Statistical Analysis of Networks

Projection and Weighted Networks

Learning Goals

- Understand *projection* of bipartite graphs to unipartite graphs.
- Examine *dichotomized* projections.
- Examine *summation weighted* and *Newman weighted* projections.

Projection

- The process by which we map the connectivity between modes to a single mode.
  - Example
    - Two-mode network is people in groups.
      - By projecting, we get:
        - One-mode network of people connected to people *by* groups.
        - One-mode network of groups connected *by* people.

Empirical Example

Journal of Contemporary Criminal Justice
2015, Vol. 31(3) 243–261

DOI: 10.1177/1043986214553380
ccj.sagepub.com
SAGE

Diffusion of Ideas and Technology: The Role of Networks in Influencing the Endorsement and Use of On-Officer Video Cameras

Jacob T. N. Young[1] and Justin T. Ready[1]

- https://journals.sagepub.com/doi/10.1177/1043986214553380

Empirical Example

Journal of Contemporary Criminal Justice
2015, Vol. 31(3) 243–261

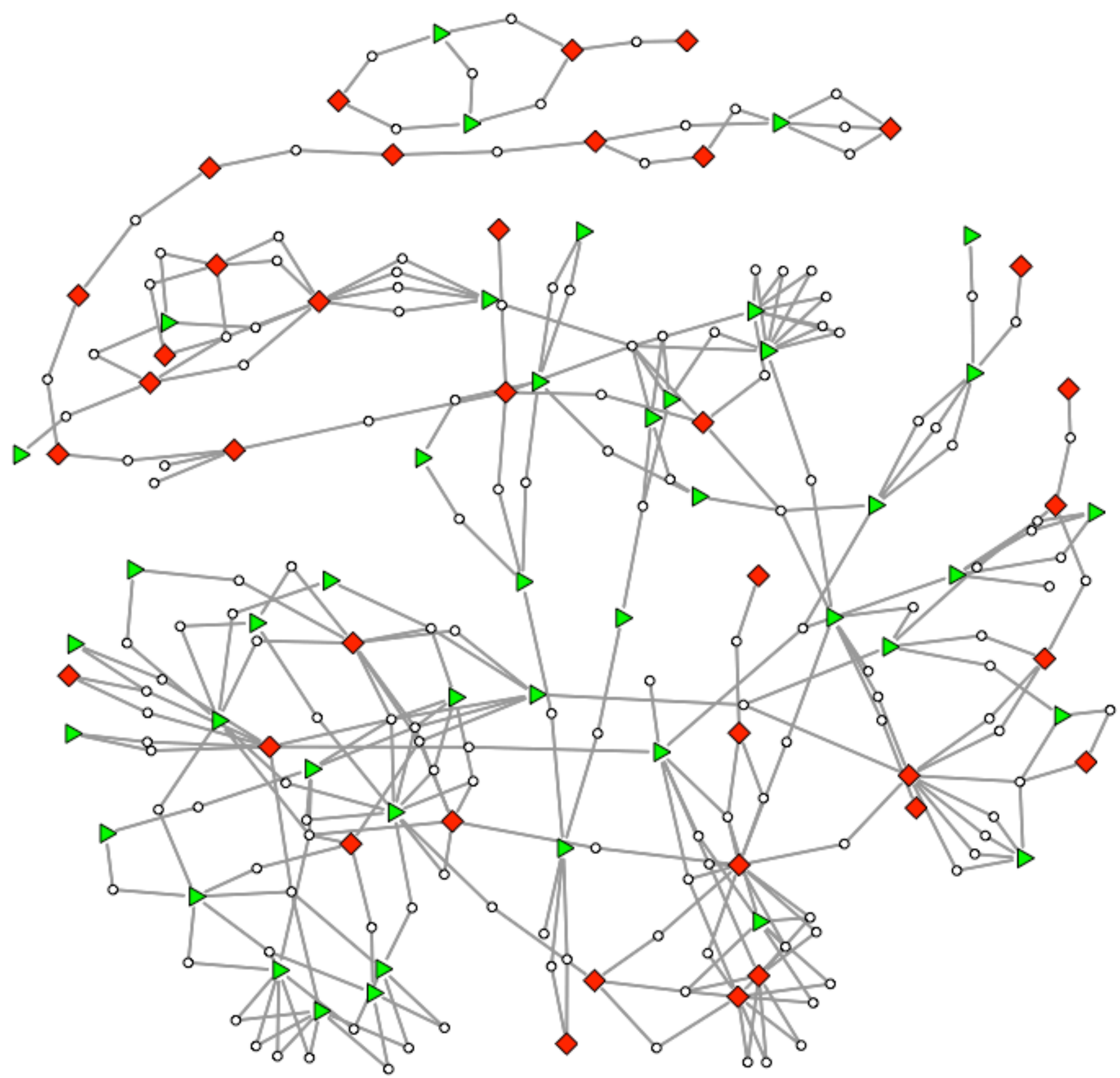
DOI: 10.1177/1043986214553380
ccj.sagepub.com
SAGE

Diffusion of Ideas and Technology: The Role of Networks in Influencing the Endorsement and Use of On-Officer Video Cameras

Jacob T. N. Young[1] and Justin T. Ready[1]

- <u>Questions</u>:
    - How do police officers “frame” body-worn cameras?
    - Is the meaning officers attribute to cameras created and transmitted in groups?

Bipartite Graph of Incidents and Officers by Treatment or Control Condition



Red/Square=Treatment Condition
Green/Triangle=Control Condition
White/Circle=Incidents

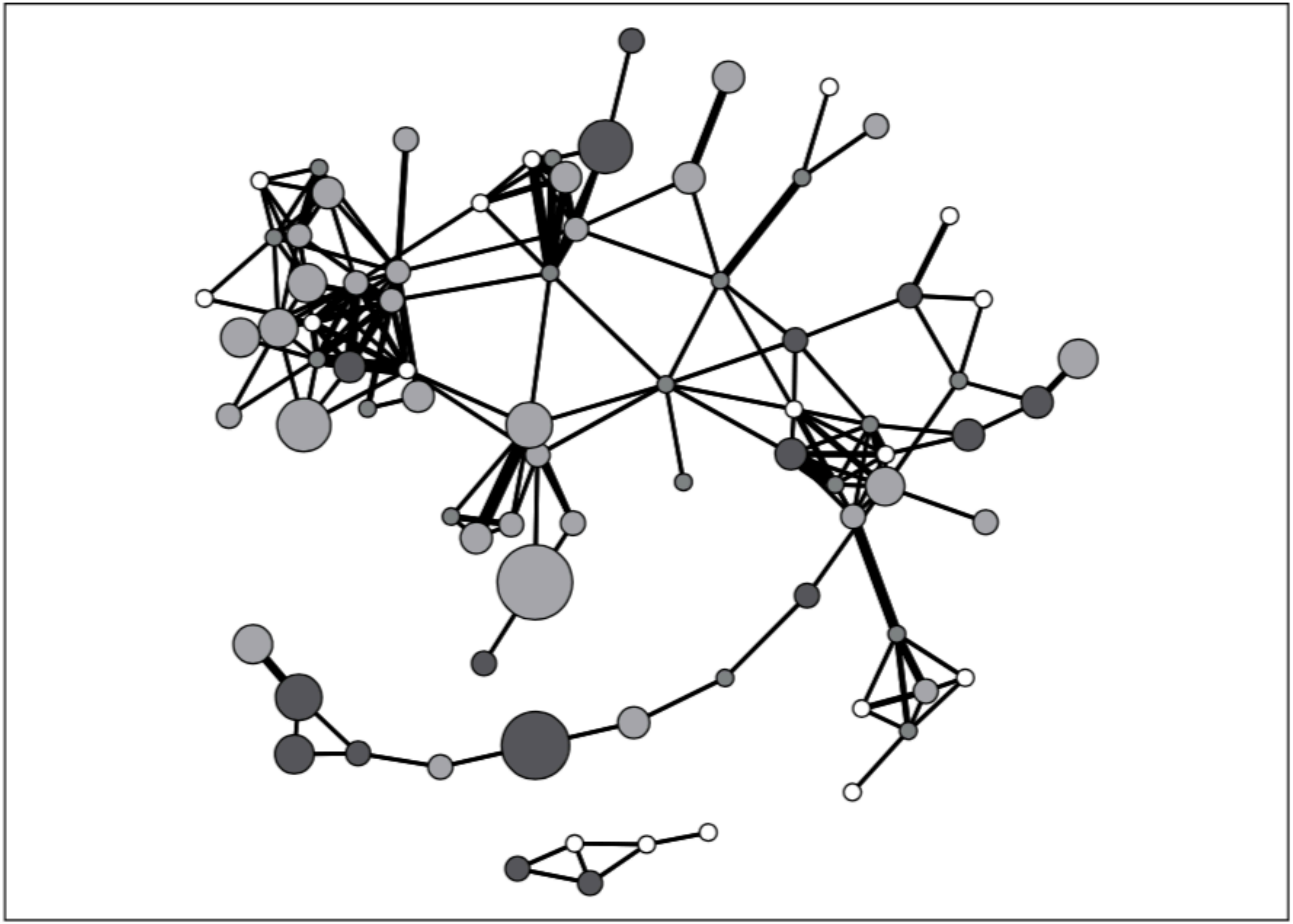


Figure 5. One-mode network of officers.

Note. Node size is proportional to change in legitimacy: darker = more negative; white = no change. Line size is proportional to number of shared incidents.

So where does this network come from?

Projection

- Breiger (1974)
  - We can build the adjacency matrix for each projected network through matrix algebra.
    - Specifically, multiplying an adjacency matrix by it's **transpose**.
      - The transpose of a matrix simply reverses the columns and rows:
        - $A^{T}_{ij} = A_{ji}$

Projection

- Breiger (1974)
    - The two-mode, *NxM*, adjacency matrix, when multiplied by it's **transpose**, produces either:
        - An *MxM* matrix (ties among *M* nodes via *N*).
        - An *NxN* matrix (ties among *N* nodes via *M*).

Transposition

Matrix **A**

	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

order is 6x5

Matrix $\mathbf{A}^{\mathbf{T}}$

	A	B	C	D	E	F
1	1	1	1	0	0	0
2	1	0	1	1	0	0
3	0	0	0	1	1	1
4	0	0	0	1	0	0
5	0	0	0	1	0	1

order is 5x6

Projection

- Matrix Multiplication Rules
  - To multiply two matrices, the number of columns in the first matrix must match the number of rows in the second matrix.
    - Example: 5x6 X 6x5 works, but not 5x6 X 5x6
  - The **product** matrix has the number of rows equal to the first matrix and the number of columns equal to the second matrix.
    - Example: 5x6 X 6x5 = 5x5

Projection

- Product Matrix
  - The product matrix is the projected graph.
    - Recall that there are two:
      - $\mathbf{A} \times \mathbf{A}^t$ (the "people" matrix **P**)
      - And the $\mathbf{A}^t \times \mathbf{A}$ (the "group" matrix **G**)
        - *What does each one mean?*

Matrix Multiplication

Matrix **A**

	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

order is 6x5

Matrix Multiplication

Matrix **A**

	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

order is 6x5

X

Matrix $\mathbf{A}^\mathbf{T}$

	A	B	C	D	E	F
1	1	1	1	0	0	0
2	1	0	1	1	0	0
3	0	0	0	1	1	1
4	0	0	0	1	0	0
5	0	0	0	1	0	1

order is 5x6

Matrix Multiplication

Matrix **A**

	1	2	3	4	5
A	1	1	0	0	0
B	1	0	0	0	0
C	1	1	0	0	0
D	0	1	1	1	1
E	0	0	1	0	0
F	0	0	1	0	1

order is 6x5

X

Matrix $\mathbf{A}^{\mathbf{T}}$

	A	B	C	D	E	F
1	1	1	1	0	0	0
2	1	0	1	1	0	0
3	0	0	0	1	1	1
4	0	0	0	1	0	0
5	0	0	0	1	0	1

order is 5x6

The product matrix is
6x6

Projection by Multiplication

$$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

6x5 X 5x6 = 6x6

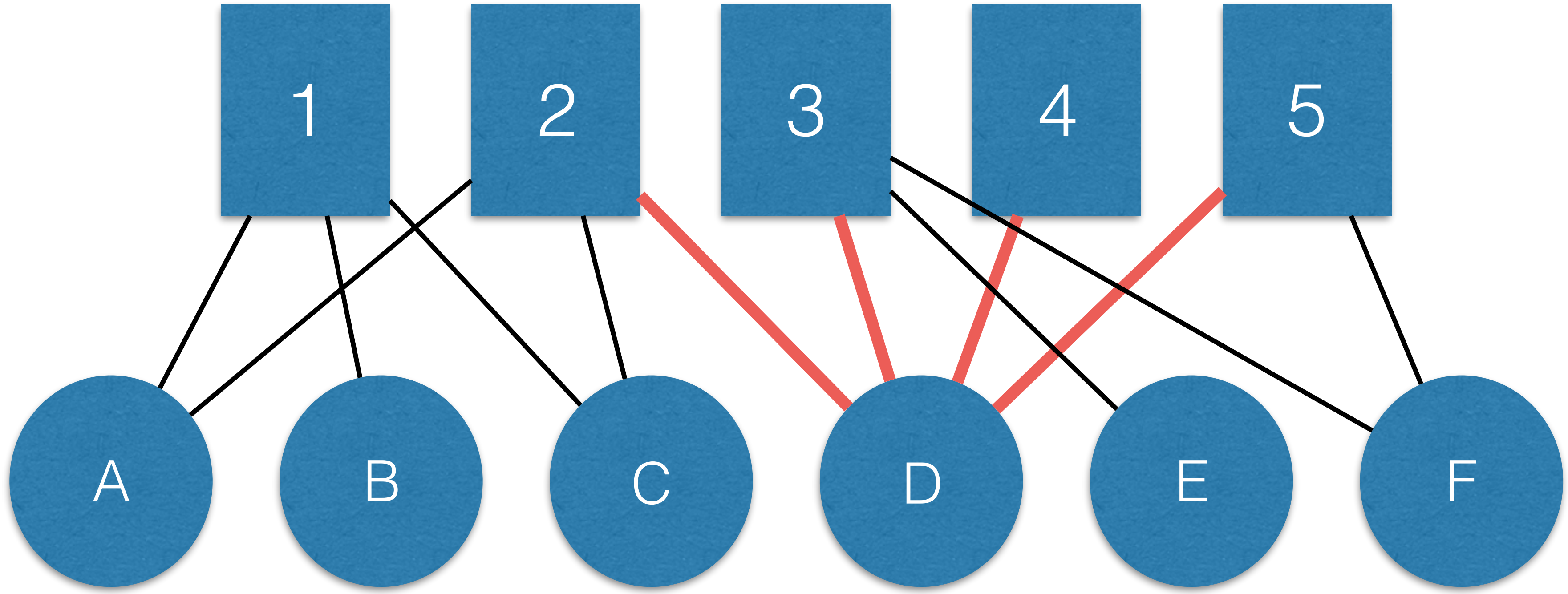
Projection by Multiplication

We want to know how people are connected by groups (i.e. the rows of our two-mode adjacency matrix)

$$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

6x5 X 5x6 = 6x6



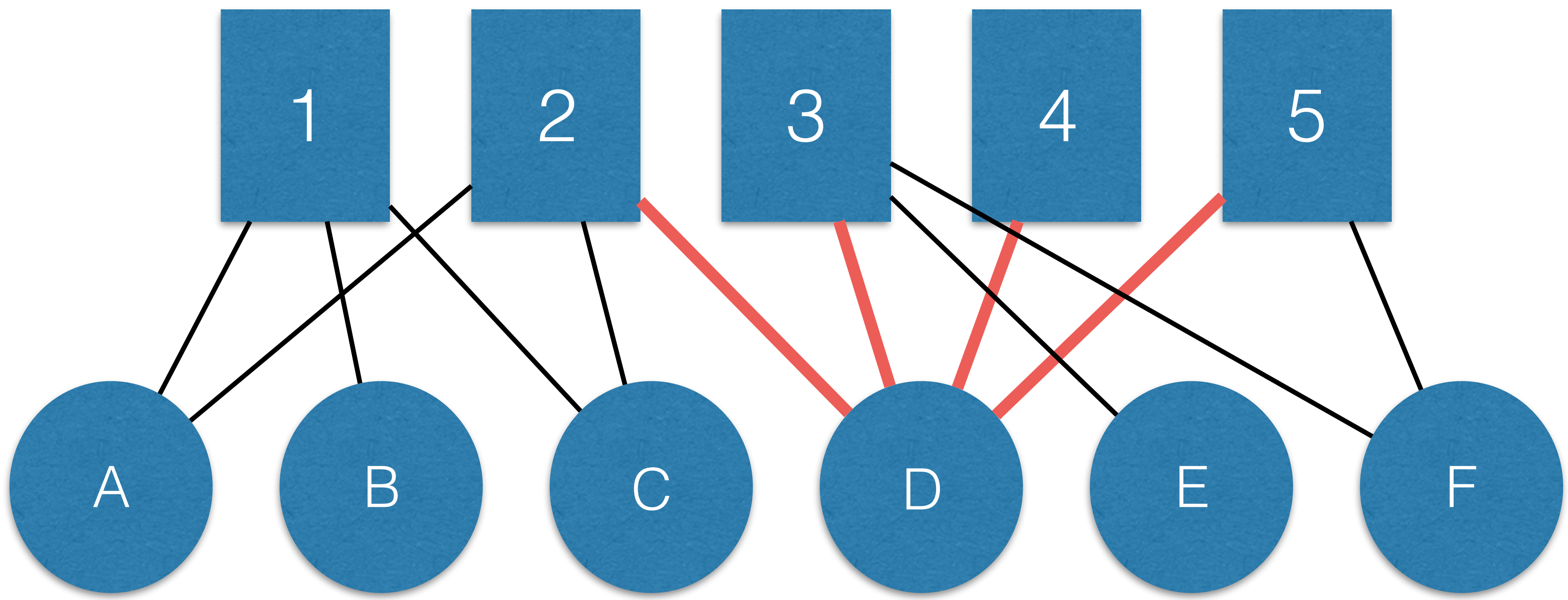
$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

The diagonal is the count of ties **the person** has with two-mode vertices

For example, D is in 4 groups

6x5 X 5x6 = 6x6

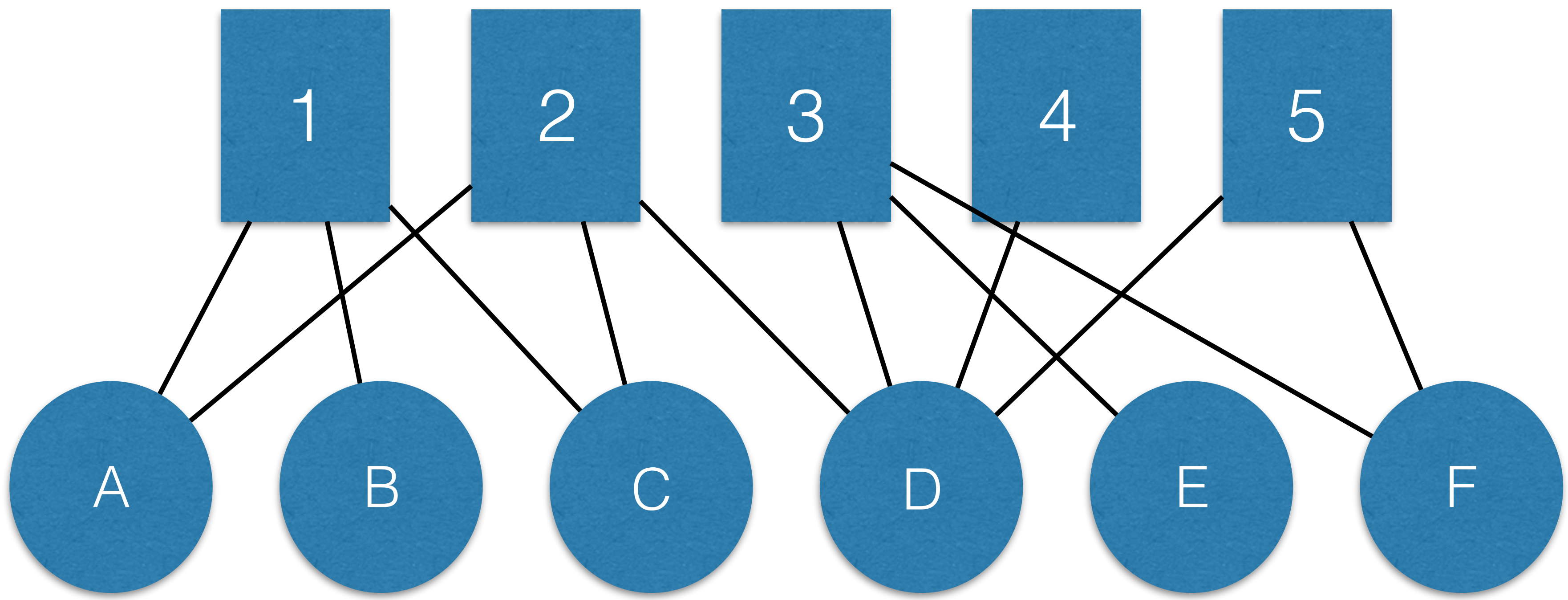


$$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

What statistic does the diagonal give us?

6x5 X 5x6 = 6x6

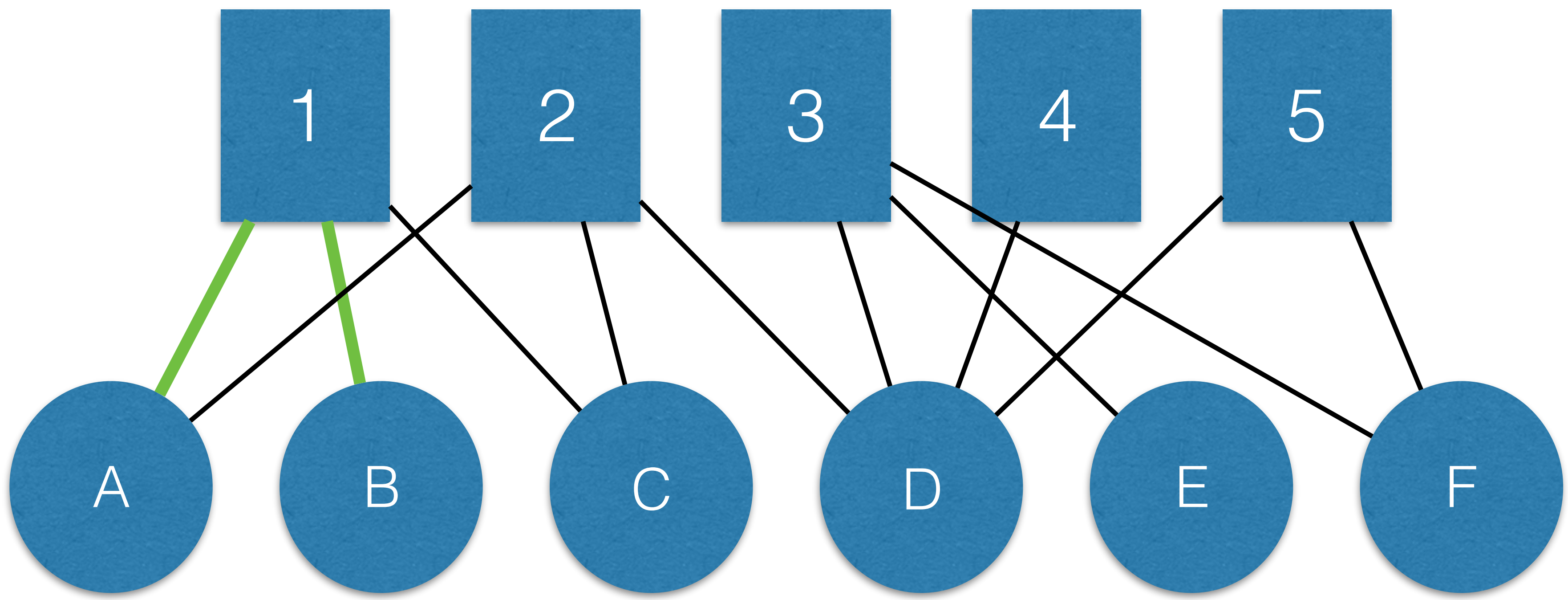


$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

Note, that the projection forces the product matrix to be symmetric (i.e. undirected graph)

6x5 X 5x6 = 6x6



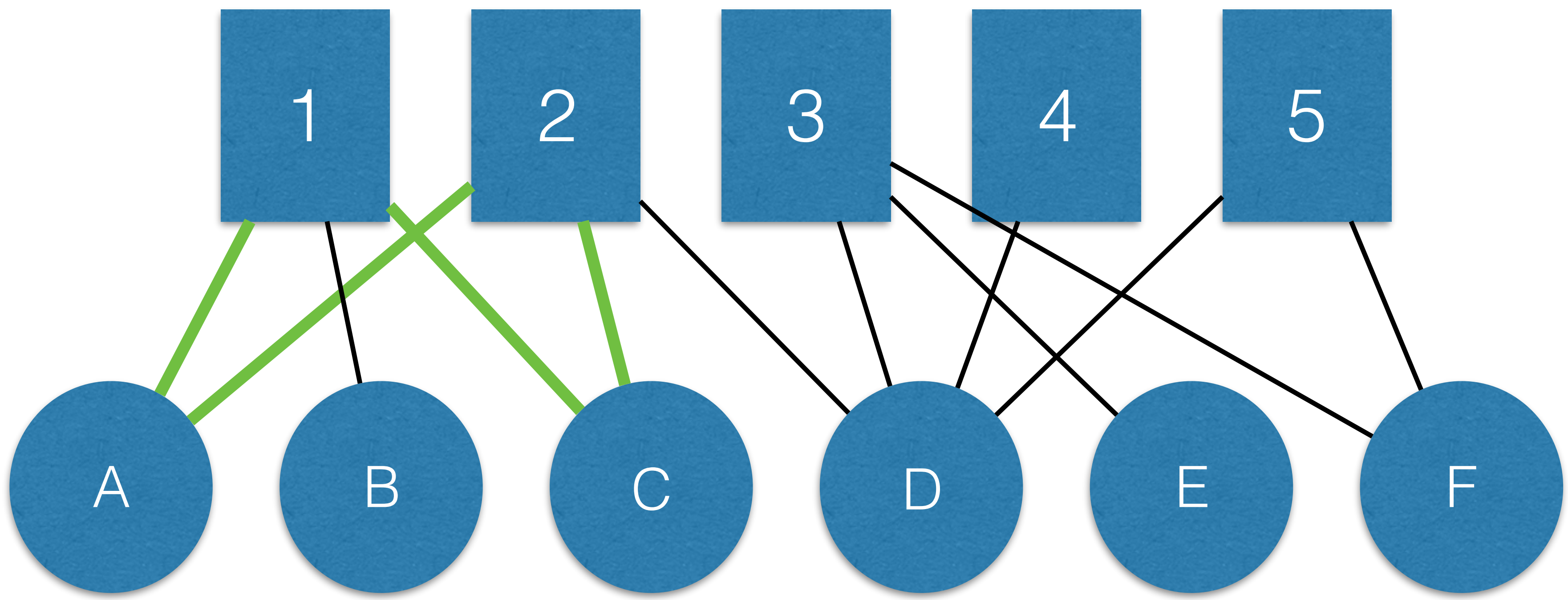
$$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

6x5 X 5x6 = 6x6

The off-diagonal entries indicate the number of ways that vertices in the first mode are connected by vertices in the second mode

A and B are linked through a single vertex, 1



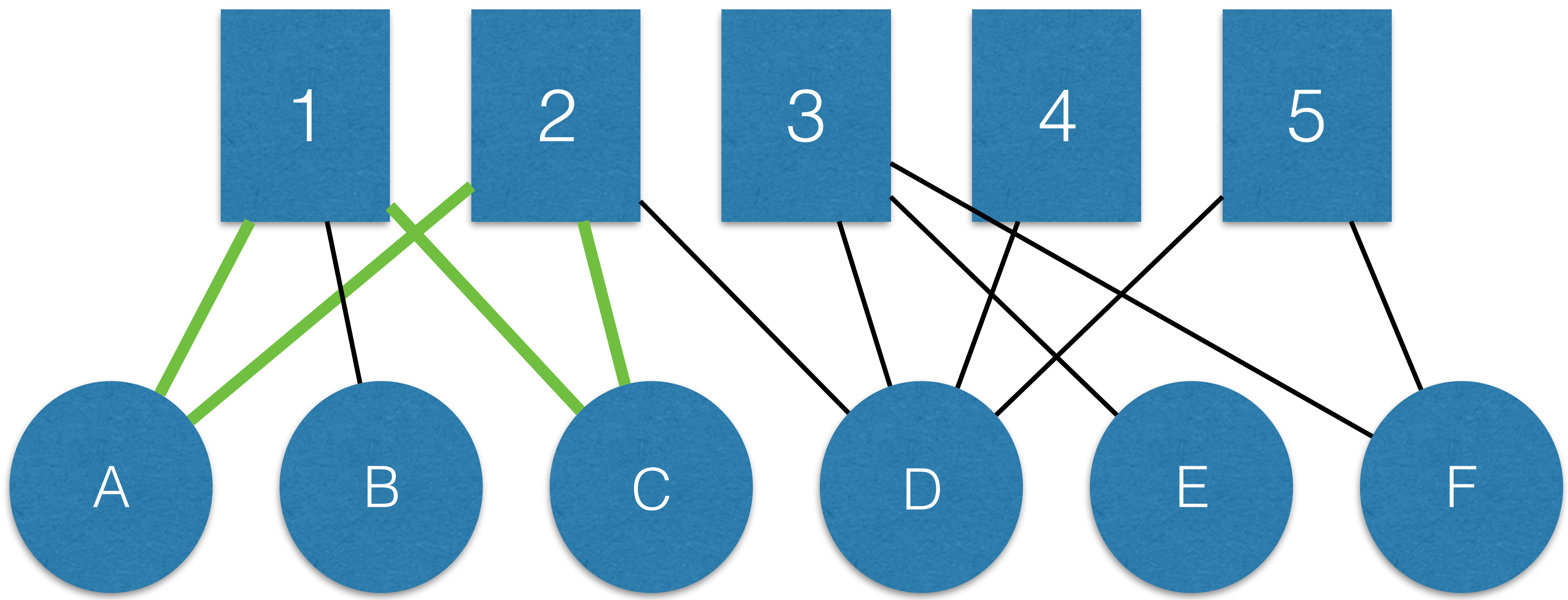
$$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

6x5 X 5x6 = 6x6

A and C are linked through two vertices, 1 and 2

So, if these are groups, A and C are members of 2 of the same groups



$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

NOTE: these are counts of shared vertices, not edge counts

6x5 X 5x6 = 6x6

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

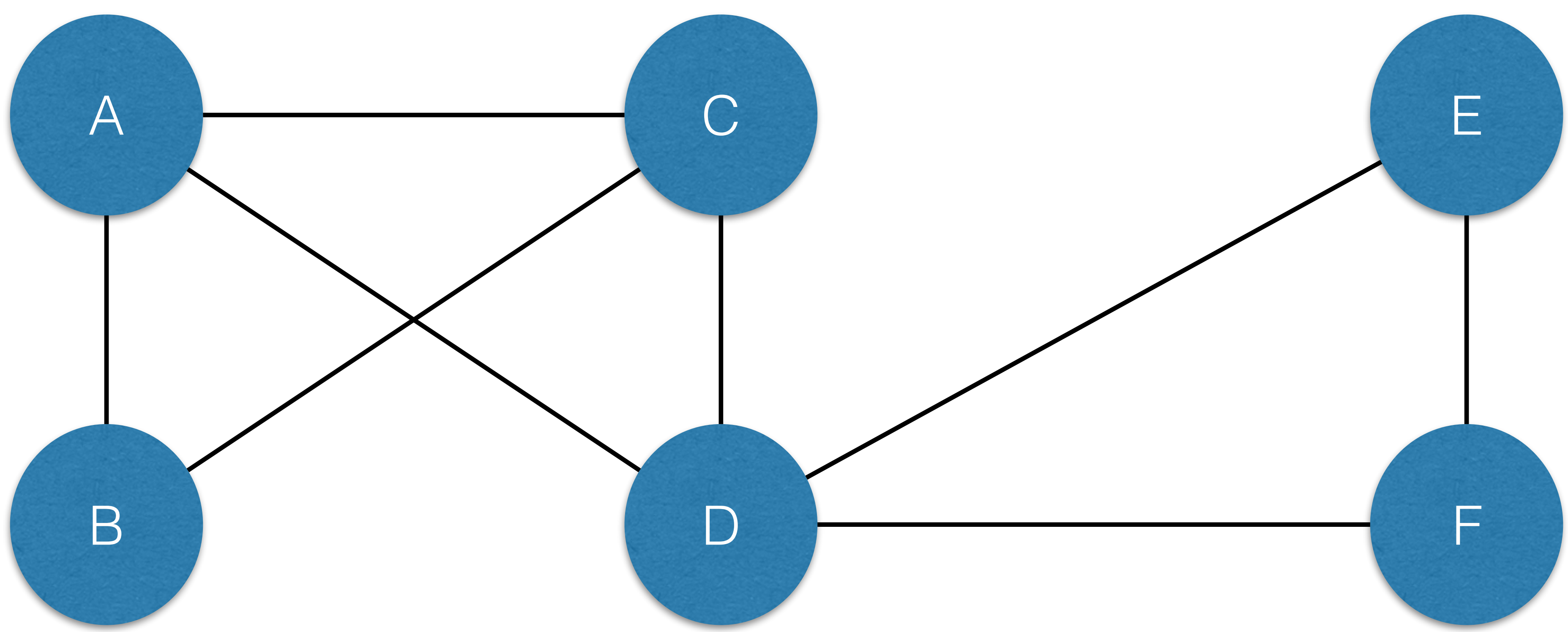
	A	B	C	D	E	F
A	0	1	1	1	0	0
B	1	0	1	0	0	0
C	1	1	0	1	0	0
D	1	0	1	0	1	1
E	0	0	0	1	0	1
F	0	0	0	1	1	0

If we treat any tie greater than 0 as binary, *called* ***dichotomizing***, and recode the diagonal as 0, we get an undirected, one-mode network

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

	A	B	C	D	E	F
A	0	1	1	1	0	0
B	1	0	1	0	0	0
C	1	1	0	1	0	0
D	1	0	1	0	1	1
E	0	0	0	1	0	1
F	0	0	0	1	1	0

If we treat any tie greater than 0 as binary, *called* ***dichotomizing***, and recode the diagonal as 0, we get an undirected, one-mode network



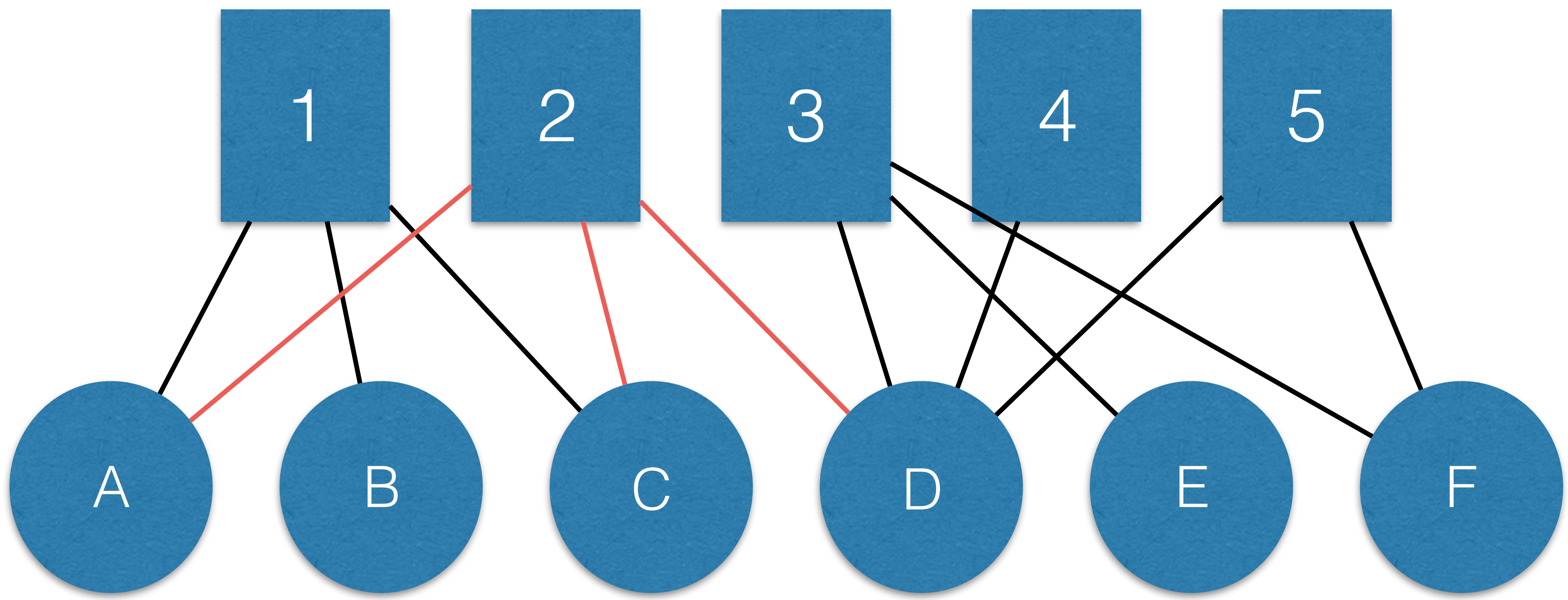
Projection by Multiplication

We want to know how groups are connected by people
(i.e. the columns of our two-mode adjacency matrix)

$$\mathbf{A^T} \times \mathbf{A} = \mathbf{G}$$

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

5x6 X 6x5 = 5x5



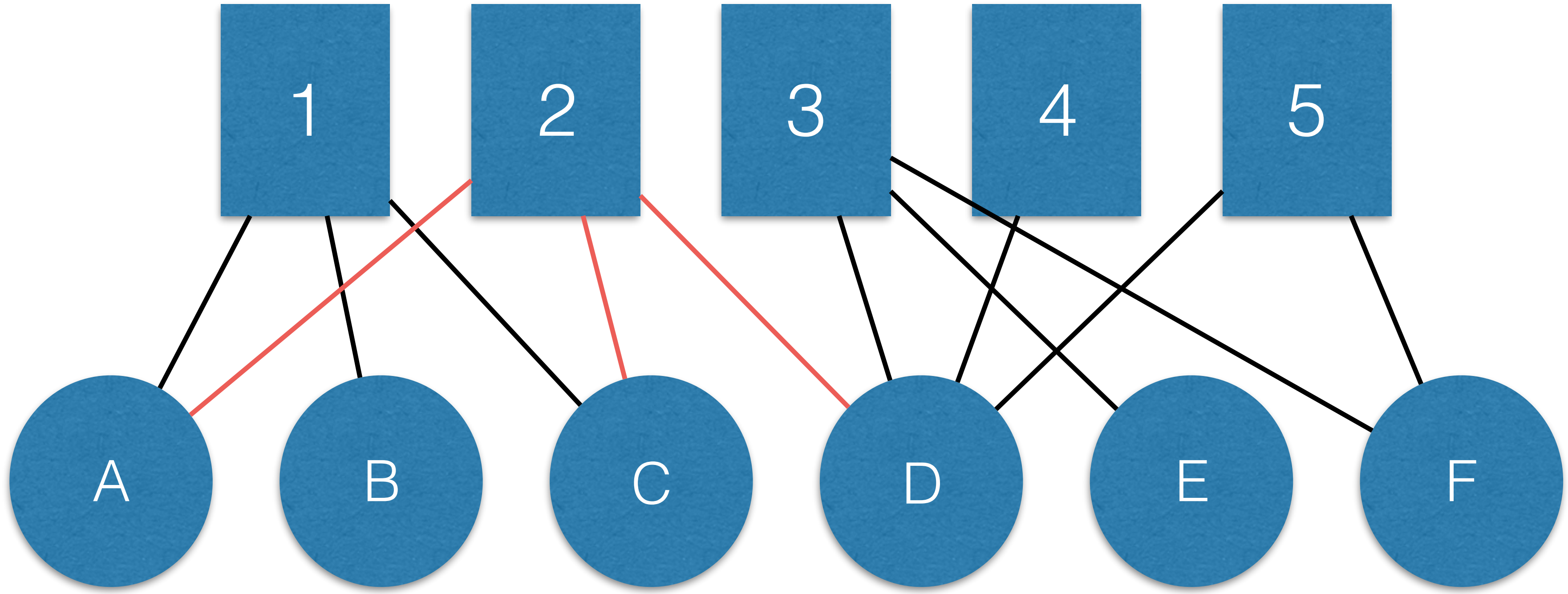
$$\mathbf{A^T} \times \mathbf{A} = \mathbf{G}$$

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

5x6 X 6x5 = 5x5

The diagonal is the count of ties the **group** has with two-mode vertices

For example, 2 has 3 people

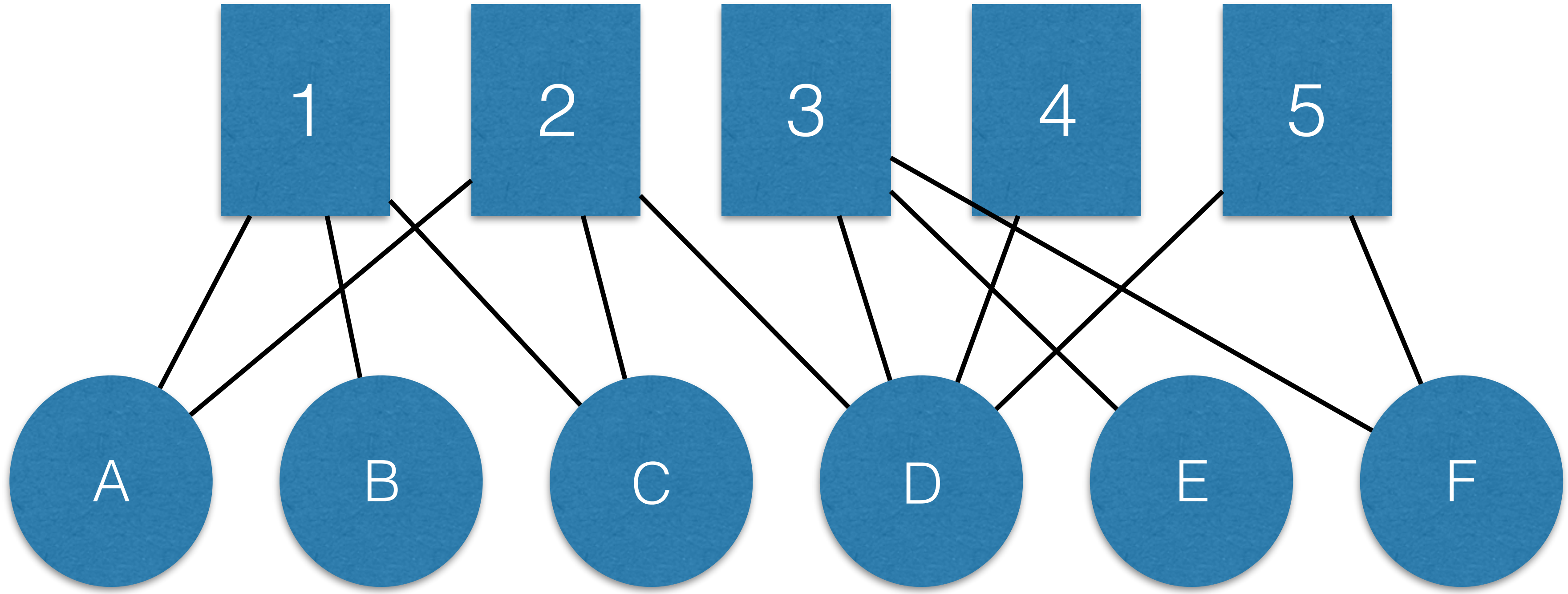


$\mathbf{A}^\mathbf{T} \times \mathbf{A} = \mathbf{G}$

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

What statistic does the diagonal give us?

5x6 X 6x5 = 5x5

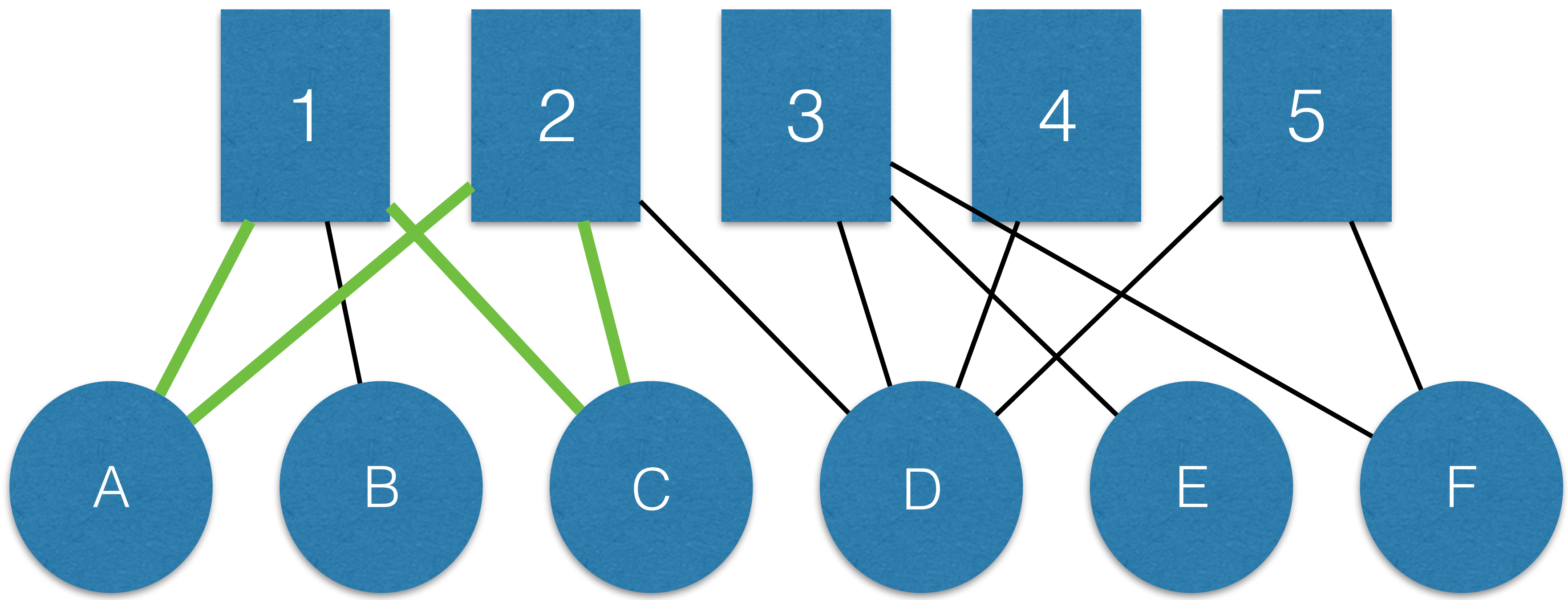


$\mathbf{A^T} \times \mathbf{A} = \mathbf{G}$

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

5x6 X 6x5 = 5x5

Note, that the projection forces the product matrix to be symmetric (i.e. undirected graph)



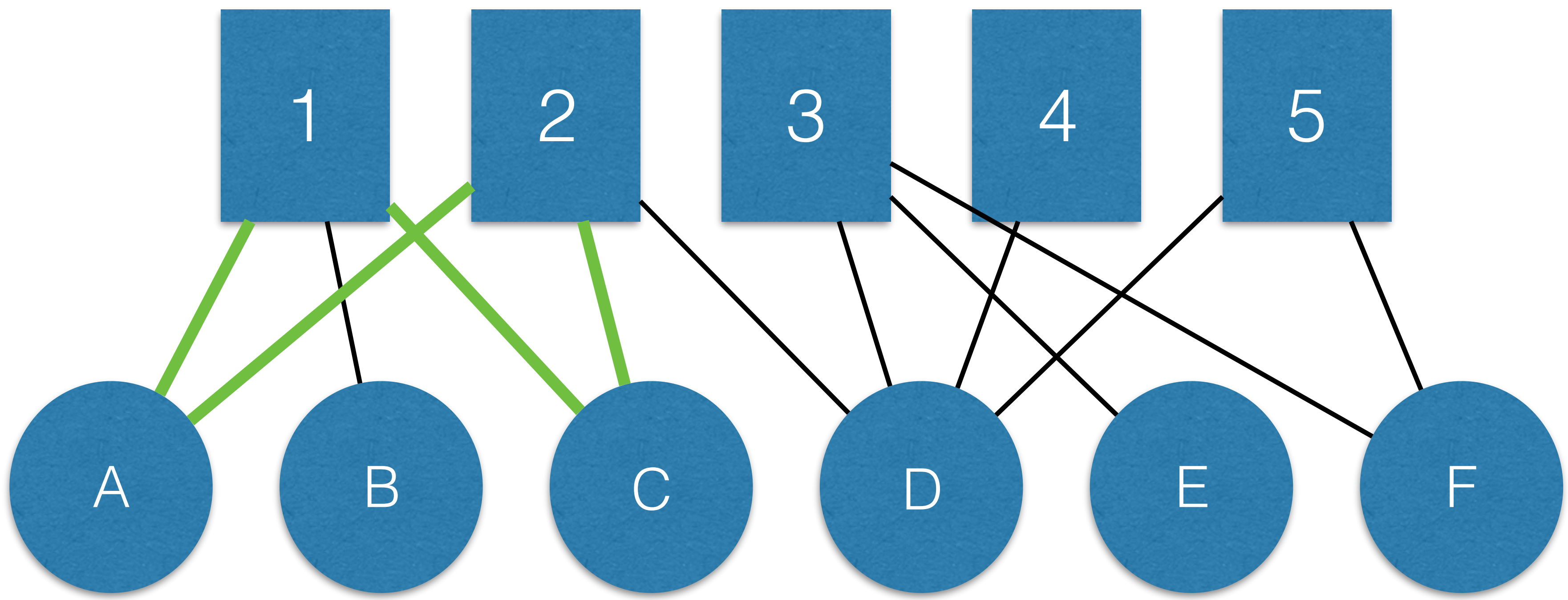
$$\mathbf{A}^{\mathbf{T}} \times \mathbf{A} = \mathbf{G}$$

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

5x6 X 6x5 = 5x5

The off-diagonal entries indicate the number of ways that vertices in the second mode are connected by vertices in the first mode

1 and 2 are connected by 2 vertices, A and C



$$\mathbf{A}^{\mathbf{T}} \times \mathbf{A} = \mathbf{G}$$

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

$5\text{x}6 \text{ X } 6\text{x}5 = 5\text{x}5$

NOTE: these are counts of shared vertices, not edge counts

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

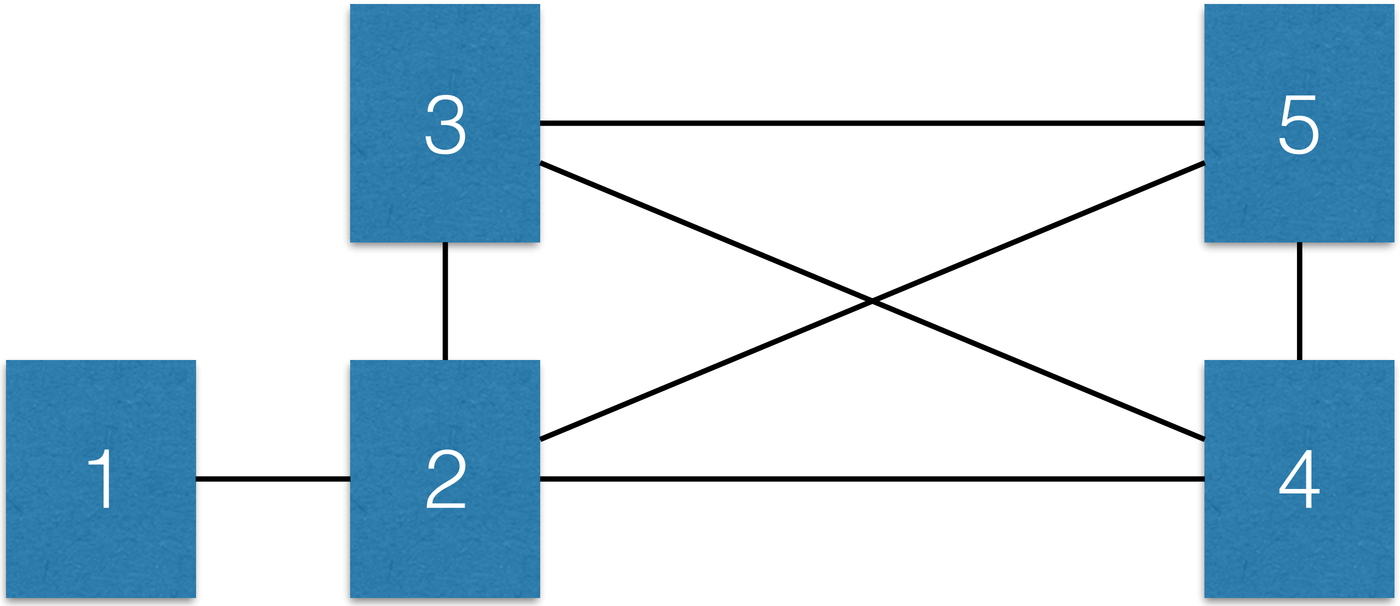
	1	2	3	4	5
1	0	1	0	0	0
2	1	0	1	1	1
3	0	1	0	1	1
4	0	1	1	0	1
5	0	1	1	1	0

If we treat any tie greater than 0 as binary, *called* ***dichotomizing***, and recode the diagonal as 0, we get an undirected, one-mode network

	1	2	3	4	5
1	3	2	0	0	0
2	2	3	1	1	1
3	0	1	3	1	2
4	0	1	1	1	1
5	0	1	2	1	2

	1	2	3	4	5
1	0	1	0	0	0
2	1	0	1	1	1
3	0	1	0	1	1
4	0	1	1	0	1
5	0	1	1	1	0

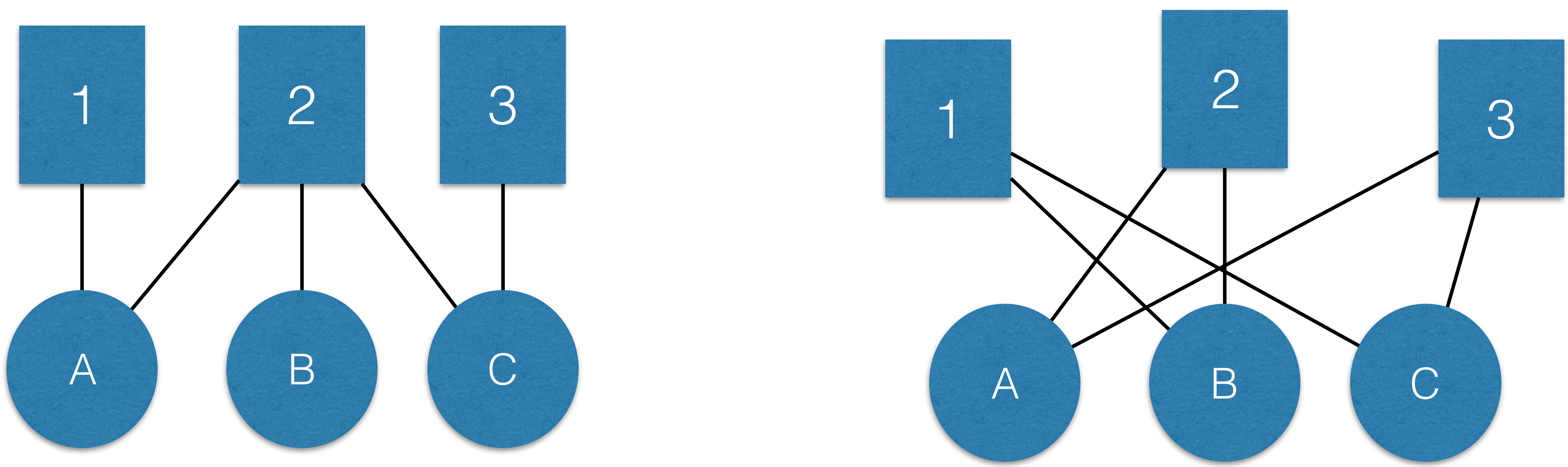
If we treat any tie greater than 0 as binary, *called* ***dichotomizing***, and recode the diagonal as 0, we get an undirected, one-mode network



Projection

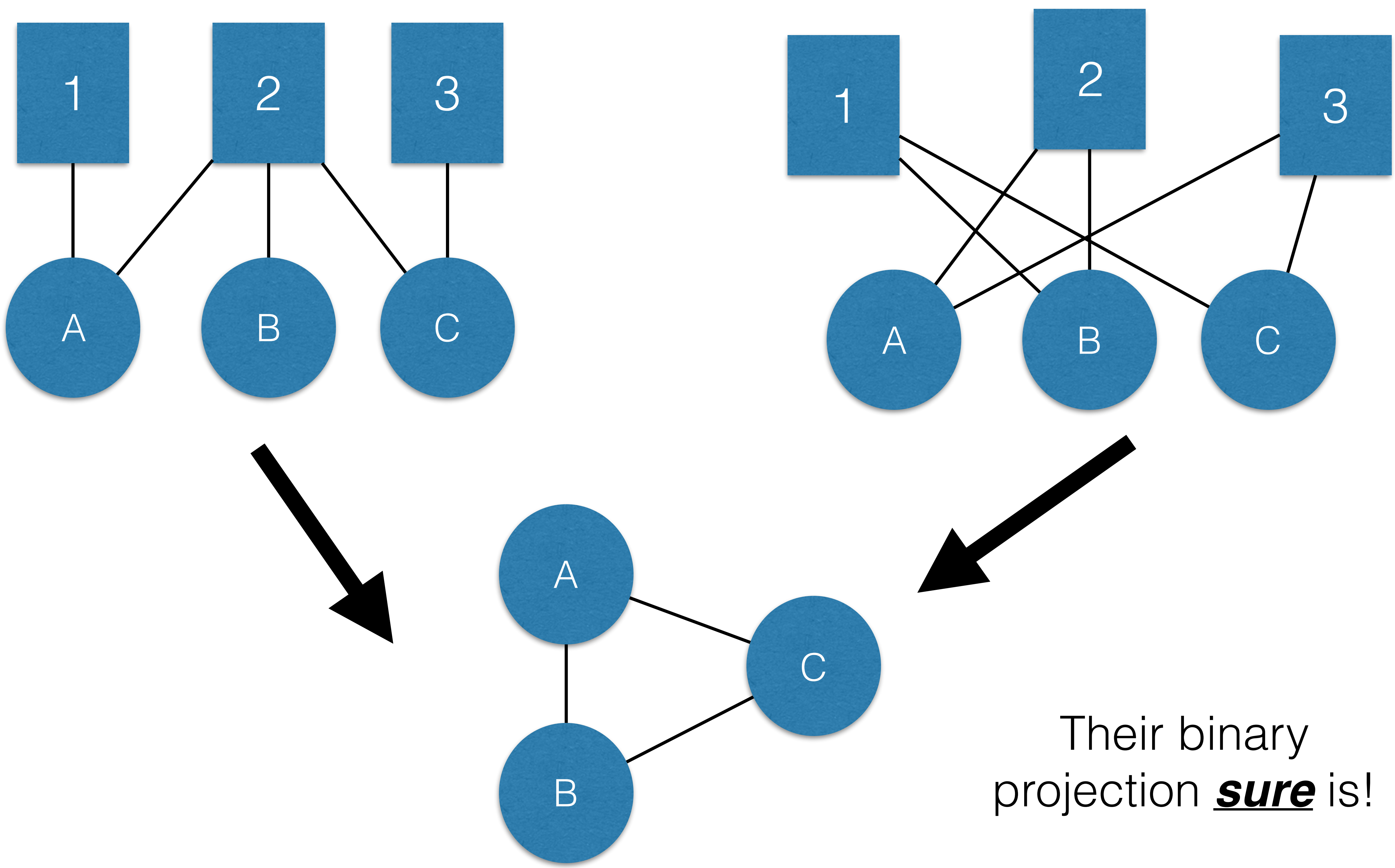
- To project, or not to project?
  - As noted by many scholars, there is data loss when we project and binarize the data.
    - Sometimes, this can be misleading.

Projection Example



Are these bipartite graphs the same?

Projection Example



Their binary projection ***sure*** is!

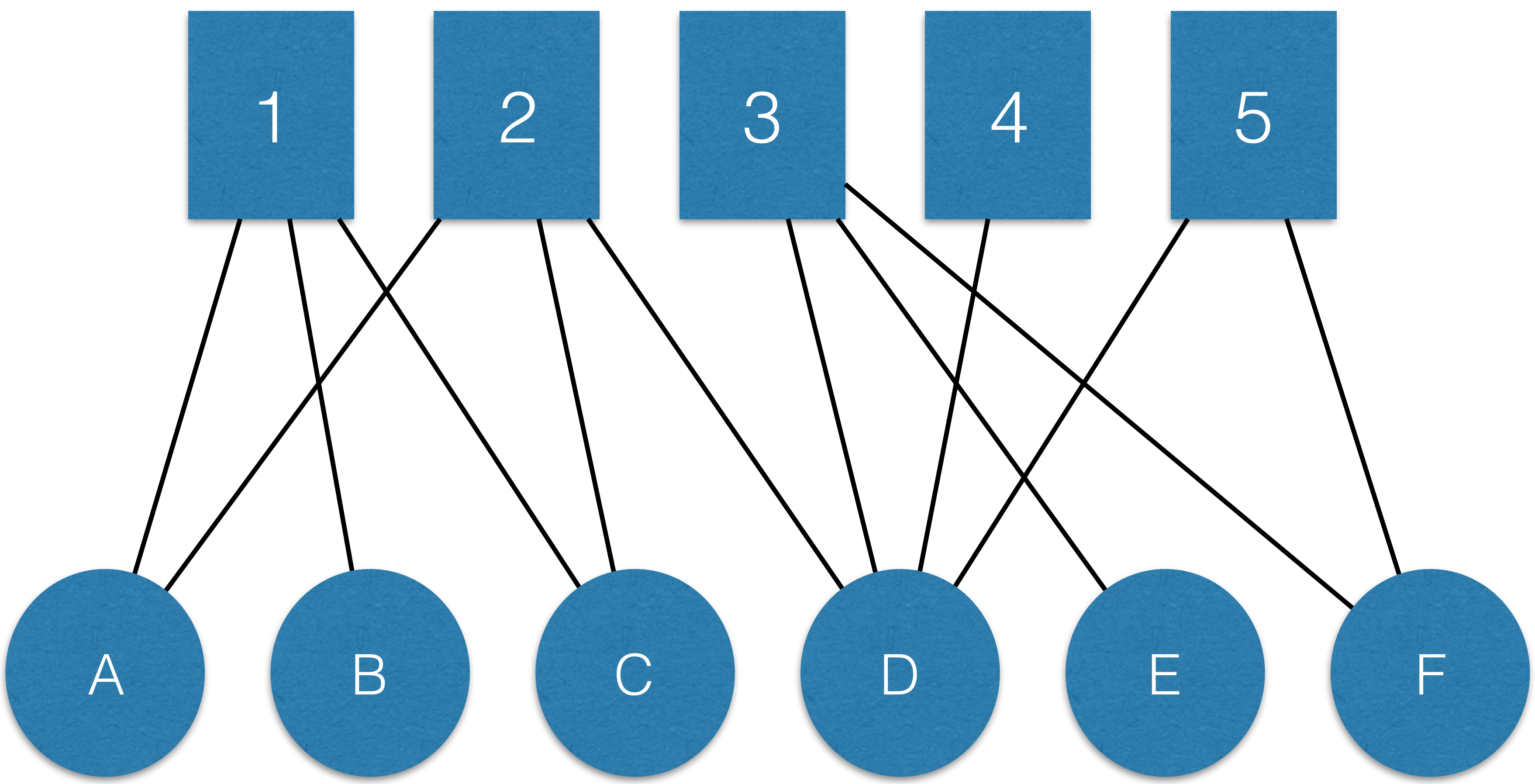
Projection

- To project, or not to project?
  - As we have seen, and as noted by many scholars, there is data loss when we project.
    - *So what to do?*
      - When you can, "keep it real" by keeping it two-mode.
      - If you must project, minimize data loss by weighting edges.

Weighted Edges

- We can use the information from the bipartite graph to weight the edges in the network.
  - This can be the same of the ties between two actors (i.e. *summation method*).

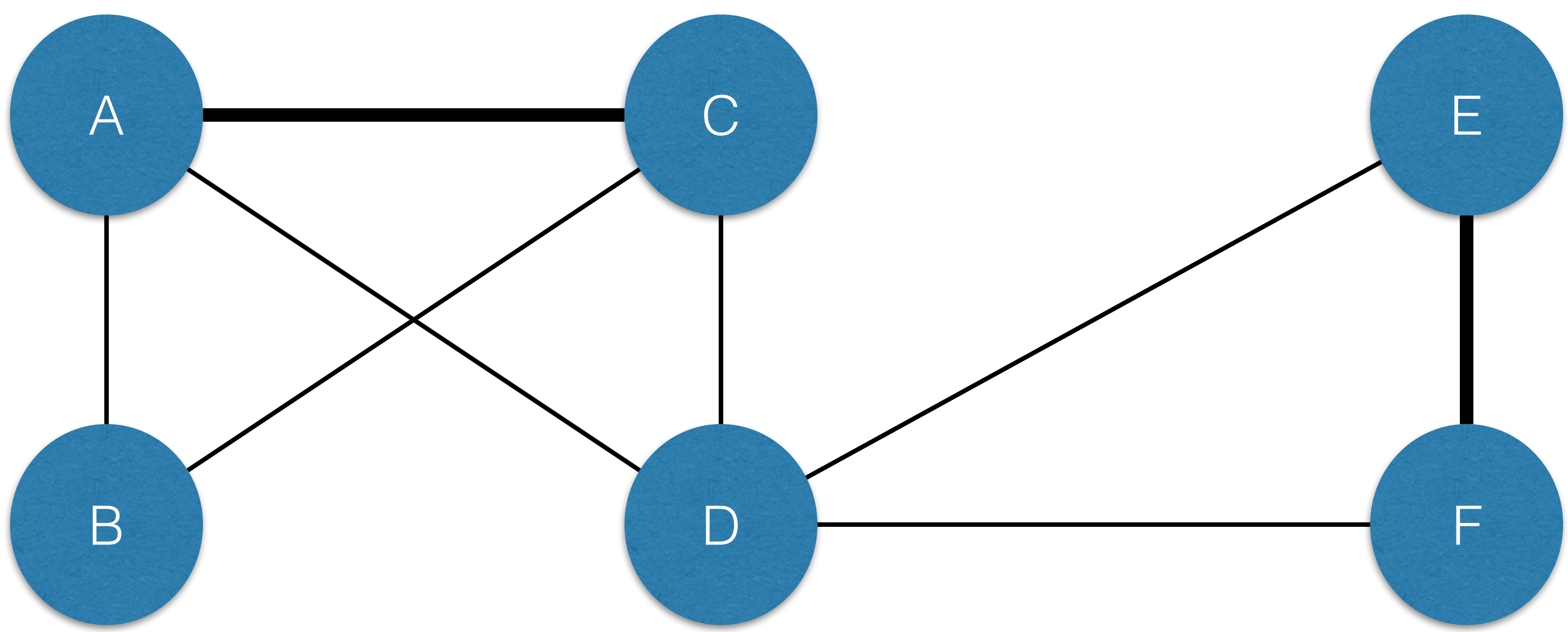
Projection



$\mathbf{A} \times \mathbf{A}^T = \mathbf{P}$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

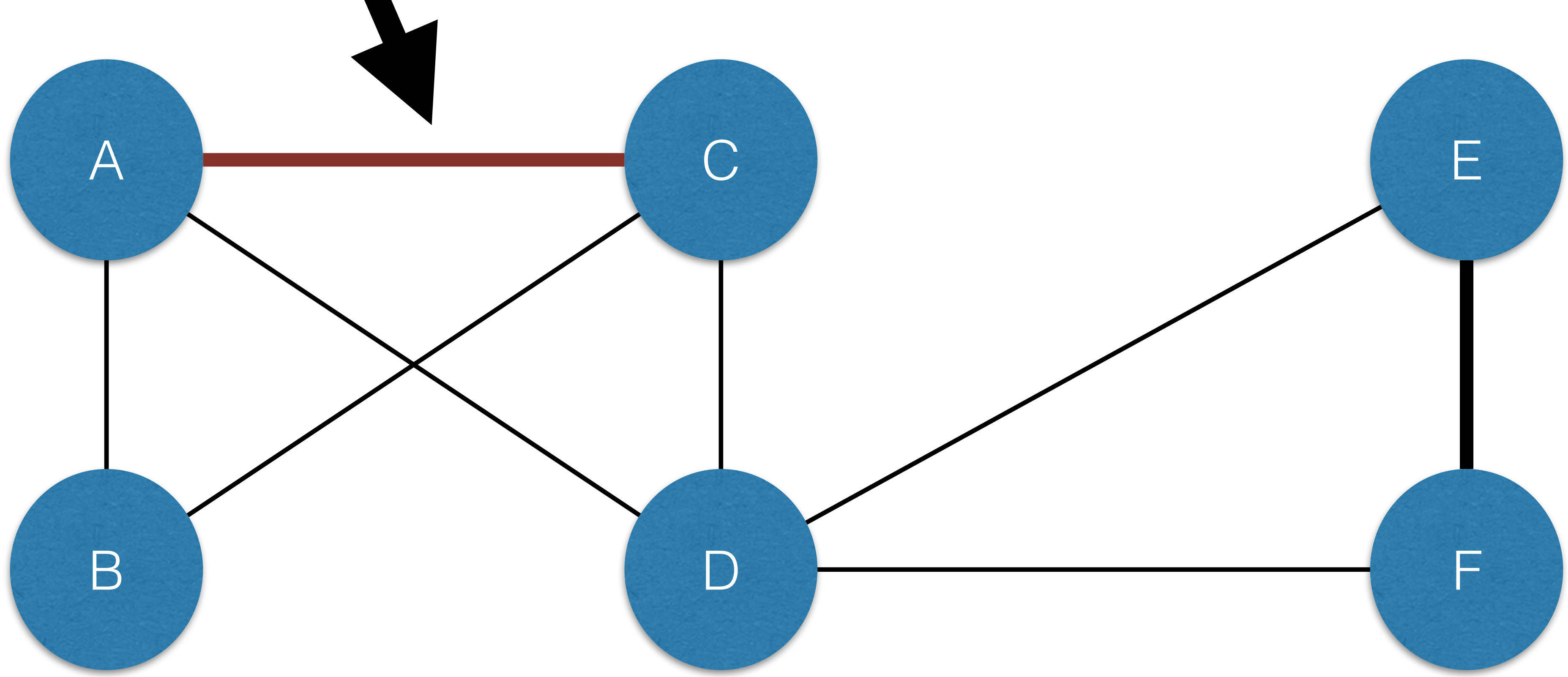
The off-diagonal entries are the tie weights



$$\mathbf{A} \times \mathbf{A}^{\mathbf{T}} = \mathbf{P}$$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	[illegible]	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

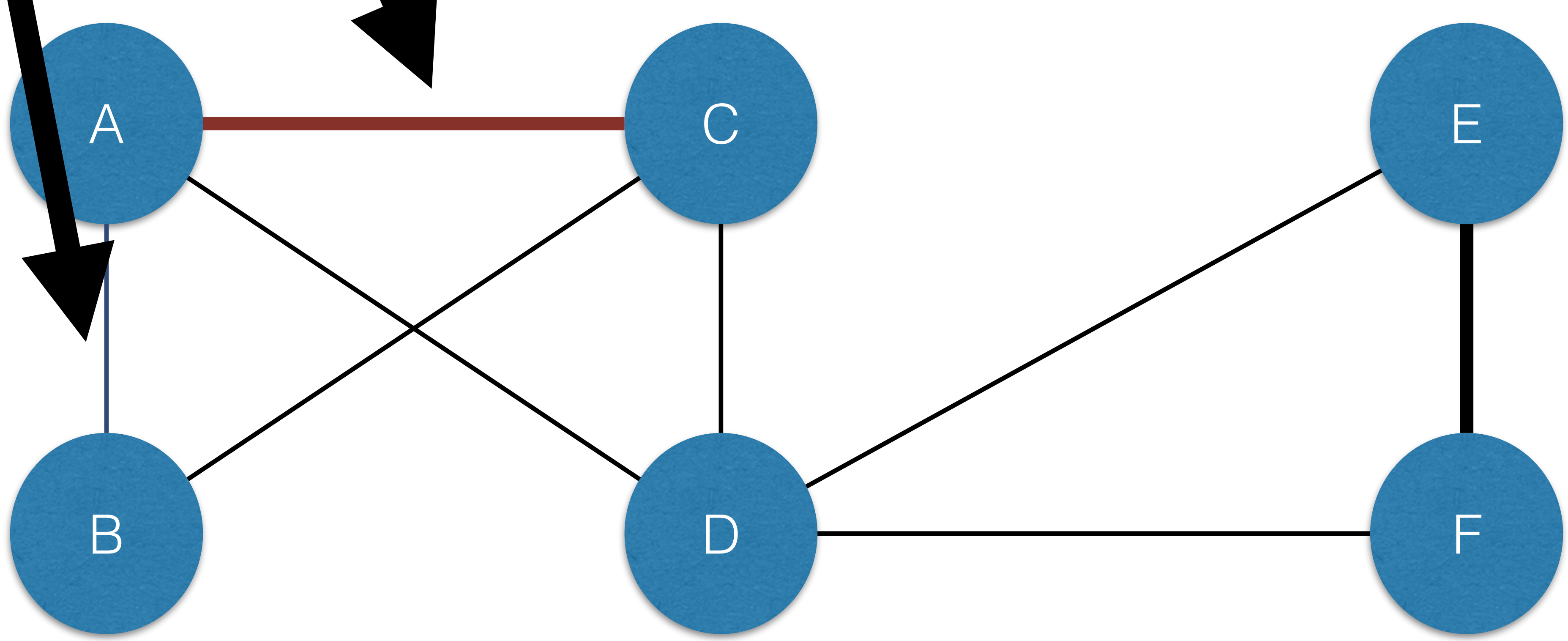
The weight corresponds to the number of nodes in the second mode that the two actors share in the bipartite graph.



$\mathbf{A} \times \mathbf{A}^\mathbf{T} = \mathbf{P}$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1		0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

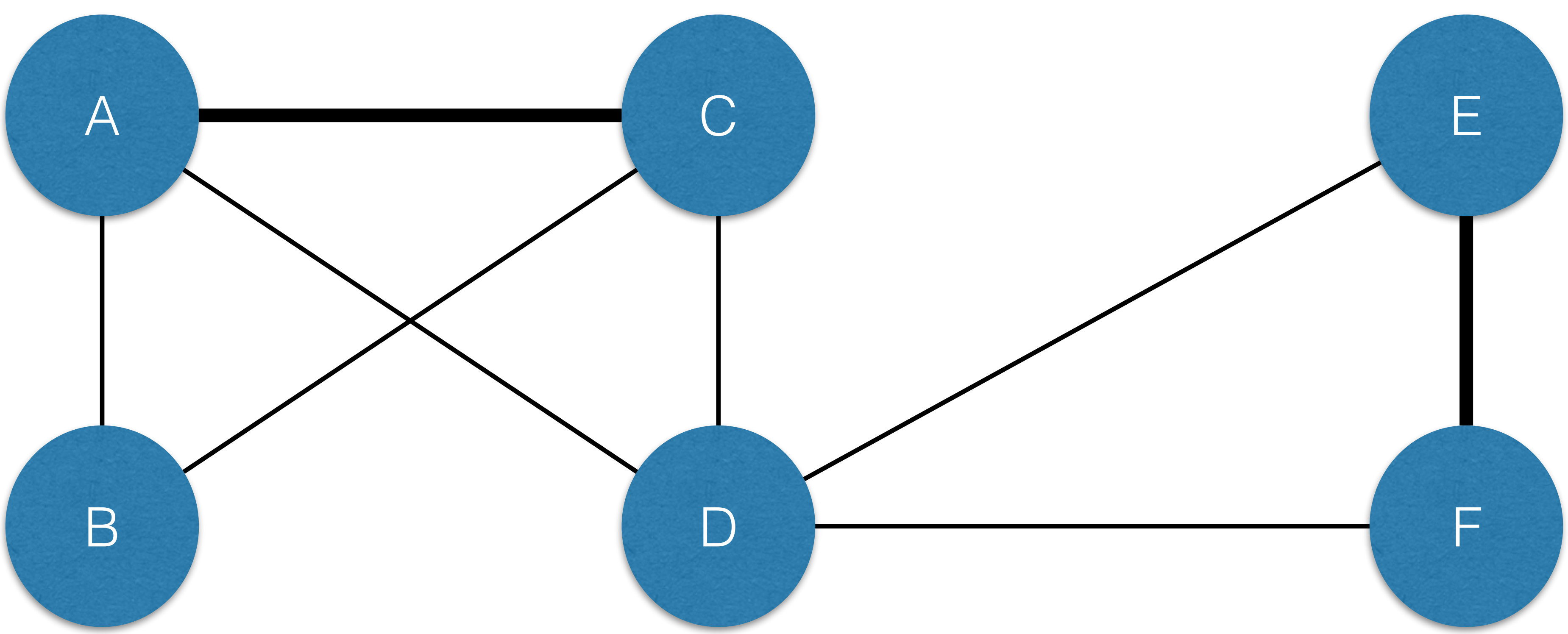
The weight corresponds to the number of nodes in the second mode that the two actors share in the bipartite graph.



$\mathbf{A} \times \mathbf{A}^{T} = \mathbf{P}$

	A	B	C	D	E	F
A	2	1	2	1	0	0
B	1	1	1	0	0	0
C	2	1	2	1	0	0
D	1	0	1	4	1	2
E	0	0	0	1	1	1
F	0	0	0	2	1	2

These weights are returned as the product matrix.



Weighted Edges

- We can use the information from the bipartite graph to weight the edges in the network.
  - This can be the sum of the ties between two actors (i.e. *summation method*).
  - We can also take into account the degrees for the second mode in the projection (i.e. *Newman method*).

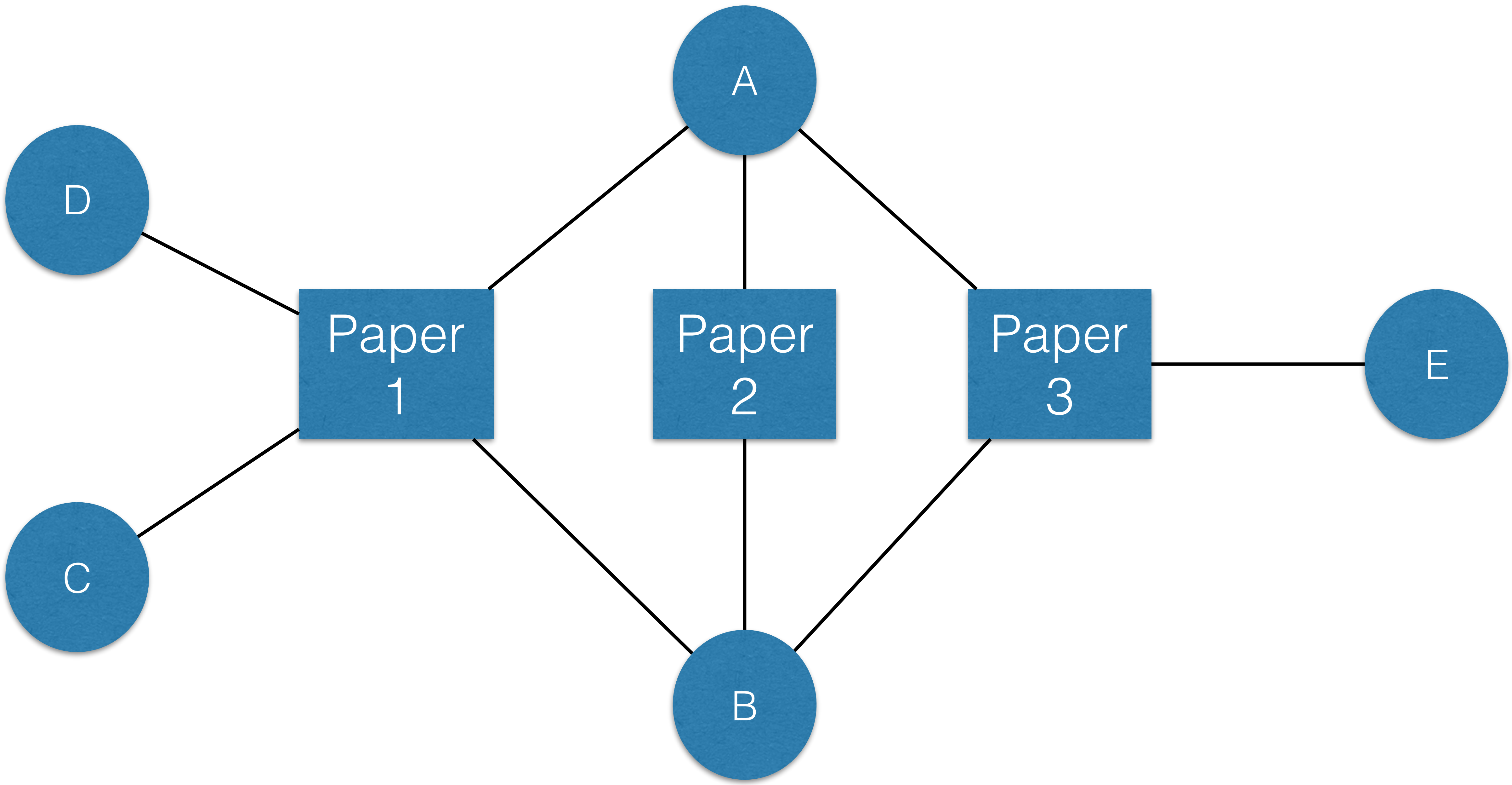
Newman Method

- Being connected to a node in the second mode that has few people may be more important than being connected to a node in the second node with many people.
  - For example: Co-authorship network
    - If I share a paper I wrote only with you, should we have the same edge weight as two people who share a paper that has five total authors?
      - In the weighted projection, the edges are the same weight, 1.

Newman Method

- Newman (2001: 5) and scientific collaboration
  - “it is probably the case…that two scientists whose names appear on a paper together with many other coauthors know one another less well on average than two who were the sole author of a paper”
    - *What argument is he making?*

Projection

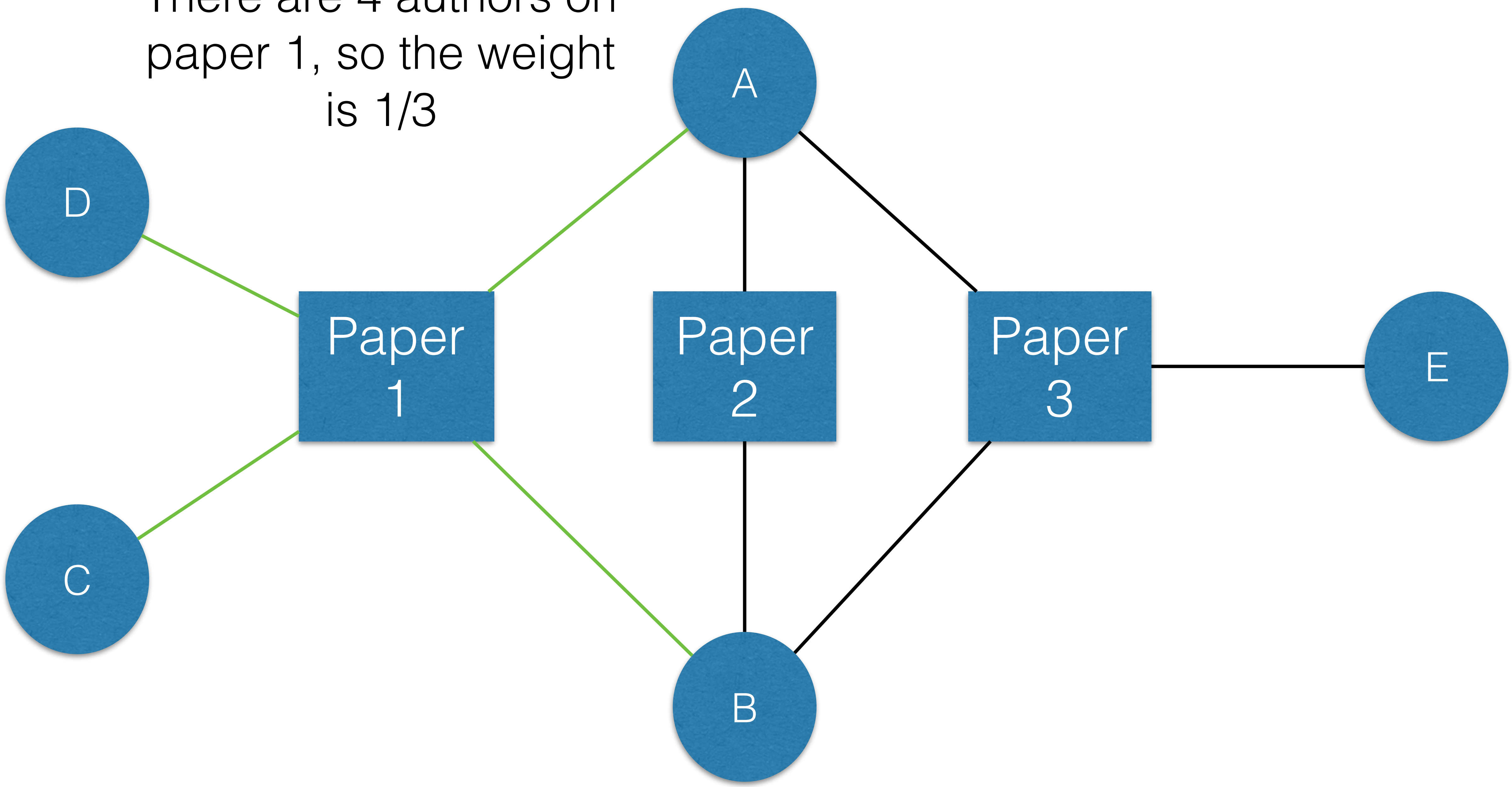


Newman Method

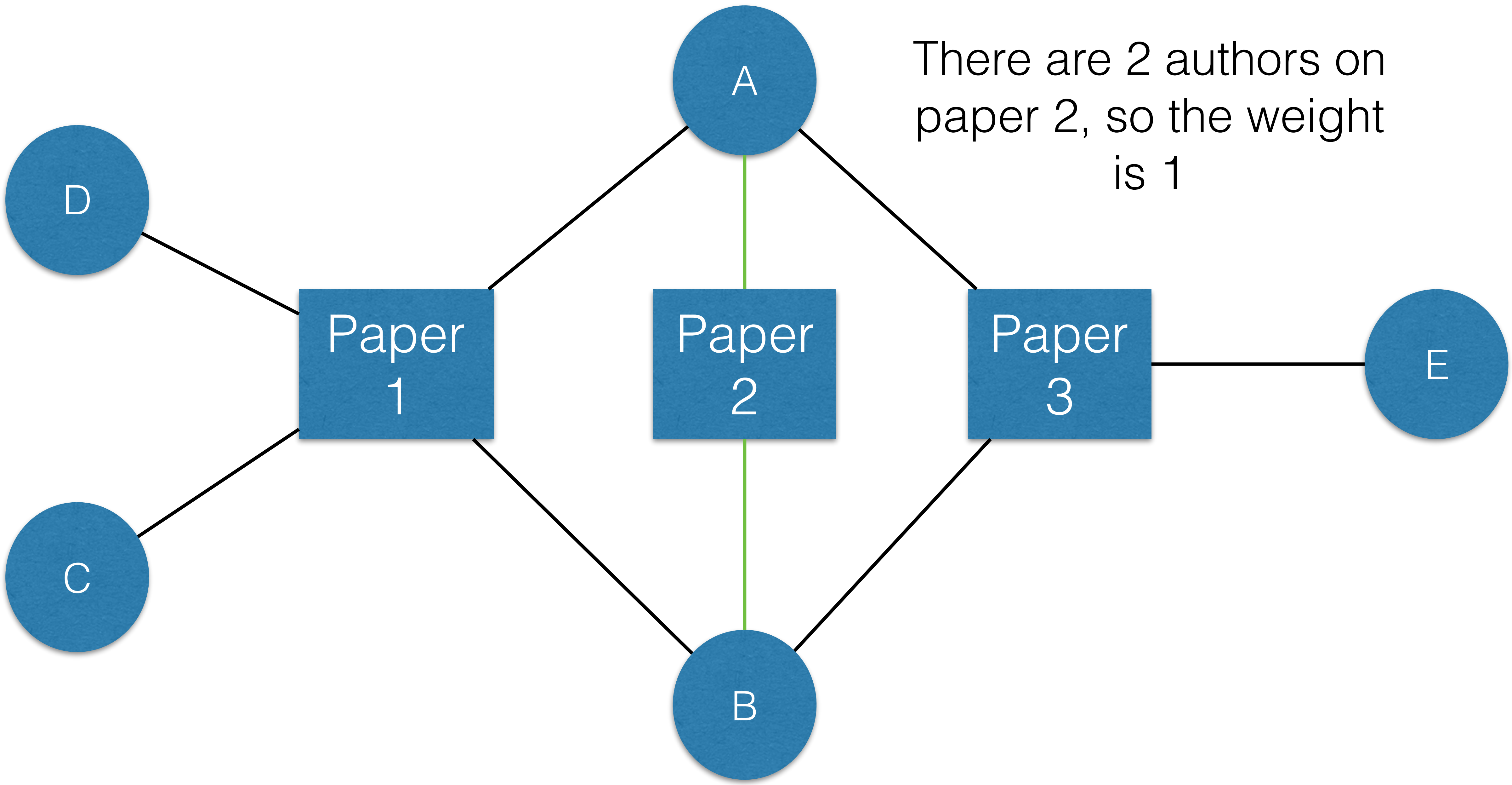
- The weight of a tie is then just:
  - *1/n-1* where *n* is the number of authors
    - If there are 5 authors on 1 paper, then the tie weight is 1/4 or 0.25.
    - If there are 2 authors on 1 paper, then the tie weight is 1/1 or 1.00.
      - Then, the projected edge weight between two authors is the sum of these weights.

Newman Method Projection

There are 4 authors on
paper 1, so the weight
is 1/3

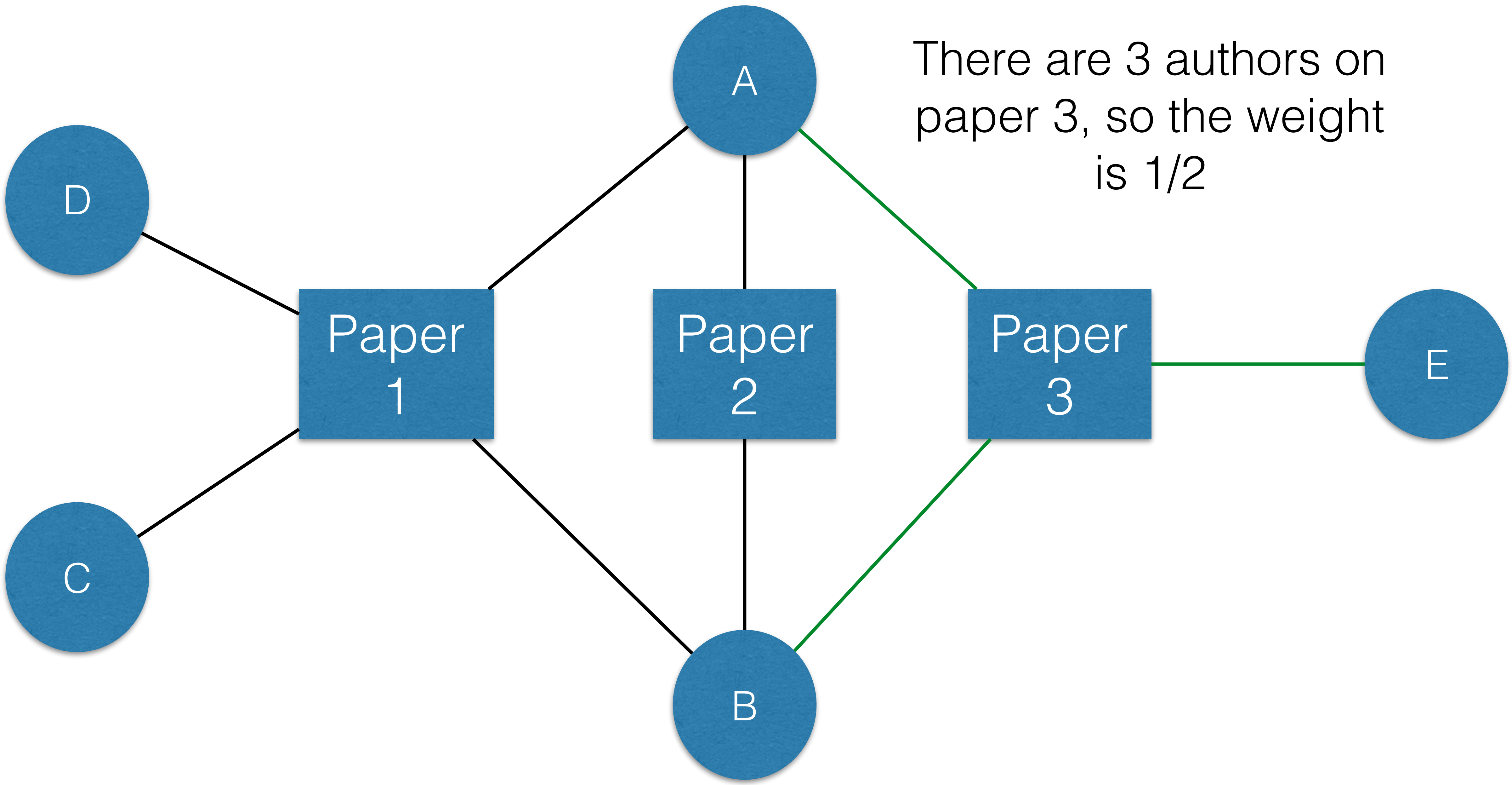


Projection



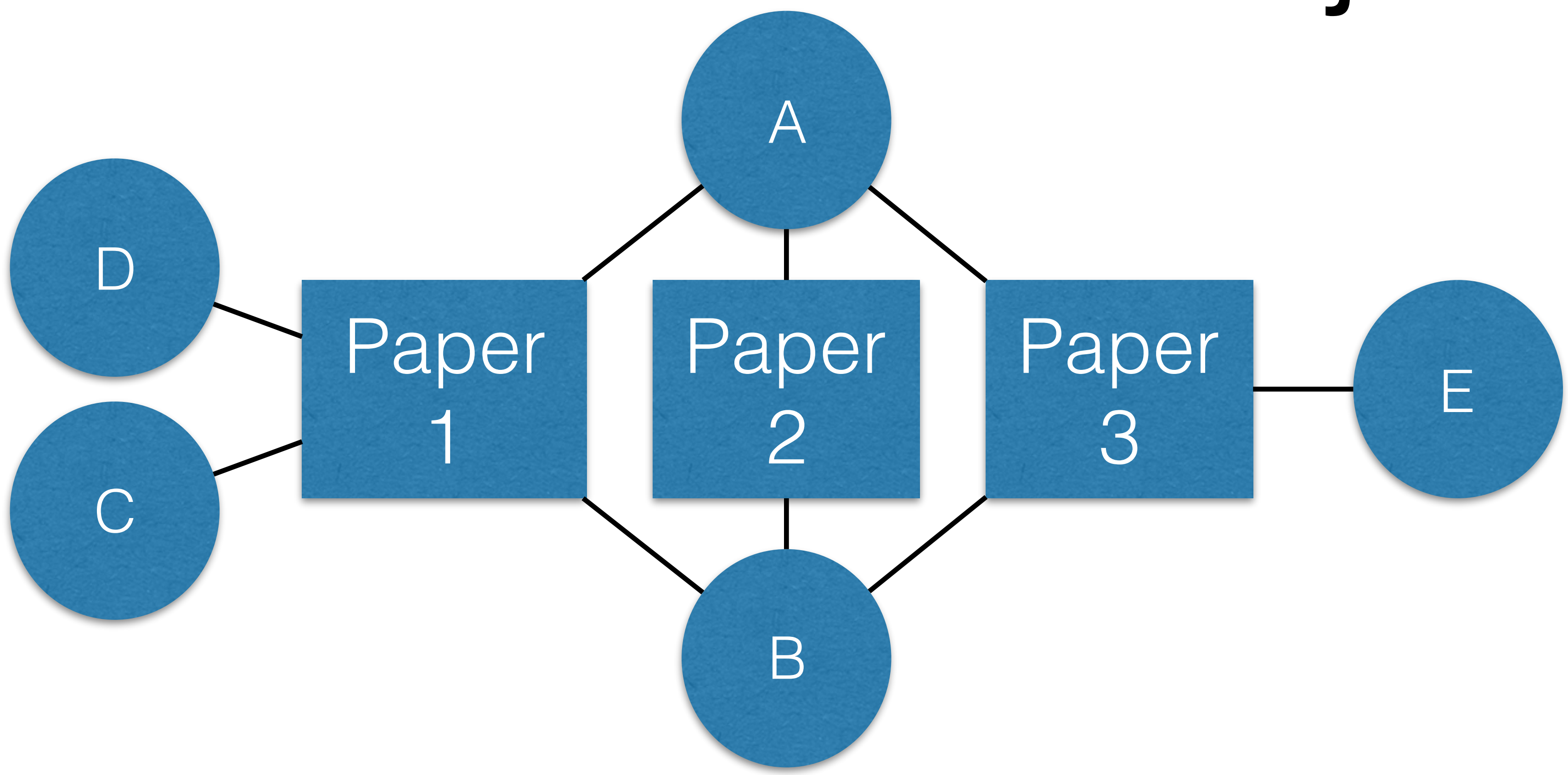
There are 2 authors on paper 2, so the weight is 1

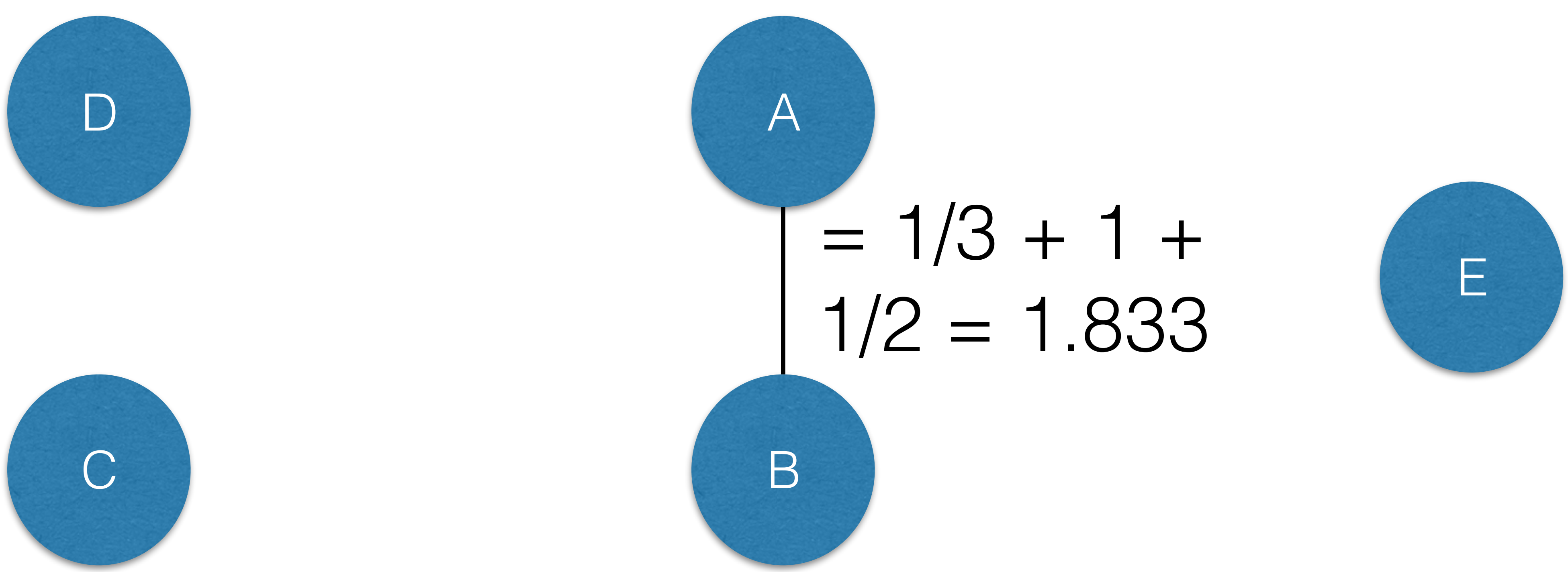
Projection



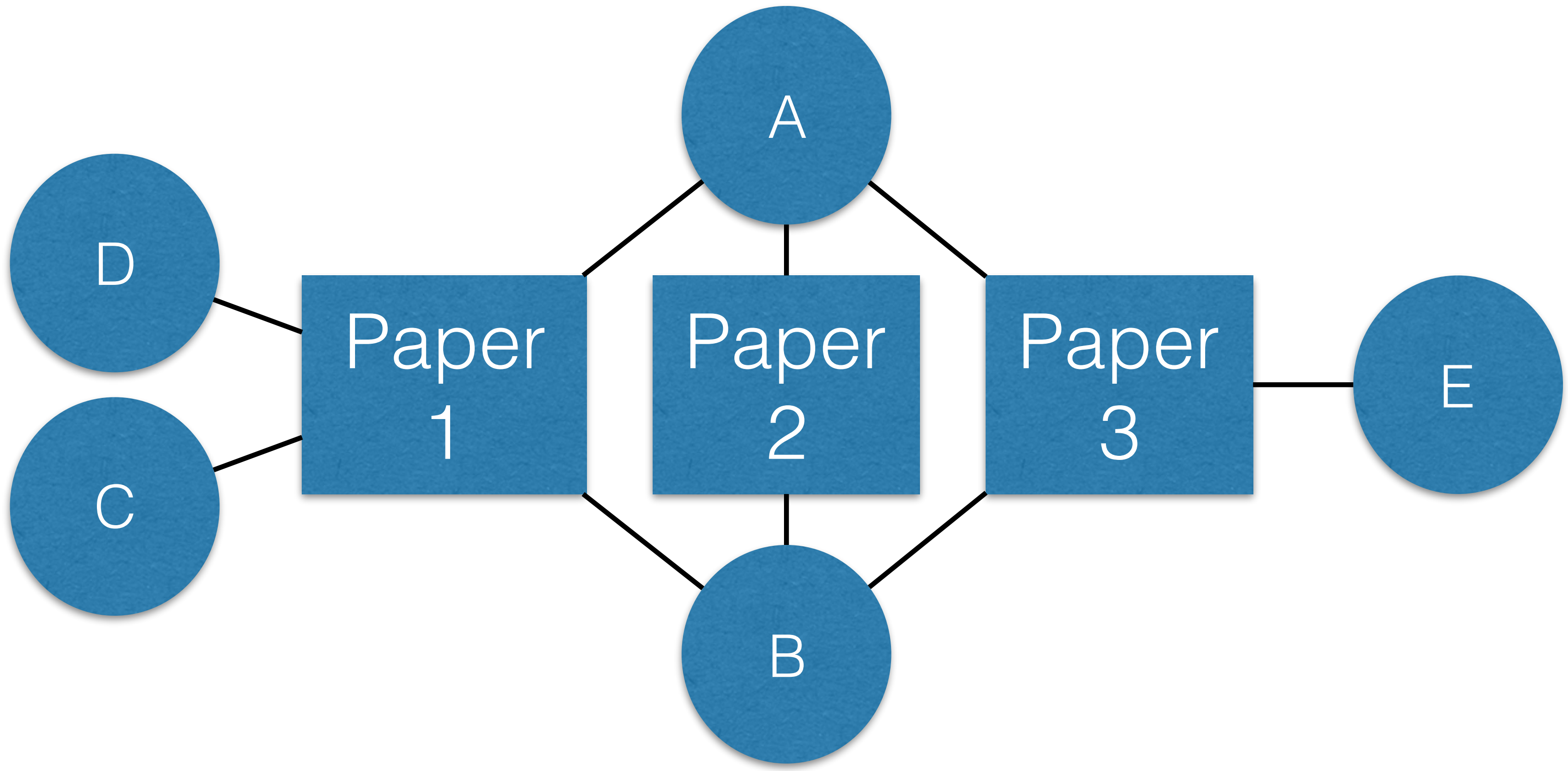
There are 3 authors on paper 3, so the weight is 1/2

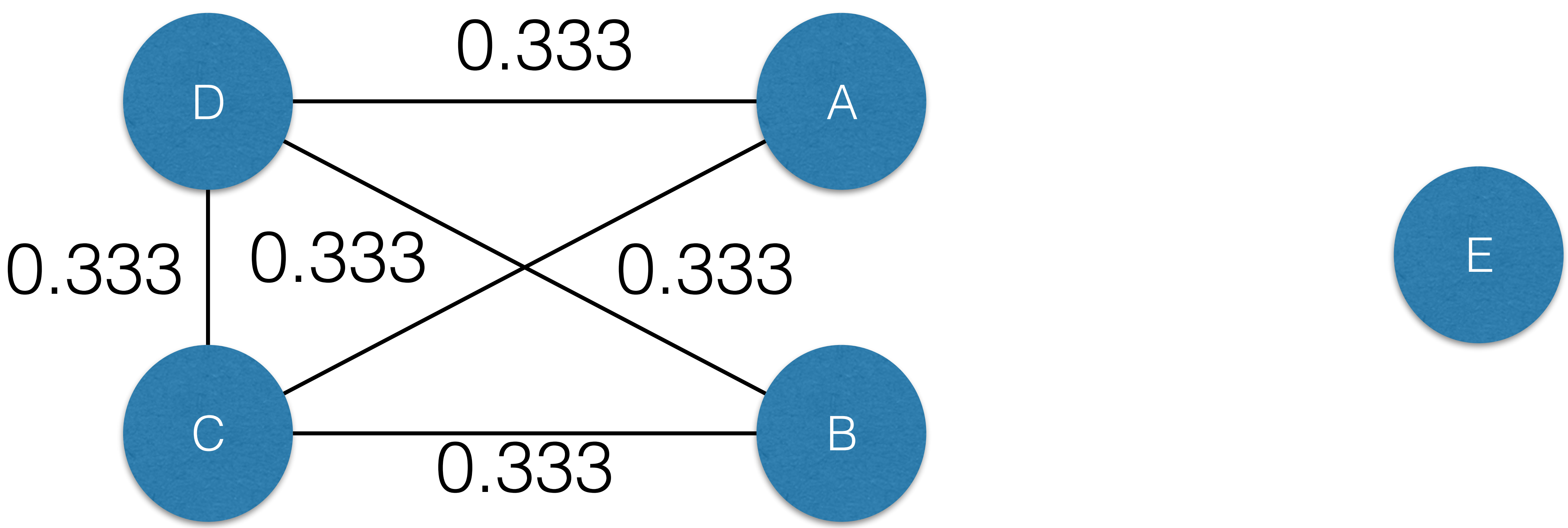
Newman Method Projection



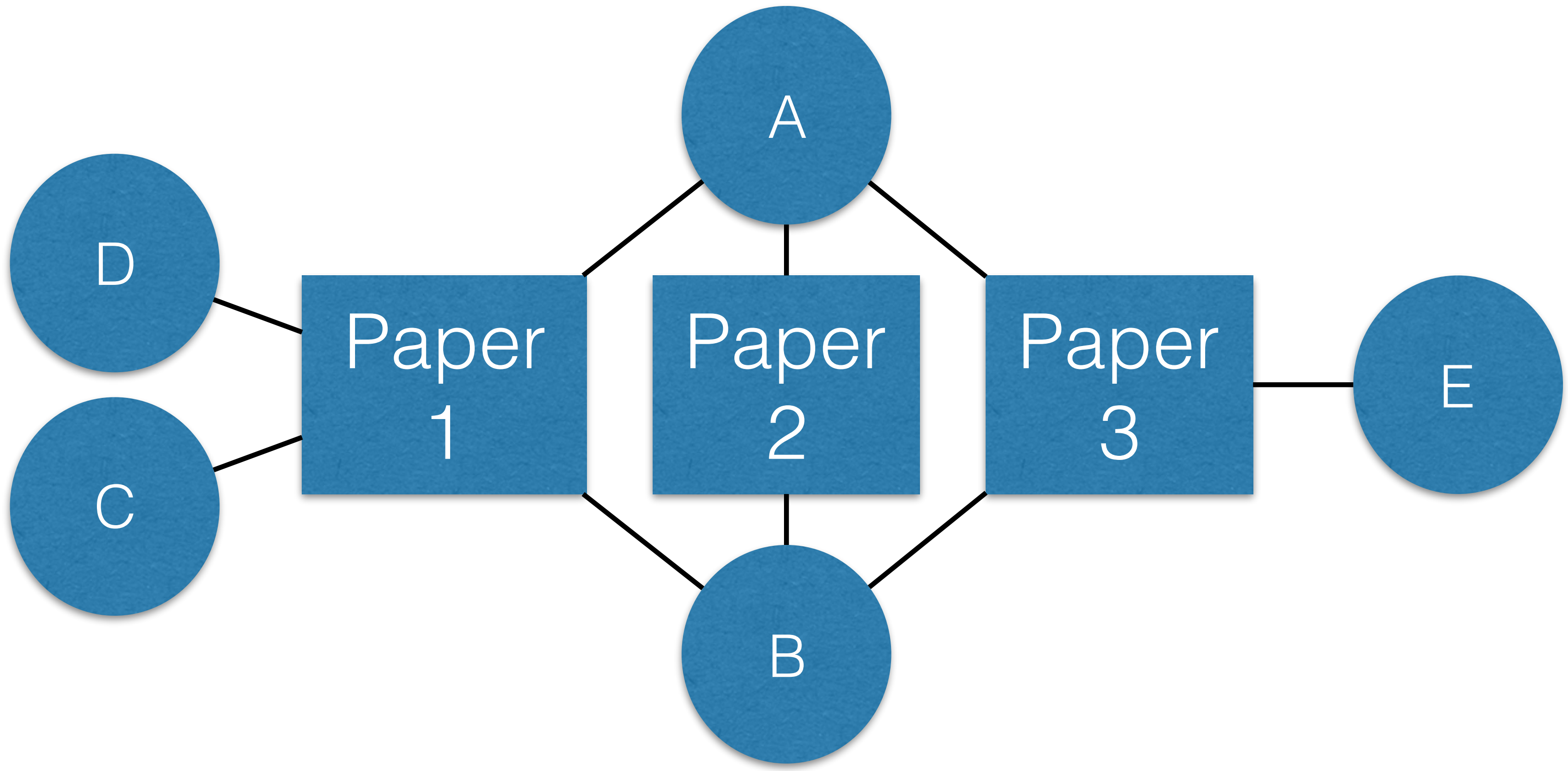


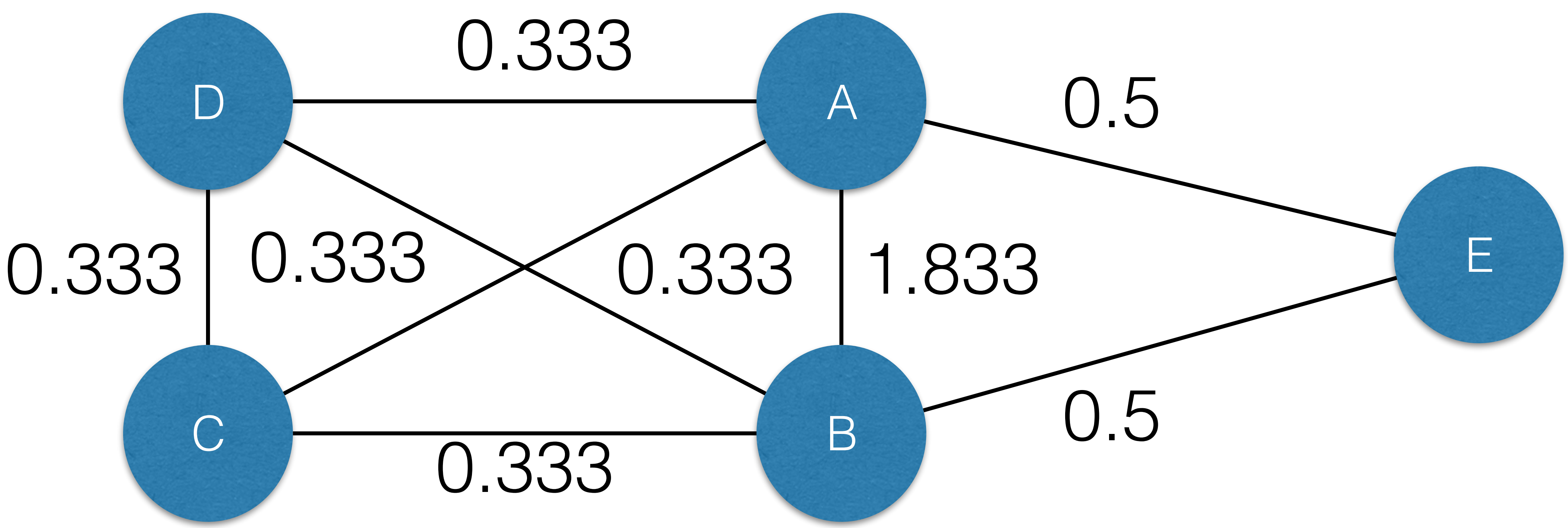
Newman Method Projection





Newman Method Projection

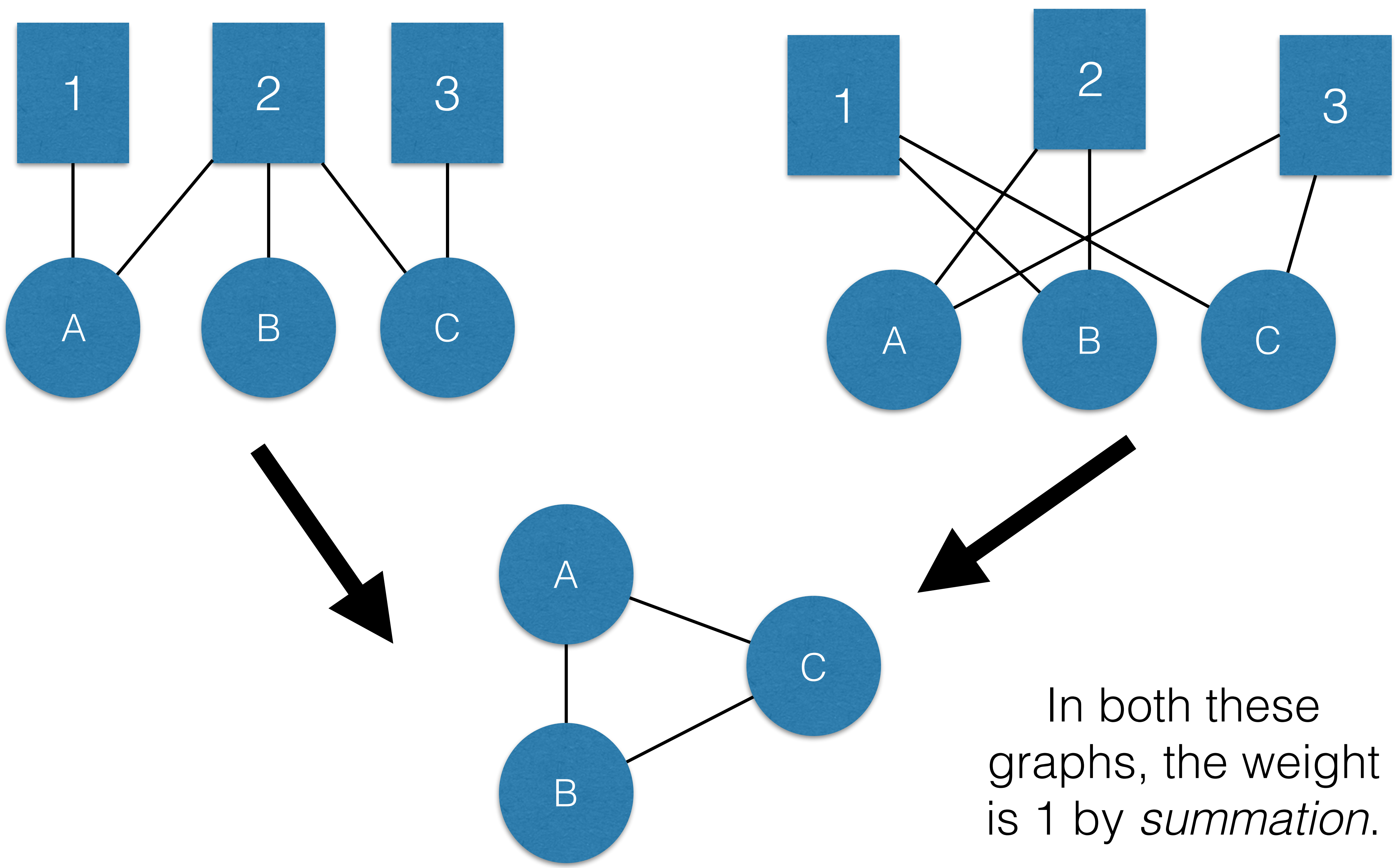




Newman Method

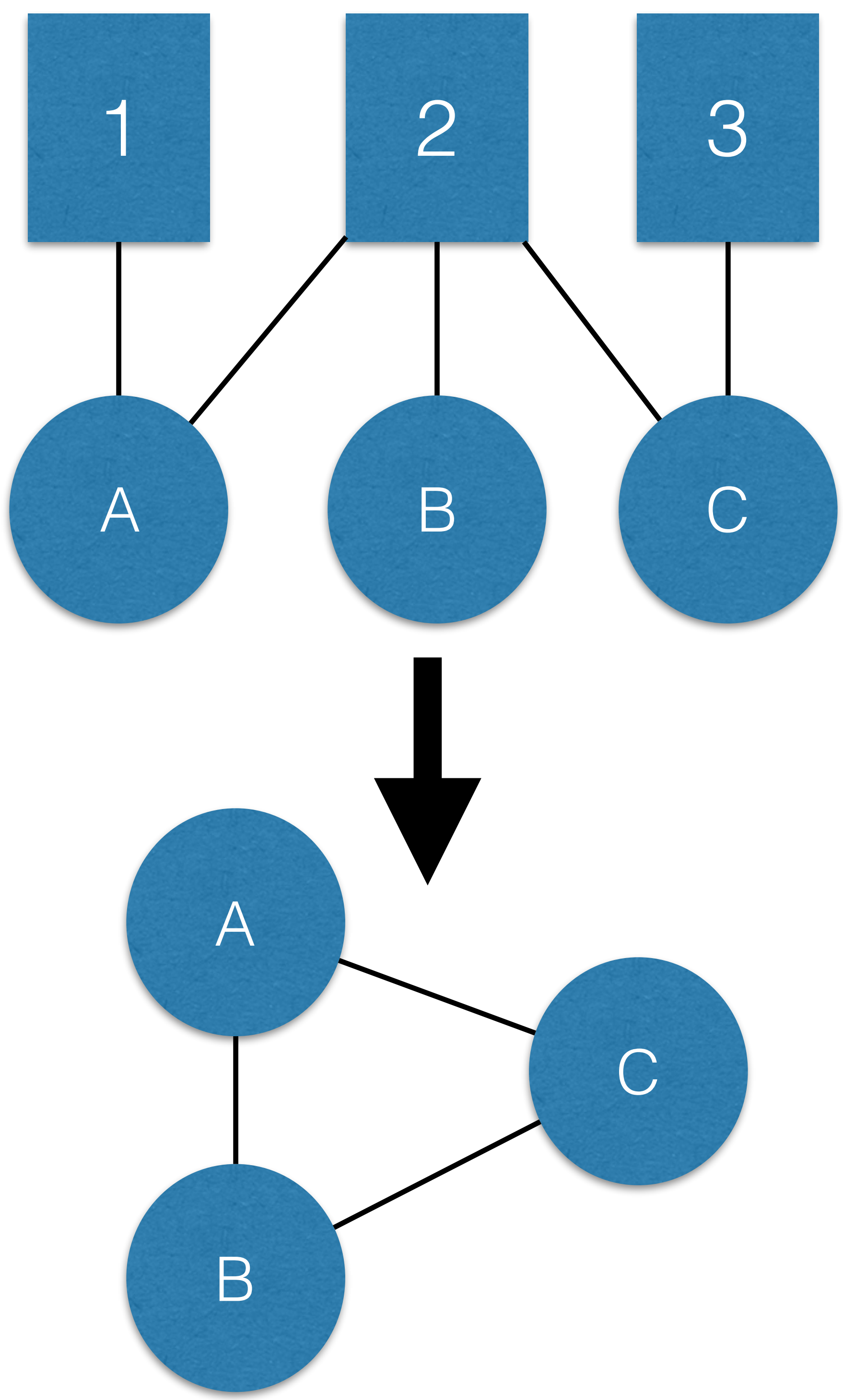
- Note that we can generalize this past papers to any two-mode network.
    - For example, being arrested with 1 person in a co-offending network versus being arrested with 49 people in a co-offending network.
- Also, note that compared to the *summation* method, the *Newman* method shows how nodes in the first mode are connected by **different** modes in the second mode.

Projection Example



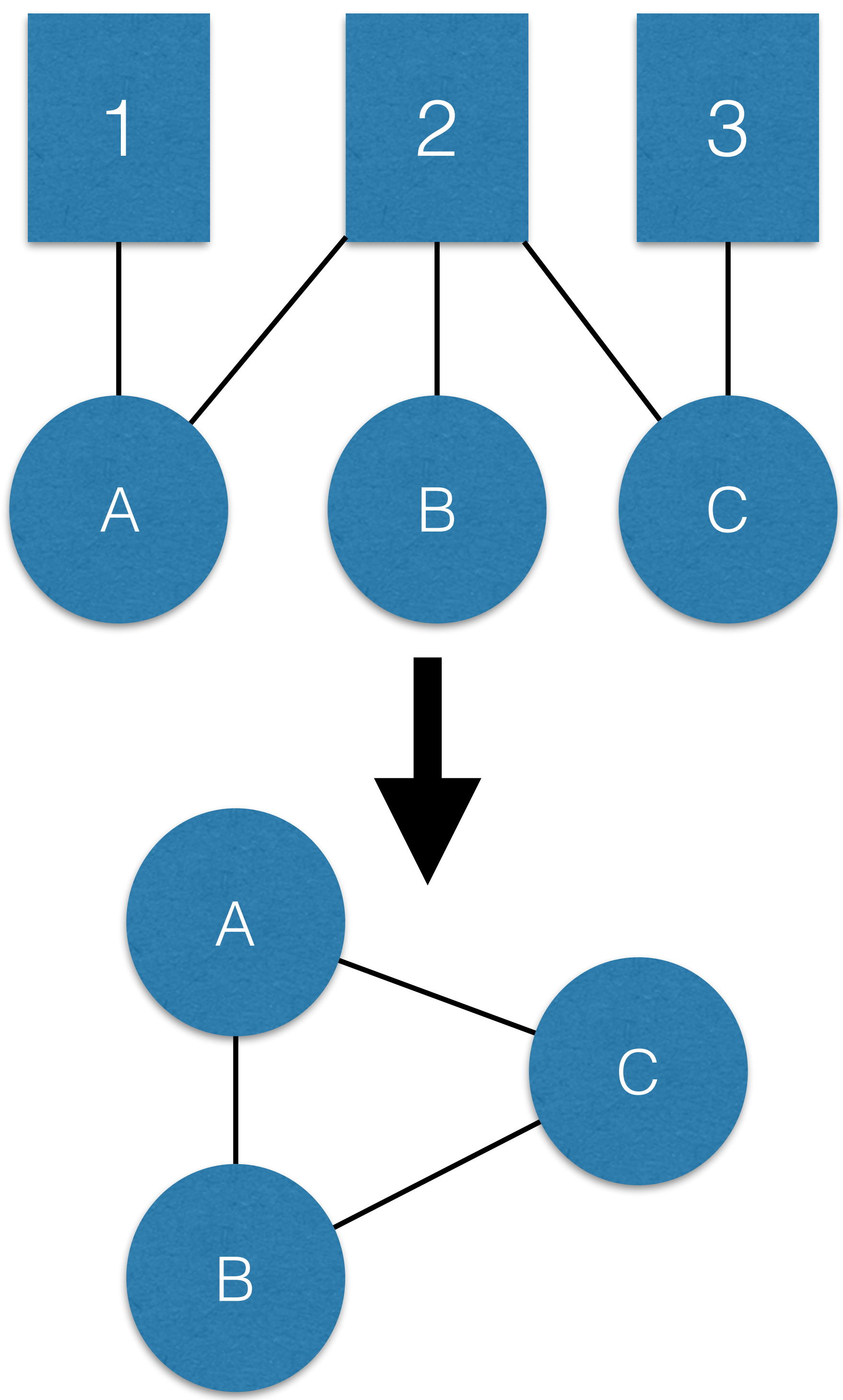
In both these graphs, the weight is 1 by *summation*.

Projection Example



Using the *Newman* method, what are the edge weights for this graph?

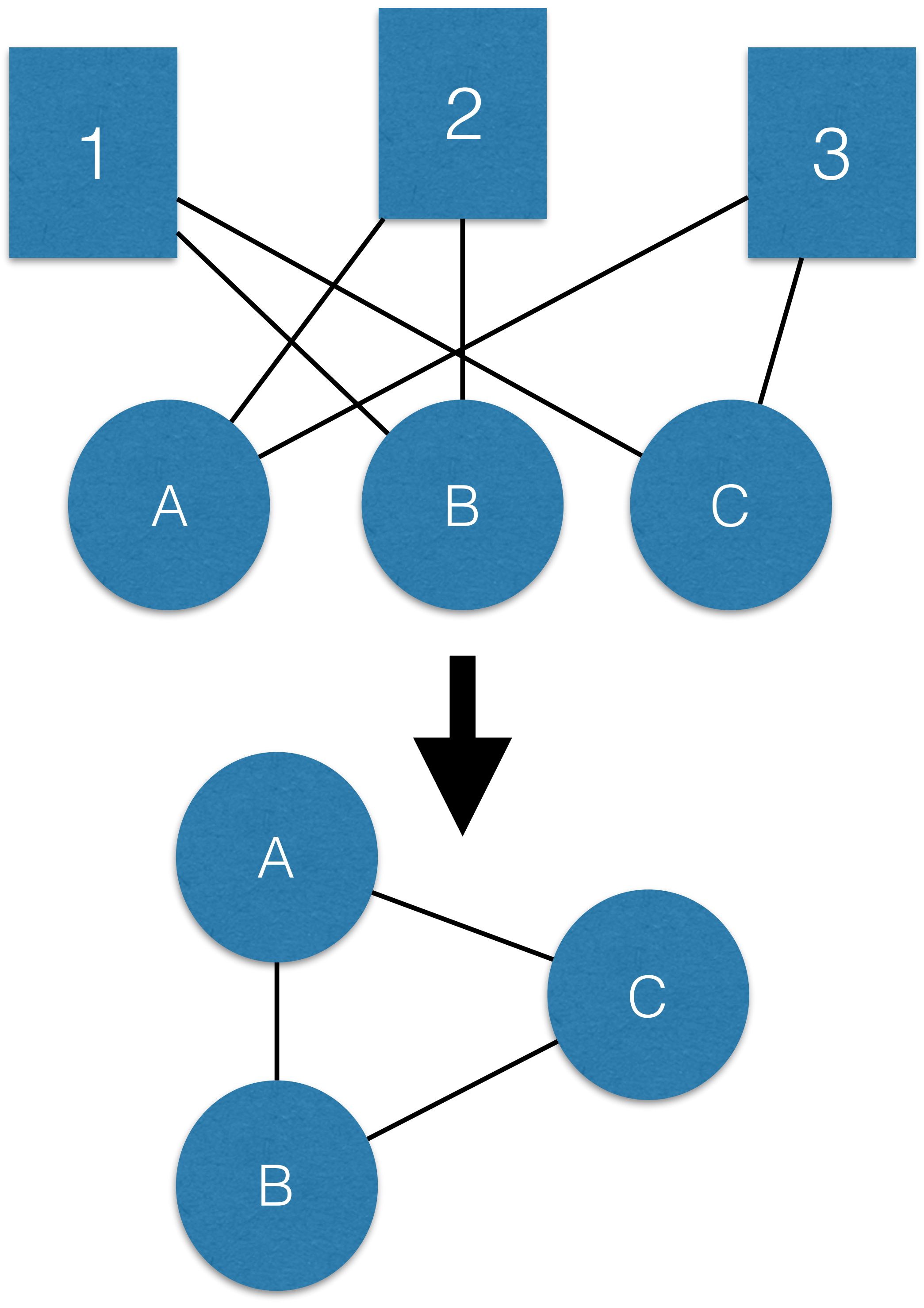
Projection Example



Node 2 has degree 3, so the weight is 1/3-1 = 1/2 = 0.5

Projection Example

Using the *Newman* method, what are the edge weights for this graph?

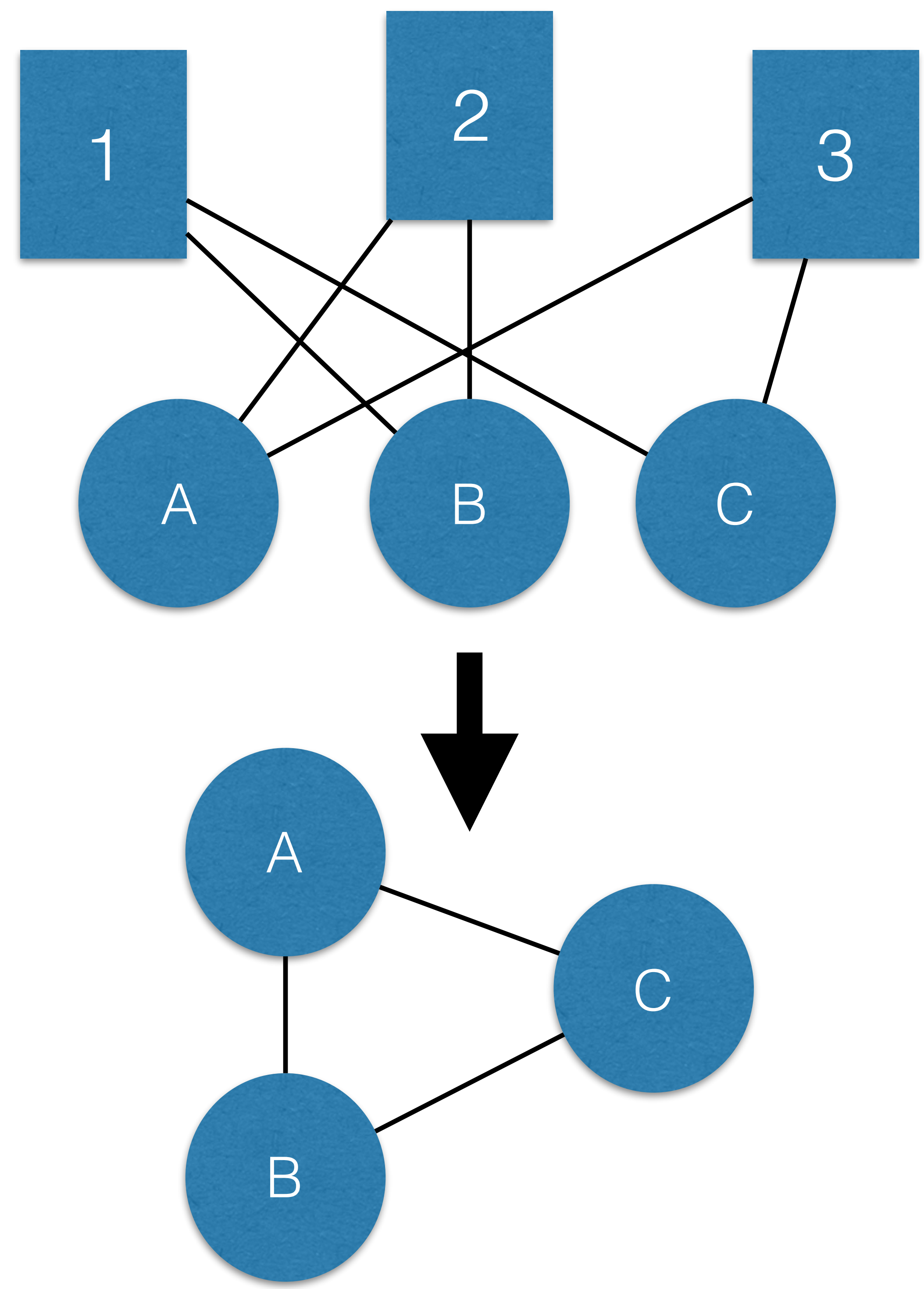


Projection Example

Node 1 has degree 2, so the weight is 1/2-1 = 1/1 = 1.0

Node 2 has degree 2, so the weight is 1/2-1 = 1/1 = 1.0

Node 3 has degree 2, so the weight is 1/2-1 = 1/1 = 1.0



Learning Goals

- Understand *projection* of bipartite graphs to unipartite graphs.
- Examine *dichotomized* projections.
- Examine *summation weighted* and *Newman weighted* projections.

Questions?