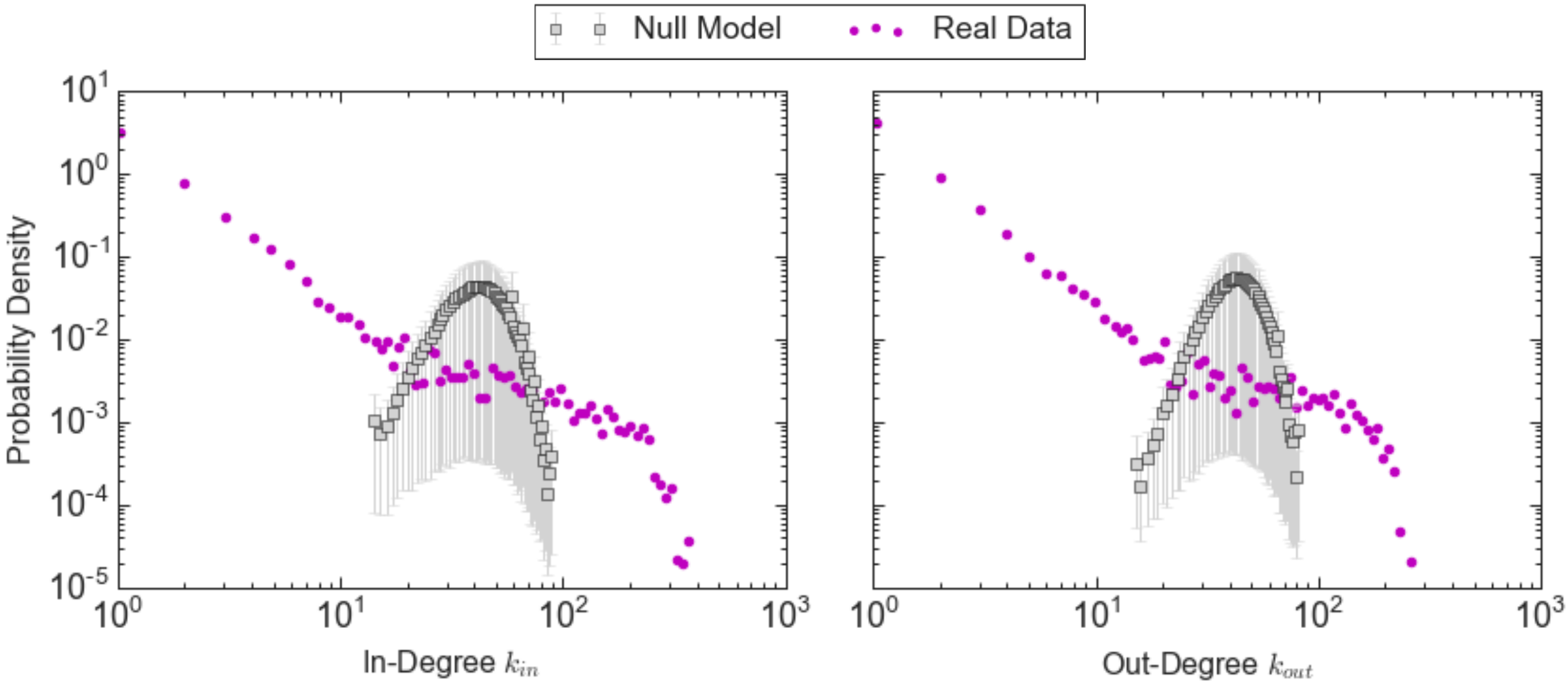
(players are indistinguishable)

- Randomly choose players from the pool
- Each player wins a set with probability 1/2



Tennis Professionals Pool

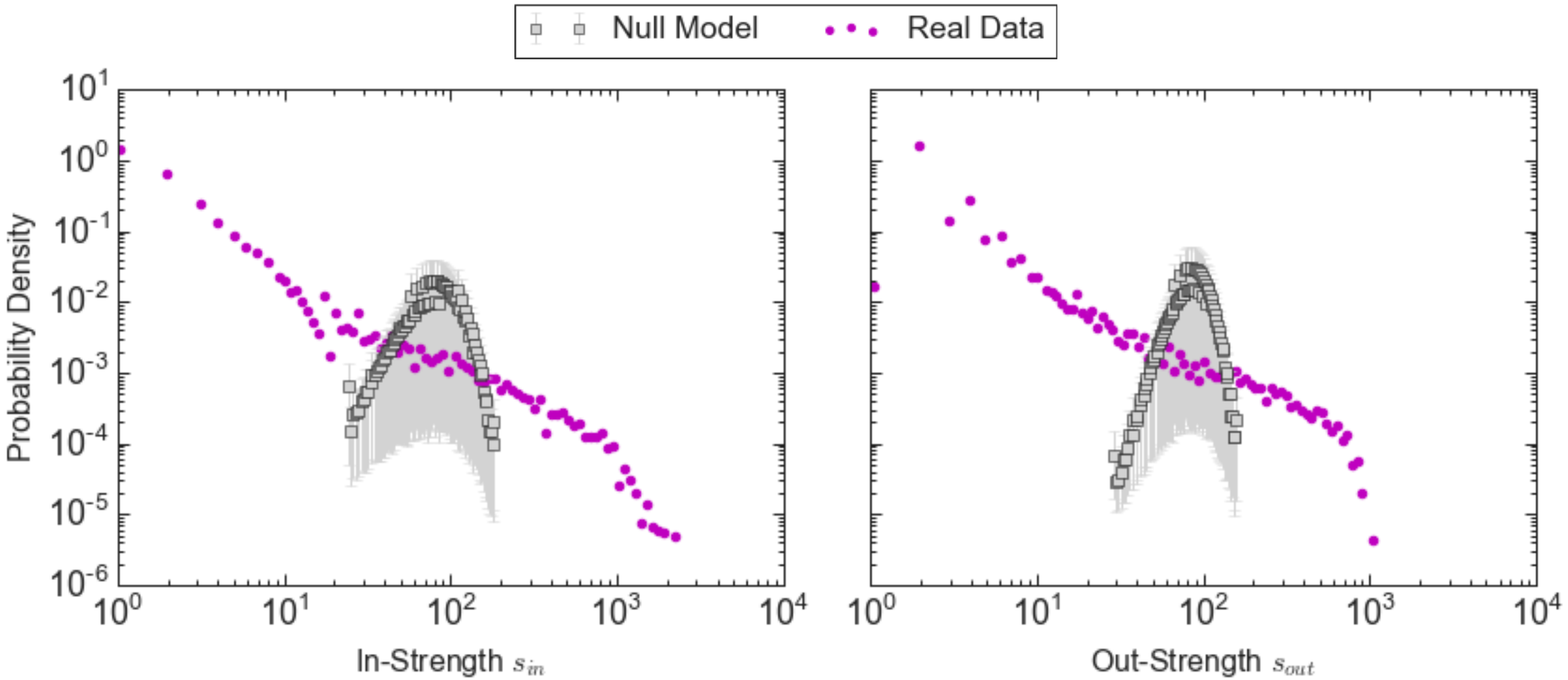
Degree Distribution



Real Data $\langle k_{in} \rangle = 33.040$

Null Model $\langle k_{in} \rangle = 46.493 \pm 0.045$

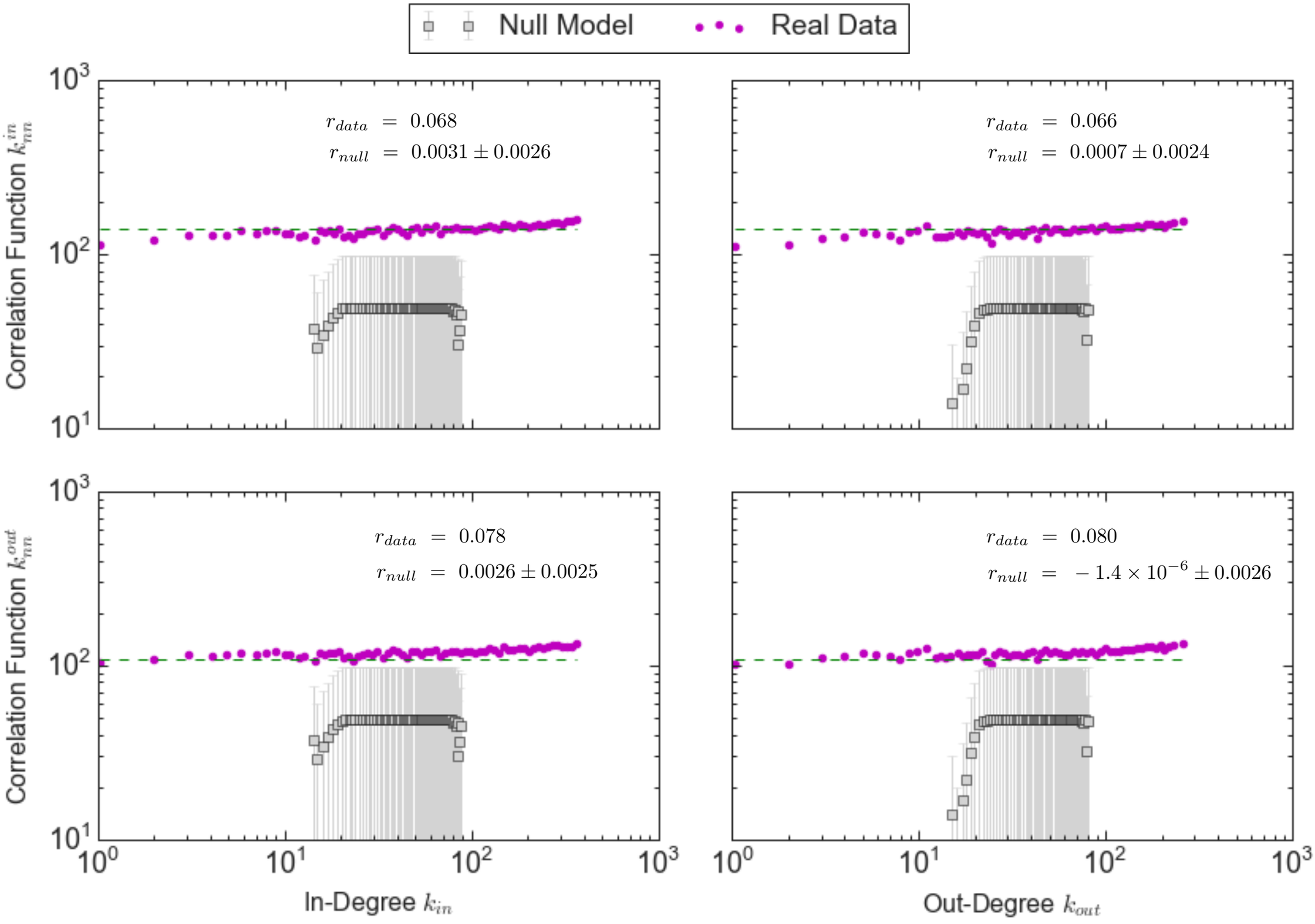
Strength Distribution



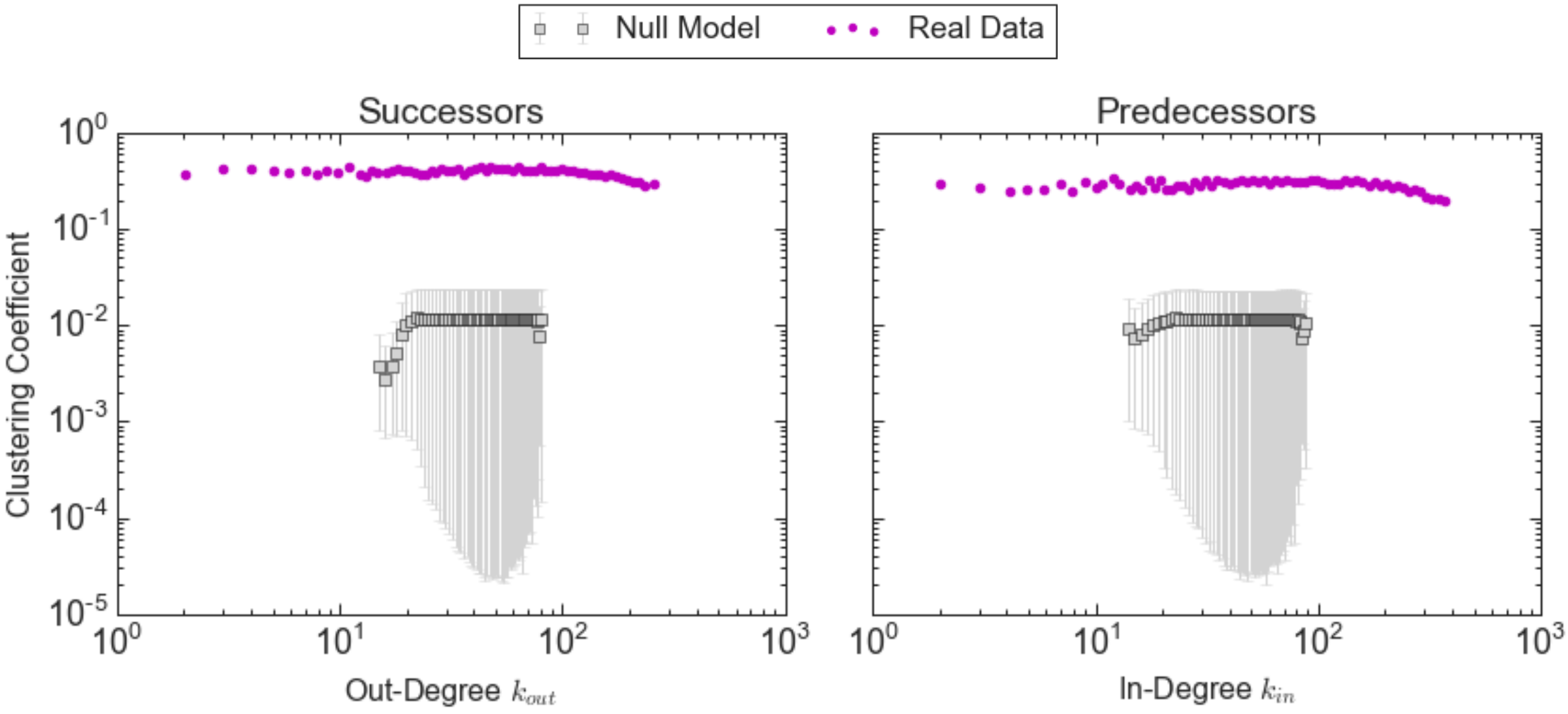
Real Data $\langle s_{in} \rangle = 95.369$

Null Model $\langle s_{in} \rangle = 90.244 \pm 0.053$

Degree Correlation



Clustering Coefficient



Real Data $\langle C_{out} \rangle = 0.297$ $\langle C_{in} \rangle = 0.173$

Null Model $\langle C_{out} \rangle = 0.012 \pm 8.0 \times 10^{-5}$ $\langle C_{in} \rangle = 0.012 \pm 7.7 \times 10^{-5}$

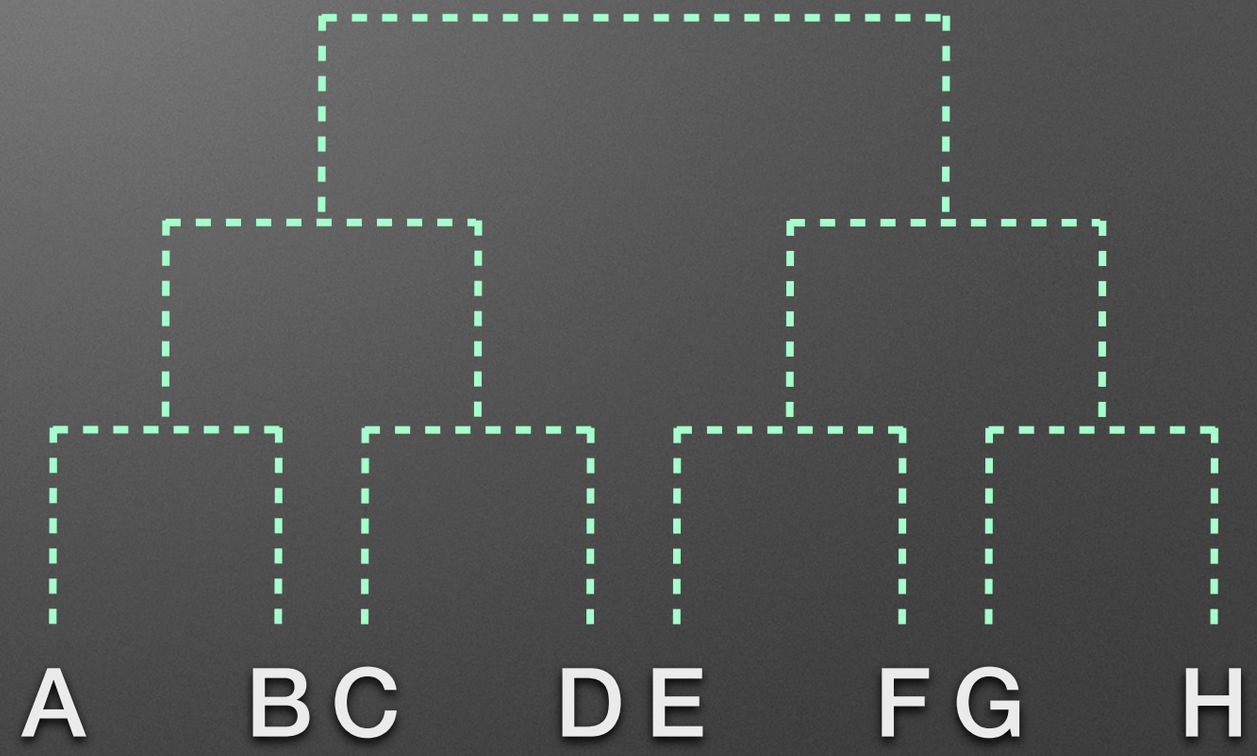
Components

	Number of strongly connected components
Real Data	841
Null Model	1

	Component size	Frquency	Average distance
Real Data	1	830	0.00
	2	10	1.00
	2368	1	3.15 (2.58 for null model)

Intrinsic Skill Modelling

- Assign each player a intrinsic skill parameter ψ following some distribution $\rho(\psi)$.
- Simulate tournaments given their initial draw from the data.
- The probability that a player wins a set is proportional to his intrinsic skill.



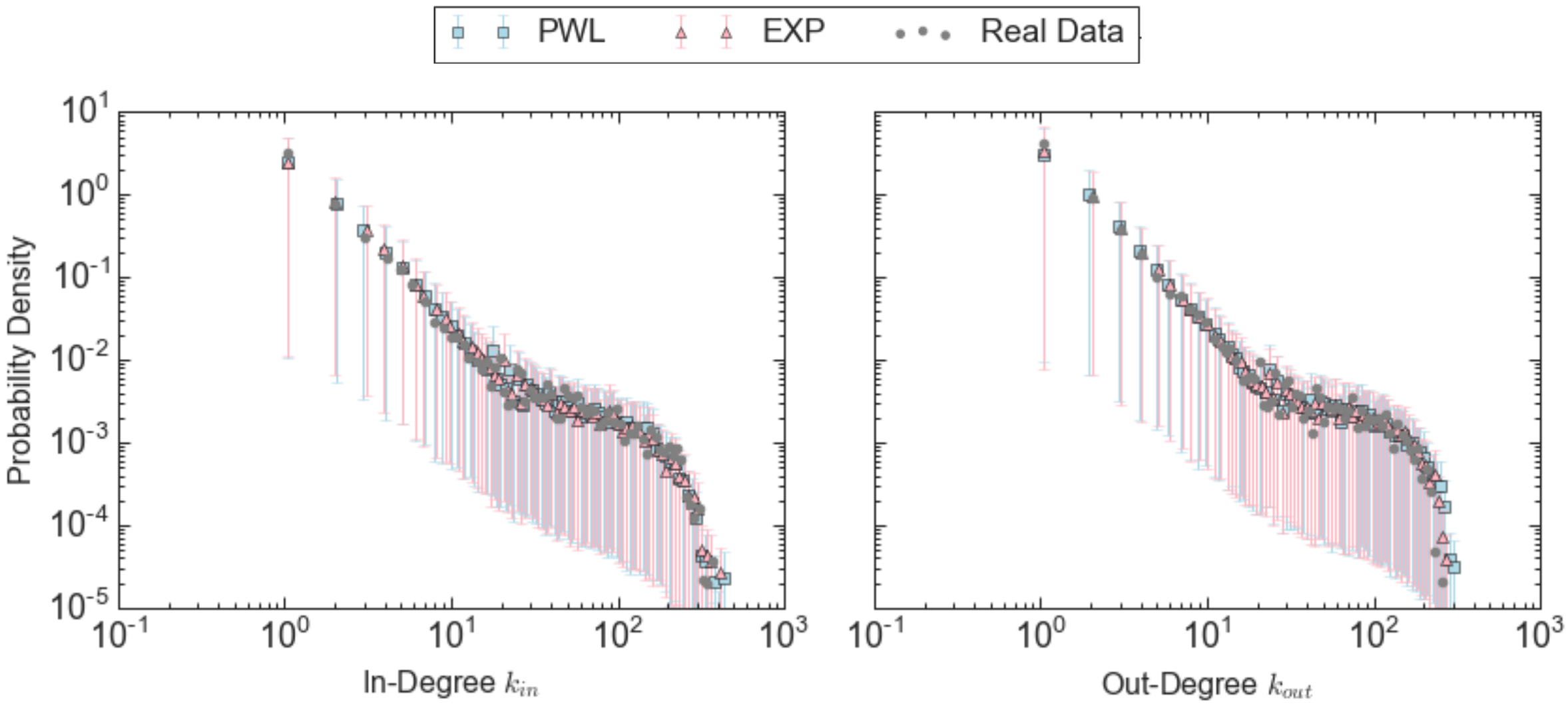
$$P\{A\ wins\} = \frac{\psi(A)}{\psi(A) + \psi(B)}$$

$$P\{B\ wins\} = \frac{\psi(B)}{\psi(A) + \psi(B)}$$

Does the intrinsic skill distribution have heavy-tail?

Power-law vs Exponential

Degree Distribution

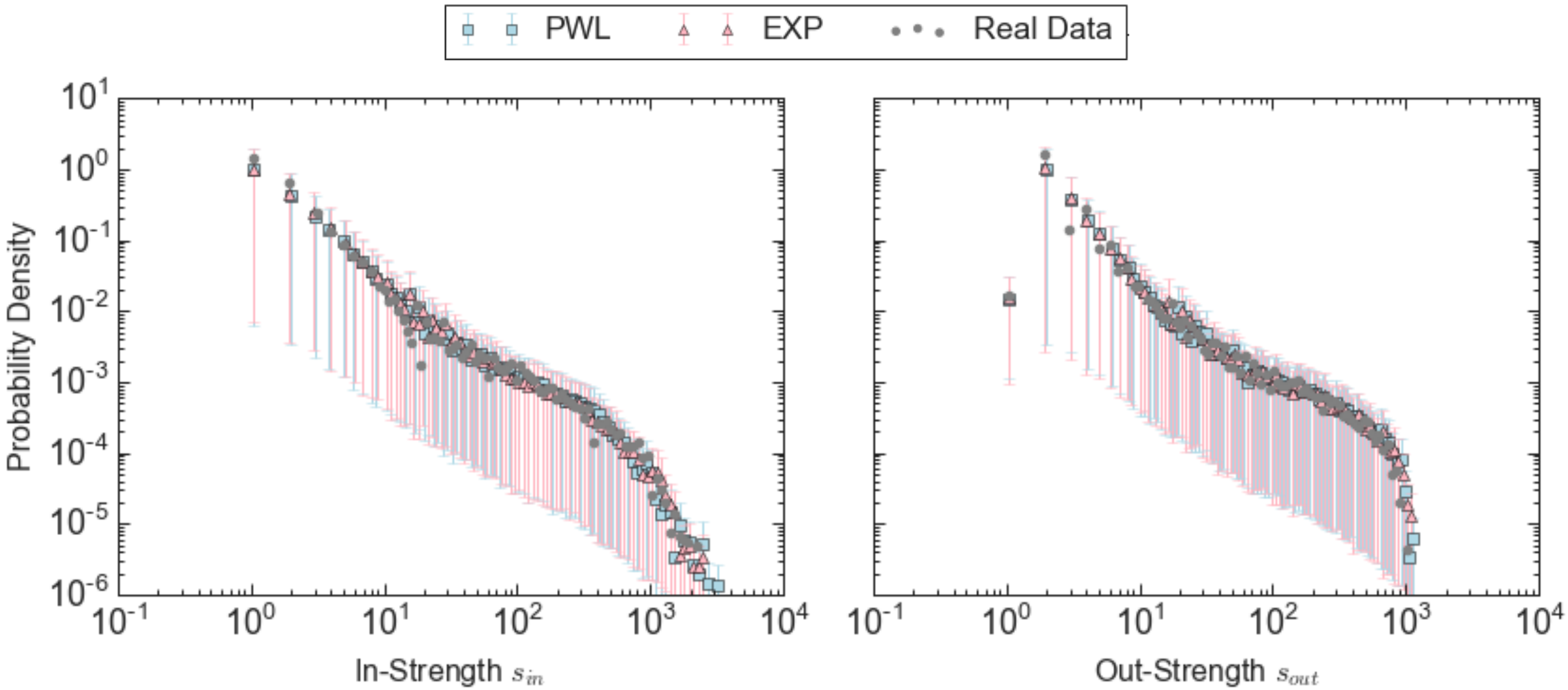


Real Data $\langle k_{in} \rangle = 33.040$

PWL $\langle k_{in} \rangle = 37.845 \pm 0.058$

EXP $\langle k_{in} \rangle = 36.452 \pm 0.044$

Strength Distribution

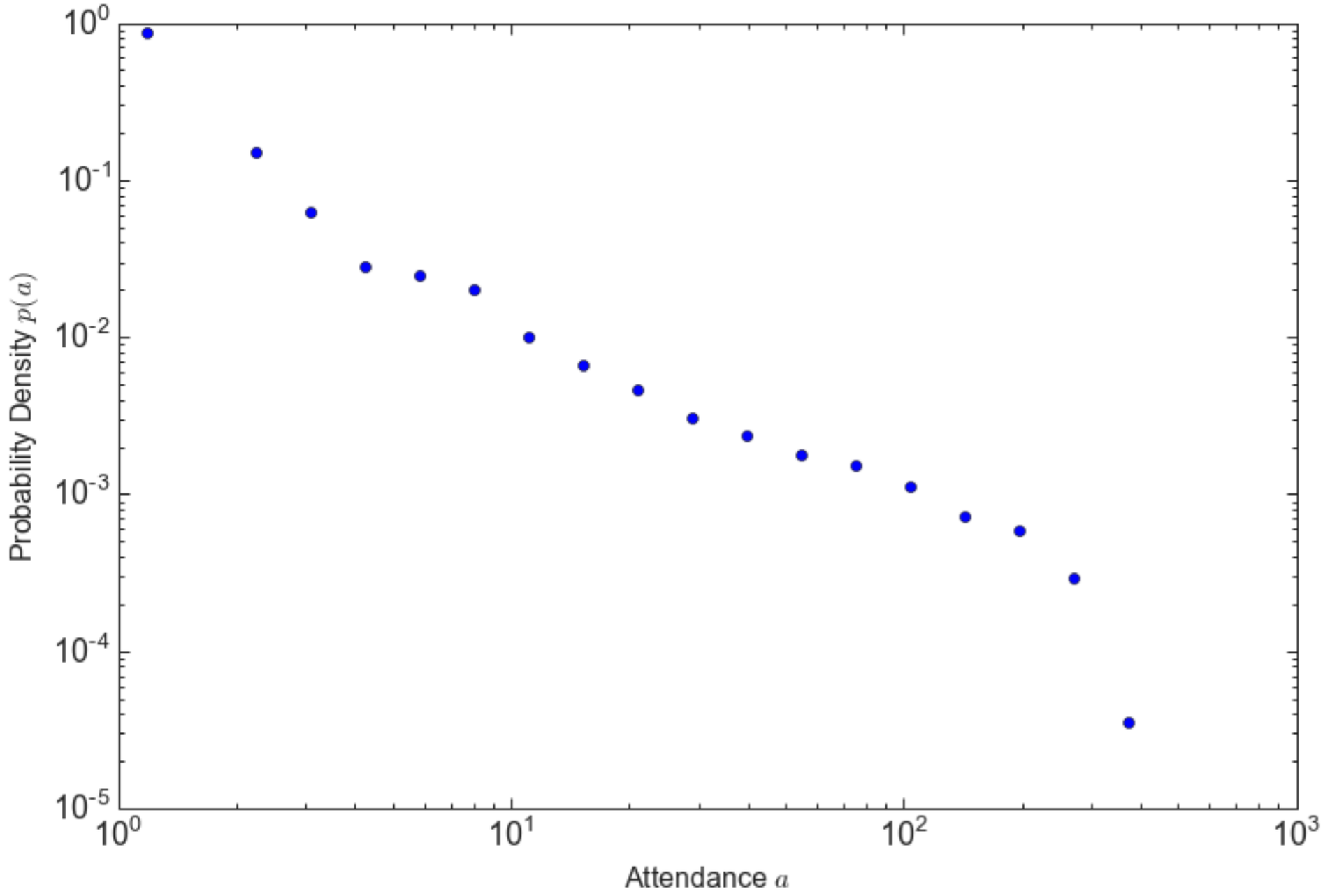


Real Data $\langle s_{in} \rangle = 95.369$

PWL $\langle s_{in} \rangle = 111.811 \pm 0.067$

EXP $\langle s_{in} \rangle = 110.295 \pm 0.061$

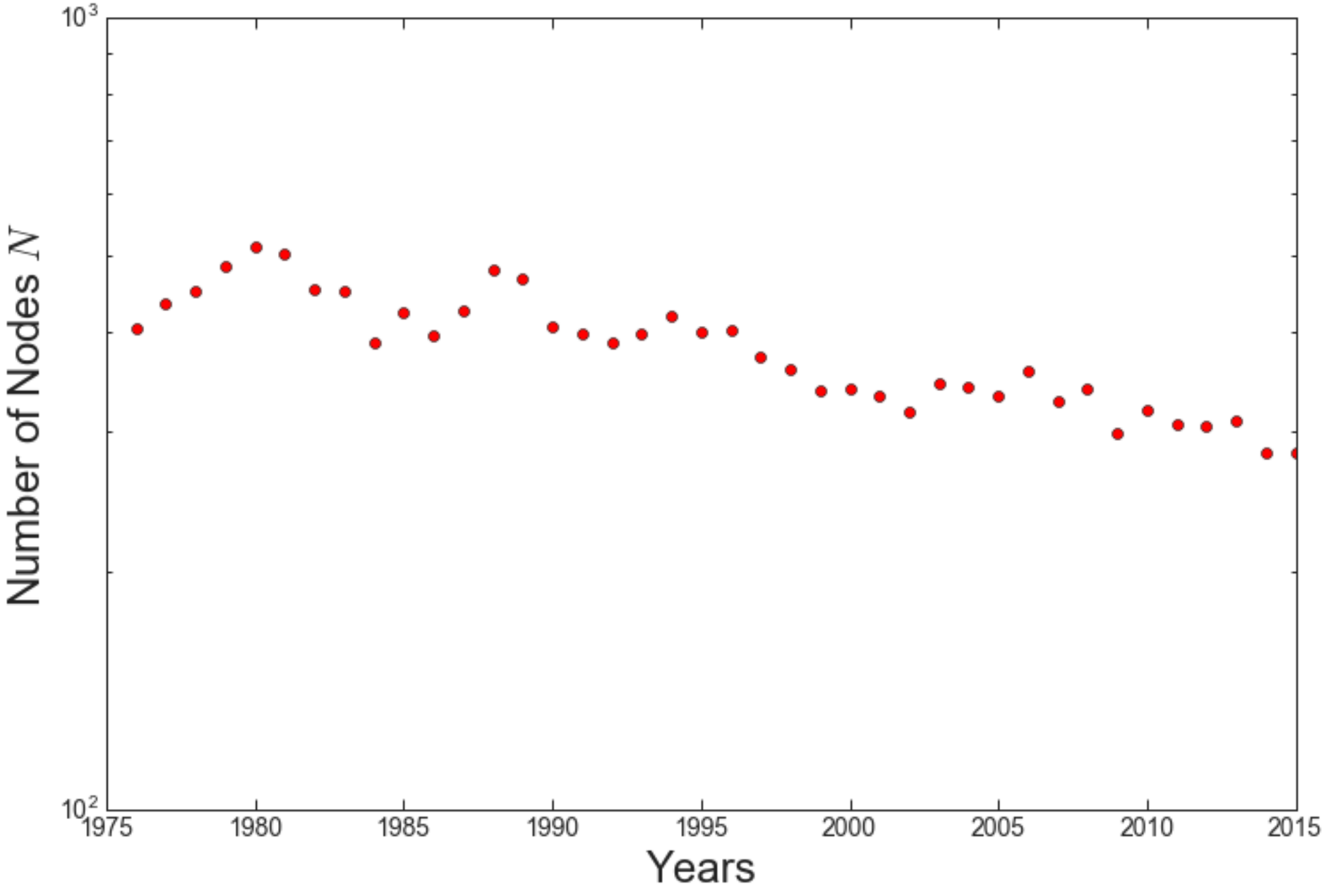
Tournament Attendance

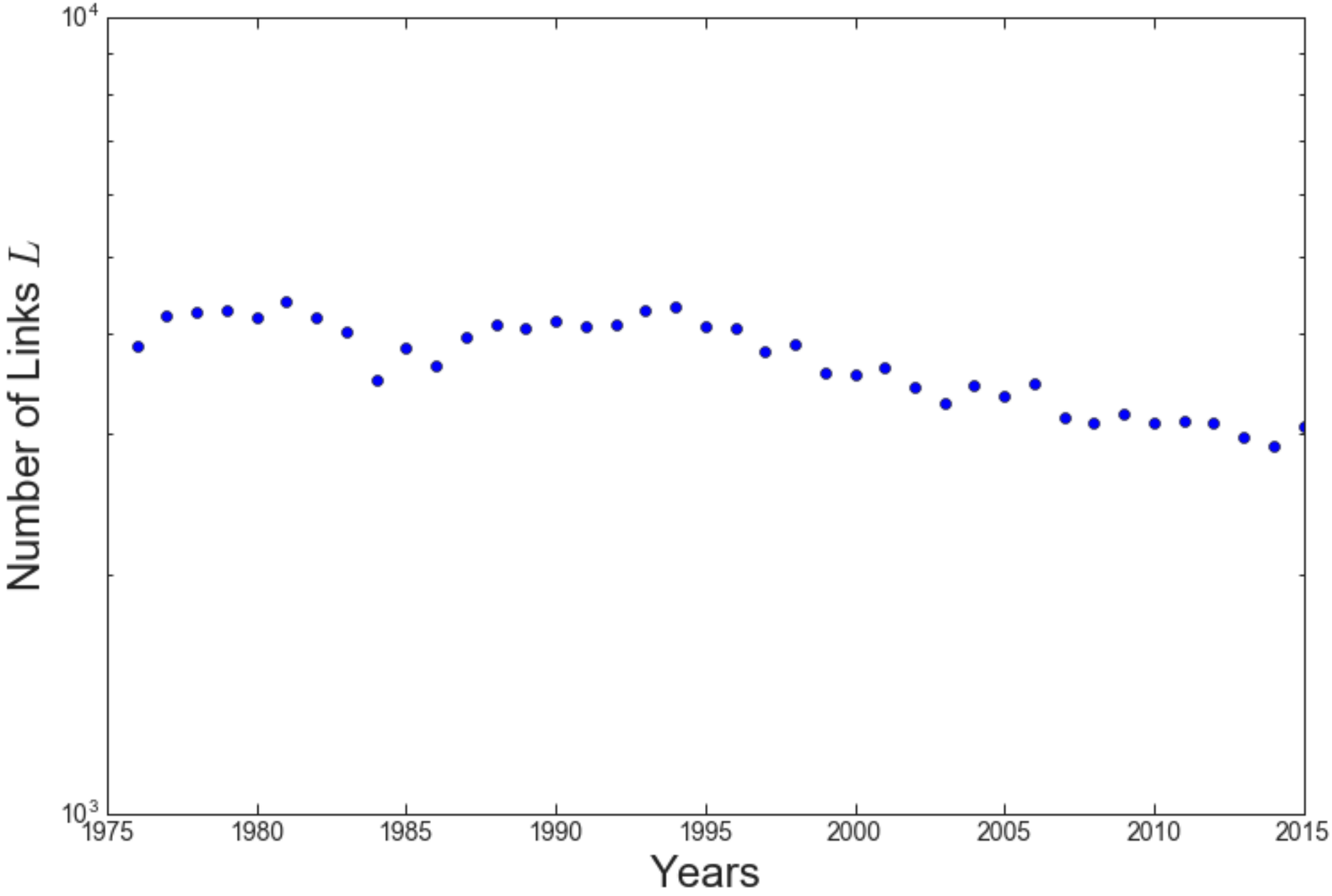


Conclusion

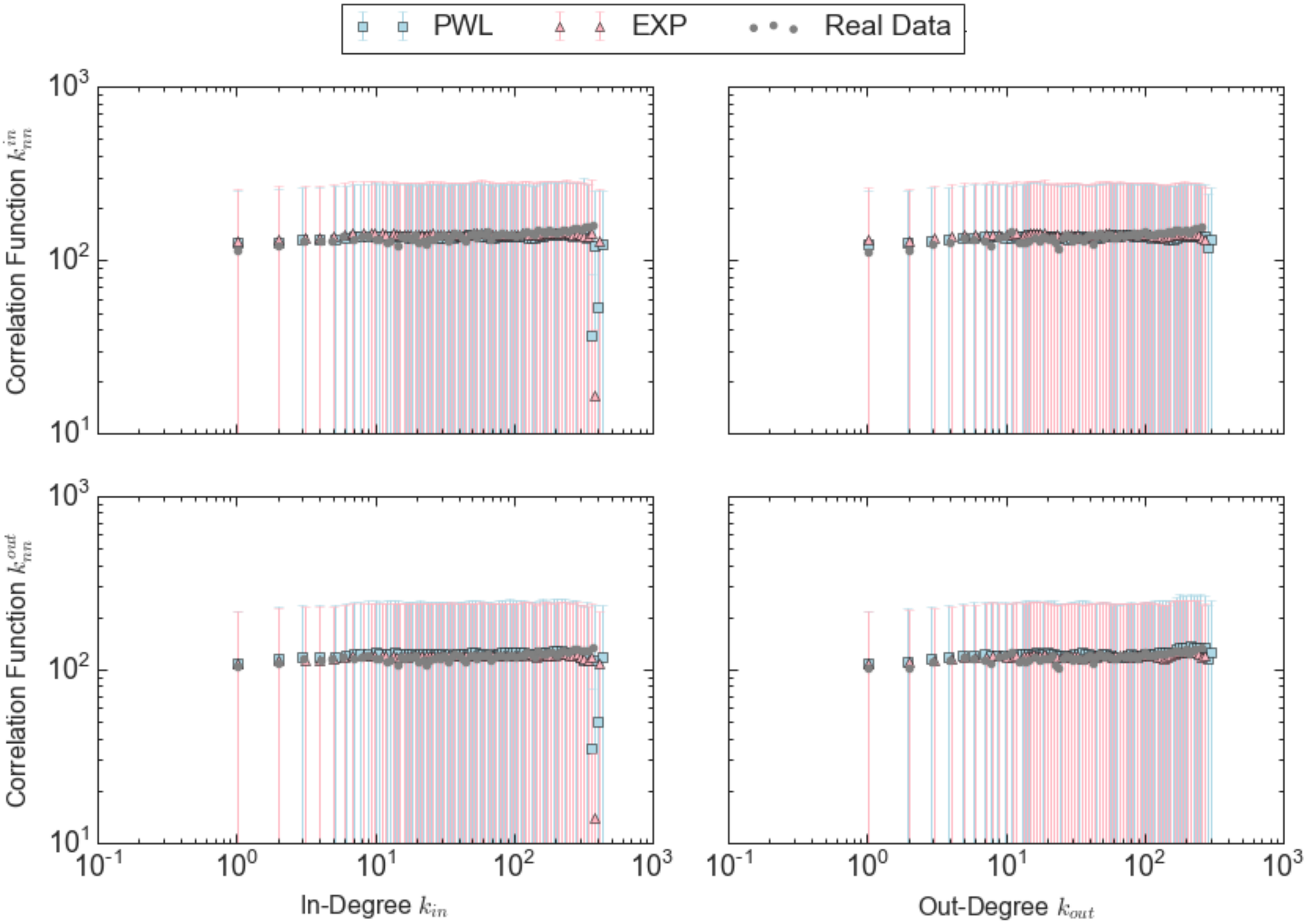
- The community structure reflects players of different generation.
- The tennis player network is neutral but highly clustered.
- The numbers of sets that players won own heterogeneity.

Appendix


10^3
10^2
Number of Nodes N
1975
1980
1985
1990
1995
2000
2005
2010
2015
Years


10^4
Number of Links L
10^3
1975
1980
1985
1990
1995
2000
2005
2010
2015
Years

Degree Correlation



Clustering Coefficient