Ellipsoid Fitting Up to a Constant

Jeff Xu
Carnegie Mellon University

Joint Work with Tim Hsieh (CMU) , Pravesh Kothari (CMU) , and Aaron Potechin (UChicago)


Overview

1. **Ellipsoid Fitting Conjecture**
2. Constructing an Ellipsoid
3. Analysis via Graph Matrices
4. A Local Machinery for Tight Norm Bounds

Ellipsoid Fitting Conjecture

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$

Normalized to be roughly unit norm

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Is there a symmetric matrix $\Lambda \in R^{d \times d}$ such that
  - $v_i^T \Lambda v_i = 1$ for all v_i
  - $\Lambda \succeq 0$ (Positive-semidefinite) ?

Normalized to be roughly unit norm

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$

Normalized to be roughly unit norm

- Is there a symmetric matrix $\Lambda \in R^{d \times d}$ such that
    - $v_i^T \Lambda v_i = 1$ for all v_i
    - $\Lambda \succeq 0$ (Positive-semidefinite) ?
- Geometrically, $\Lambda \in R^{d \times d}$ is an **ellipsoid** centered at origin that **passes through all m points on boundary.**

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Is there a matrix $\Lambda \in R^{d \times d}$ such that
  - $v_i^T \Lambda v_i = 1$ for all v_i, and $\Lambda \succeq 0$ (Positive-semidefinite) ?

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Is there a matrix $\Lambda \in R^{d \times d}$ such that
    - $v_i^T \Lambda v_i = 1$ for all v_i, and $\Lambda \succeq 0$ (Positive-semidefinite) ?
- **[SCPW Conjecture]** For all $\epsilon > 0$, and for sufficiently large d,

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Is there a matrix $\Lambda \in R^{d \times d}$ such that
  - $v_i^T \Lambda v_i = 1$ for all v_i, and $\Lambda \succeq 0$ (Positive-semidefinite) ?
- **[SCPW Conjecture]** For all $\epsilon > 0$, and for sufficiently large d,
  - **(Positive)** If $m \leq (1 - \epsilon)\frac{d^2}{4}$, there exists such an ellipsoid w.h.p.

Ellipsoid Fitting Conjecture

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Is there a matrix $\Lambda \in R^{d \times d}$ such that
  - $v_i^T \Lambda v_i = 1$ for all v_i, and $\Lambda \succeq 0$ (Positive-semidefinite) ?
- **[SCPW Conjecture]** For all $\epsilon > 0$, and for sufficiently large d,
  - **(Positive)** If $m \leq (1-\epsilon)\frac{d^2}{4}$, there exists such an ellipsoid w.h.p.
  - **(Negative)** If $m \geq (1+\epsilon)\frac{d^2}{4}$, there does not exist such an ellipsoid w.h.p.

Ellipsoid Fitting Conjecture

- Given m random points $v_1, ..., v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Is there a matrix $\Lambda \in R^{d \times d}$ such that
  - $v_i^T \Lambda v_i = 1$ for all v_i, and $\Lambda \succeq 0$ (Positive-semidefinite) ?
- **[SCPW Conjecture]** For all $\epsilon > 0$, and for sufficiently large d,
  - **(Positive)** If $m \leq (1 - \epsilon)\frac{d^2}{4}$, there exists such an ellipsoid w.h.p.
  - **(Negative)** If $m \geq (1 + \epsilon)\frac{d^2}{4}$, there does not exist such an ellipsoid w.h.p.

Sharp transition!

Ellipsoid Fitting Conjecture

(Negative side via dimension-argument)

Conjecture:

Impossible when $m \geq (1+\epsilon)\frac{d^2}{4}$

Ellipsoid Fitting Conjecture

(Negative side via dimension-argument)

Conjecture:

Impossible when $m \geq (1+\epsilon)\frac{d^2}{4}$

- **m** equations $v_i^T \Lambda v_i = 1$ for symmetric $\Lambda \in R^{d \times d}$

Ellipsoid Fitting Conjecture

(Negative side via dimension-argument)

Conjecture:

Impossible when $m \geq (1+\epsilon)\frac{d^2}{4}$

- **m** equations $v_i^T \Lambda v_i = 1$ for symmetric $\Lambda \in R^{d \times d}$
  - View each as a linear constraint on $\frac{d(d+1)}{2}$ entries of Λ

Ellipsoid Fitting Conjecture

(Negative side via dimension-argument)

Conjecture:

Impossible when $m \geq (1+\epsilon)\frac{d^2}{4}$

- **m** equations $v_i^T \Lambda v_i = 1$ for symmetric $\Lambda \in R^{d \times d}$
  - View each as a linear constraint on $\frac{d(d+1)}{2}$ entries of Λ
  - When randomly chosen, as long as $m > \frac{d(d+1)}{2}$, more constraints than the degree of freedom

Ellipsoid Fitting Conjecture

(Negative side via dimension-argument)

Conjecture:

Impossible when $m \geq (1+\epsilon)\frac{d^2}{4}$

- **m** equations $v_i^T \Lambda v_i = 1$ for symmetric $\Lambda \in R^{d\times d}$
  - View each as a linear constraint on $\frac{d(d+1)}{2}$ entries of Λ
  - When randomly chosen, as long as $m > \frac{d(d+1)}{2}$, more constraints than the degree of freedom

—> w.h.p., the linear system cannot be satisfied.

Ellipsoid Fitting Conjecture

(Negative side via dimension-argument)

Conjecture:

Impossible when $m \geq (1+\epsilon)\frac{d^2}{4}$

- **m** equations $v_i^T \Lambda v_i = 1$ for symmetric $\Lambda \in R^{d\times d}$
  - View each as a linear constraint on $\frac{d(d+1)}{2}$ entries of Λ
  - When randomly chosen, as long as $m > \frac{d(d+1)}{2}$, more constraints than the degree of freedom

    —> w.h.p., the linear system cannot be satisfied.
- **Open Question**: Can we improve upon this bound? PSDness?

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky '13] $m = O(d^{6/5-c})$ for any constant c>0

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky ’13] $m = O(d^{6/5-c})$ for any constant c>0
- Recently, motivated from studies in Sum-of-Squares Algorithms

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky '13] $m = O(d^{6/5-c})$ for any constant c>0
- Recently, motivated from studies in Sum-of-Squares Algorithms
  - [Ghosh, Jeronimo, Jones, Potechin, Rajendran '20] $m = \tilde{O}(d^{3/2})$

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky '13] $m = O(d^{6/5-c})$ for any constant c>0
- Recently, motivated from studies in Sum-of-Squares Algorithms
  - [Ghosh, Jeronimo, Jones, Potechin, Rajendran '20] $m = \tilde{O}(d^{3/2})$
    - Implicitly extends to $m = O(\frac{d^2}{polylog(d)})$

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky '13] $m = O(d^{6/5-c})$ for any constant c>0
- Recently, motivated from studies in Sum-of-Squares Algorithms
  - [Ghosh, Jeronimo, Jones, Potechin, Rajendran '20] $m = \tilde{O}(d^{3/2})$
    - Implicitly extends to $m = O(\frac{d^2}{polylog(d)})$
    - Obtained also by [Potechin, Turner, Venkat, Wein '22] and [Kane, Diakonikolas '22]

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky '13] $m = O(d^{6/5-c})$ for any constant c>0
- Recently, motivated from studies in Sum-of-Squares Algorithms
  - [Ghosh, Jeronimo, Jones, Potechin, Rajendran '20] $m = \tilde{O}(d^{3/2})$
    - Implicitly extends to $m = O(\frac{d^2}{polylog(d)})$
    - Obtained also by [Potechin, Turner, Venkat, Wein '22] and [Kane, Diakonikolas '22]
  - Best-known bound $m = O(\frac{d^2}{log^4(d)})$

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky '13] $m = O(d^{6/5-c})$ for any constant c>0
- Recently, motivated from studies in Sum-of-Squares Algorithms
  - [Ghosh, Jeronimo, Jones, Potechin, Rajendran '20] $m = \tilde{O}(d^{3/2})$
    - Implicitly extends to $m = O(\frac{d^2}{polylog(d)})$
    - Obtained also by [Potechin, Turner, Venkat, Wein '22] and [Kane, Diakonikolas '22]
  - Best-known bound $m = O(\frac{d^2}{log^4(d)})$

Ellipsoid Fitting Conjecture

(Positive side)

Conjecture:

Possible when $m \leq (1+\epsilon)\frac{d^2}{4}$

- [Saunderson, Parrilo, Willsky '13] $m = O(d^{6/5-c})$ for any constant c>0
- Recently, motivated from studies in Sum-of-Squares Algorithms
  - [Ghosh, Jeronimo, Jones, Potechin, Rajendran '20] $m = \tilde{O}(d^{3/2})$
    - Implicitly extends to $m = O(\frac{d^2}{polylog(d)})$
    - Obtained also by [Potechin, Turner, Venkat, Wein '22] and [Kane, Diakonikolas '22]
  - Best-known bound $m = O(\frac{d^2}{log^4(d)})$

Is the polylog dependence tight?

Our result

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$

Our result

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- For $m \leq cd^2$ for some absolute constant c>0, there is a symmetric matrix $\Lambda \in R^{d \times d}$ such that

Our result

- Given m random points $v_1, ..., v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- For $m \leq cd^2$ for some absolute constant c>0, there is a symmetric matrix $\Lambda \in R^{d \times d}$ such that
    - $v_i^T \Lambda v_i = 1$ for all v_i,

Our result

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- For $m \le cd^2$ for some absolute constant c>0, there is a symmetric matrix $\Lambda \in R^{d \times d}$ such that
  - $v_i^T \Lambda v_i = 1$ for all v_i,
  - $\Lambda \succeq 0.$

Our result

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- For $m \leq cd^2$ for some absolute constant c>0, there is a symmetric matrix $\Lambda \in R^{d \times d}$ such that
    - $v_i^T \Lambda v_i = 1$ for all v_i,
    - $\Lambda \succeq 0$.
- **Constant-factor** away from the conjecture

Proof Overview

- High-level idea

Proof Overview

- High-level idea
  - Construct a matrix $\Lambda \in R^{d \times d}$ explicitly from $v_1, \ldots, v_m$

Proof Overview

- High-level idea
    - Construct a matrix $\Lambda \in R^{d \times d}$ explicitly from $v_1, \ldots, v_m$
    - Show $\Lambda \in R^{d \times d}$ satisfies both constraints

Proof Overview

- High-level idea
  - Construct a matrix $\Lambda \in R^{d\times d}$ explicitly from $v_1, \ldots, v_m$
  - Show $\Lambda \in R^{d\times d}$ satisfies both constraints
    - **(Quadratic-form)** $v_i^T \Lambda v_i = 1$ for all v_i,

Proof Overview

- High-level idea
    - Construct a matrix $\Lambda \in R^{d\times d}$ explicitly from $v_1, \ldots, v_m$
    - Show $\Lambda \in R^{d\times d}$ satisfies both constraints
        - **(Quadratic-form)** $v_i^T \Lambda v_i = 1$ for all v_i,
        - **(PSDness)** $\Lambda \succeq 0$.

Proof Overview

- High-level idea
  - Construct a matrix $\Lambda \in R^{d \times d}$ explicitly from $v_1, \ldots, v_m$
  - Show $\Lambda \in R^{d \times d}$ satisfies both constraints
    - **(Quadratic-form)** $v_i^T \Lambda v_i = 1$ for all v_i,
    - **(PSDness)** $\Lambda \succeq 0$.

Overview

1. Ellipsoid Fitting Conjecture
2. **Candidate Construction**
3. Graph Matrices
4. Tight Norm Bounds for Graph Matrices

Construction of ellipsoid

- Given m random points $v_1, ..., v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$

Construction of ellipsoid

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Observation: setting $\Lambda = Id_d$ **"almost"** works

Construction of ellipsoid

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Observation: setting $\Lambda = Id_d$ **“almost”** works
    - $\Lambda = Id_d \succeq 0$ (immediate);

Construction of ellipsoid

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Observation: setting $\Lambda = Id_d$ **"almost"** works
  - $\Lambda = Id_d \succeq 0$ (immediate);
  - For all $i \in [m]$, $v_i^T \Lambda v_i \approx 1$

Construction of ellipsoid

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Observation: setting $\Lambda = Id_d$ **"almost"** works
  - $\Lambda = Id_d \succeq 0$ (immediate);
  - For all $i \in [m]$, $v_i^T \Lambda v_i \approx 1$
    - The quadratic-form constraints are **"approximately"** satisfied

Construction of ellipsoid

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Observation: setting $\Lambda = Id_d$ **"almost"** works
  - $\Lambda = Id_d \succeq 0$ (immediate);
  - For all $i \in [m]$, $v_i^T \Lambda v_i \approx 1$
    - The quadratic-form constraints are **"approximately"** satisfied
    - W.h.p., $|v_i^T \Lambda v_i - 1| \leq \frac{\log d}{\sqrt{d}}$.

Construction of ellipsoid via two-steps

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Starting from Id_d, and correct all the deviations of the quadratic form

Construction of ellipsoid via two-steps

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Starting from Id_d, and correct all the deviations of the quadratic form
    - Our correction term has a nice form $\sum_{a=1}^{m} w_a v_a v_a^T$ for some $w \in \mathbf{R}^m$

Construction of ellipsoid via two-steps

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Starting from Id_d, and correct all the deviations of the quadratic form
  - Our correction term has a nice form $\sum_{a=1}^{m} w_a v_a v_a^T$ for some $w \in \mathbf{R}^m$
  - Final matrix $\Lambda = Id_d - \sum_{a \in [m]} w_a v_a v_a^T$

Construction of ellipsoid via two-steps

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Starting from Id_d, and correct all the deviations of the quadratic form
  - Our correction term has a nice form $\sum_{a=1}^{m} w_a v_a v_a^T$ for some $w \in \mathbf{R}^m$
  - Final matrix $\Lambda = Id_d - \sum_{a \in [m]} w_a v_a v_a^T$

Construction of ellipsoid via two-steps

- Given m random points $v_1, \ldots, v_m \in \mathbf{R}^d$ drawn from drawn from $\mathbf{N}(0, \frac{1}{d} Id_d)$
- Starting from Id_d, and correct all the deviations of the quadratic form
  - Our correction term has a nice form $\sum_{a=1}^{m} w_a v_a v_a^T$ for some $w \in \mathbf{R}^m$
  - Final matrix $\Lambda = Id_d - \sum_{a \in [m]} w_a v_a v_a^T$

How do we pick w?

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T$$

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T$$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T$$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint
- Consider a matrix $M \in \mathbf{R}^{m \times m}$ s.t. $M[i,j] = \langle v_i, v_j \rangle^2$

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T$$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint
- Consider a matrix $M \in \mathbf{R}^{m \times m}$ s.t. $M[i,j] = \langle v_i, v_j \rangle^2$

For any i, $$v_i^T \Lambda v_i = v_i^T Id_d v_i - \sum_{j \in [m]} w_j \langle v_i, v_j \rangle^2$$

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T$$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint
- Consider a matrix $M \in \mathbf{R}^{m \times m}$ s.t. $M[i,j] = \langle v_i, v_j \rangle^2$

For any i, $$v_i^T \Lambda v_i = v_i^T Id_d v_i - \sum_{j \in [m]} w_j \langle v_i, v_j \rangle^2$$

$$= \|v_i\|_2^2 - \langle M[i], w \rangle$$

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T$$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint
- Consider a matrix $M \in \mathbf{R}^{m \times m}$ s.t. $M[i,j] = \langle v_i, v_j \rangle^2$

For any i,

$$v_i^T \Lambda v_i = v_i^T Id_d v_i - \sum_{j \in [m]} w_j \langle v_i, v_j \rangle^2$$

$$= \|v_i\|_2^2 - \langle M[i], w \rangle$$

i-th row of M

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T$$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint
- Consider a matrix $M \in \mathbf{R}^{m \times m}$ s.t. $M[i,j] = \langle v_i, v_j \rangle^2$

For any i, $$v_i^T \Lambda v_i = v_i^T Id_d v_i - \sum_{j \in [m]} w_j \langle v_i, v_j \rangle^2$$

$$= \|v_i\|_2^2 - \langle M[i], w \rangle$$

i-th row of M

If we pick w s.t. $Mw = \eta$

Finding the correction coefficient w

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T$$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint
- Consider a matrix $M \in \mathbf{R}^{m\times m}$ s.t. $M[i,j] = \langle v_i, v_j\rangle^2$

For any i,

$$v_i^T \Lambda v_i = v_i^T Id_d v_i - \sum_{j\in[m]} w_j \langle v_i, v_j\rangle^2$$

$$= \|v_i\|_2^2 - \langle M[i], w\rangle$$

i-th row of M

$$= \|v_i\|_2^2 - (\underbrace{\|v_i\|_2^2 - 1}_{\eta_i})$$

If we pick w s.t. $Mw = \eta$

Finding the correction coefficient w

$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T$

- Let $\eta_i := \|v_i\|_2^2 - 1$ be the deviation of quadratic-form constraint
- Consider a matrix $M \in \mathbf{R}^{m \times m}$ s.t. $M[i,j] = \langle v_i, v_j \rangle^2$

For any i, $$v_i^T \Lambda v_i = v_i^T Id_d v_i - \sum_{j \in [m]} w_j \langle v_i, v_j \rangle^2$$

$$= \|v_i\|_2^2 - \langle M[i], w \rangle$$

i-th row of M

$$= \|v_i\|_2^2 - \underbrace{(\|v_i\|_2^2 - 1)}_{\eta_i}$$

If we pick w s.t. $Mw = \eta$

$= 1$ (**Exactly** satisfying the constraint)

Remaining questions

Picking $w \in \mathbf{R}^m$ s.t. $Mw = \eta$, and set $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$

Remaining questions

Construction from [KD '22]

Picking $w \in \mathbf{R}^m$ s.t. $Mw = \eta$, and set $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$

Remaining questions

Construction from [KD '22]

Picking $w \in \mathbf{R}^m$ s.t. $Mw = \eta$, and set $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$

1. Why is the vector w well-defined?

Remaining questions

Construction from [KD '22]

Picking $w \in \mathbf{R}^m$ s.t. $Mw = \eta$, and set $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$

1. Why is the vector w well-defined?

2. Why does $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$ satisfy the PSDness constraint?

Remaining questions

Construction from [KD ’22]

Picking $w \in \mathbf{R}^m$ s.t. $Mw = \eta$, and set $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$

1. Why is the vector w well-defined?

2. Why does $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$ satisfy the PSDness constraint?

Remaining questions

Construction from [KD '22]

Picking $w \in \mathbf{R}^m$ s.t. $Mw = \eta$, and set $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$

1. Why is the vector w well-defined?

2. Why does $\Lambda = Id_d - \sum_{a=1}^{m} w_a v_a v_a^T$ satisfy the PSDness constraint?

Both boil down to studying spectral norm of random matrices with polynomial entries

Why norm bounds?

(Vector w is well-defined)

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

Why norm bounds?

(Vector w is well-defined)

- Want to find $w \in \mathbf{R}^m$ s.t. $Mw = \eta$

$M \in \mathbf{R}^{m \times m} : M[i, j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

Why norm bounds?

(Vector w is well-defined)

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Want to find $w \in \mathbf{R}^m$ s.t. $Mw = \eta$
- Sufficient to show M is full-rank, and take $w = M^{-1}\eta$

Why norm bounds?

(Vector w is well-defined)

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Want to find $w \in \mathbf{R}^m$ s.t. $Mw = \eta$
- Sufficient to show M is full-rank, and take $w = M^{-1}\eta$
- Idea: decompose $M = A + B$ such that

Why norm bounds?

(Vector w is well-defined)

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$
$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Want to find $w \in \mathbf{R}^m$ s.t. $Mw = \eta$
- Sufficient to show M is full-rank, and take $w = M^{-1}\eta$
- Idea: decompose $M = A + B$ such that
    - A is a perturbed identity matrix,

Why norm bounds?

(Vector w is well-defined)

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$
$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Want to find $w \in \mathbf{R}^m$ s.t. $Mw = \eta$
- Sufficient to show M is full-rank, and take $w = M^{-1}\eta$
- Idea: decompose $M = A + B$ such that
  - A is a perturbed identity matrix,
  - $B = UCV$ is a rank-2 component

Why norm bounds?

(Vector w is well-defined)

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$
$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Want to find $w \in \mathbf{R}^m$ s.t. $Mw = \eta$
- Sufficient to show M is full-rank, and take $w = M^{-1}\eta$
- Idea: decompose $M = A + B$ such that
    - A is a perturbed identity matrix,
    - $B = UCV$ is a rank-2 component
- Woodbury Expansion:

$$(A + UCV)^{-1} = A^{-1} + A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$$

Why norm bounds?

(Vector w is well-defined)

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Want to find $w \in \mathbf{R}^m$ s.t. $Mw = \eta$
- Sufficient to show M is full-rank, and take $w = M^{-1}\eta$
- Idea: decompose $M = A + B$ such that
  - A is a perturbed identity matrix,
  - $B = UCV$ is a rank-2 component

Our focus: showing $A = Id + E$ for $\|E\| < 1$

- Woodbury Expansion:

$$(A + UCV)^{-1} = A^{-1} + A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$$

Why norm bounds?

Studying A^{-1}

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

Why norm bounds?

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

Studying A^{-1}

- Neumann Series: $(Id - T)^{-1} = \sum_{k \geq 0} T^k$ provided $||T||_{sp} < 1$

Why norm bounds?

Studying A^{-1}

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Neumann Series: $(Id - T)^{-1} = \sum_{k \geq 0} T^k$ provided $||T||_{sp} < 1$

$\|T\|_{sp} := \max_{v:\|v\|_2=1} \|Tv\|_2^2$

Why norm bounds?

Studying A^{-1}

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Neumann Series: $(Id - T)^{-1} = \sum_{k\geq 0} T^k$ provided $||T||_{sp} < 1$

$\|T\|_{sp} := \max_{v:\|v\|_2=1} \|Tv\|_2^2$

- $A \approx Id + A_1 + A_2$ (ignoring components with o(1) norm)

Why norm bounds?

Studying A^{-1}

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Neumann Series: $(Id - T)^{-1} = \sum_{k\geq 0} T^k$ provided $||T||_{sp} < 1$

$||T||_{sp} := \max_{v:\|v\|_2=1} \|Tv\|_2^2$

- $A \approx Id + A_1 + A_2$ (ignoring components with o(1) norm)
    - $A_1[a,b] = \sum_{i\neq j\in[d]} v_a[i]\cdot v_b[i]\cdot v_a[j]\cdot v_b[j]$

Why norm bounds?

Studying A^{-1}

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- Neumann Series: $(Id - T)^{-1} = \sum_{k \geq 0} T^k$ provided $\|T\|_{sp} < 1$

$\|T\|_{sp} := \max_{v : \|v\|_2 = 1} \|Tv\|_2^2$

- $A \approx Id + A_1 + A_2$ (ignoring components with o(1) norm)
  - $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$
  - $A_2[a,b] = \sum_{i \in [d]} (v_a[i]^2 - \frac{1}{d})(v_b[i]^2 - \frac{1}{d})$

Why norm bounds?

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^{m} : \eta_i := \|v_i\|_2^2 - 1$

Studying A^{-1}

- Neumann Series: $(Id - T)^{-1} = \sum_{k \geq 0} T^k$ provided $\|T\|_{sp} < 1$

$\|T\|_{sp} := \max_{v : \|v\|_2 = 1} \|Tv\|_2^2$

- $A \approx Id + A_1 + A_2$ (ignoring components with o(1) norm)
  - $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$
  - $A_2[a,b] = \sum_{i \in [d]} (v_a[i]^2 - \frac{1}{d})(v_b[i]^2 - \frac{1}{d})$

Bounding spectral norm of random matrices!

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^{m} : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T \succeq 0$$

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j\rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T \succeq 0$$

Main Theorem: for $m \leq cd^2$, $\|\sum_{a\in[m]} w_a v_a v_a^T\|_{sp} < 1$

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j\rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T \succeq 0$$

Main Theorem: for $m \leq cd^2$, $\|\sum_{a\in[m]} w_a v_a v_a^T\|_{sp} < 1$

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a \in [m]} w_a v_a v_a^T \succeq 0$$

Main Theorem: for $m \leq cd^2$, $\| \sum_{a \in [m]} w_a v_a v_a^T \|_{sp} < 1$

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j\rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T \succeq 0$$

Main Theorem: for $m \leq cd^2$, $\|\sum_{a\in[m]} w_a v_a v_a^T\|_{sp} < 1$

- Implies PSDness by a triangle inequality.

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j\rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T \succeq 0$$

Main Theorem: for $m \leq cd^2$, $\|\sum_{a\in[m]} w_a v_a v_a^T\|_{sp} < 1$

- Implies PSDness by a triangle inequality.

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j\rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T \succeq 0$$

Main Theorem: for $m \leq cd^2$, $\|\sum_{a\in[m]} w_a v_a v_a^T\|_{sp} < 1$

- Implies PSDness by a triangle inequality.

Again, bounding spectral norm of random matrices!

Why norm bounds?

Showing PSDness of Λ

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- For the given w obtained by solving $Mw = \eta$, we need to show

$$\Lambda = Id_d + \sum_{a\in[m]} w_a v_a v_a^T \succeq 0$$

Main Theorem: for $m \le cd^2$, $\| \sum_{a\in[m]} w_a v_a v_a^T \|_{sp} < 1$

- Implies PSDness by a triangle inequality.

Again, bounding spectral norm of random matrices!

Warning: w has complicated dependences on $\{v_i\}$

Overview

1. Ellipsoid Fitting Conjecture
2. Constructing an Ellipsoid
3. **Analysis via Graph Matrices**
4. A Local Machinery for Tight Norm Bounds

Random Matrix with Polynomial Entries: Graph Matrix

Random Matrix with Polynomial Entries: Graph Matrix

- General set-up:

Random Matrix with Polynomial Entries: Graph Matrix

- General set-up:
  - Assume there is some **underlying input,** eg. $v_1, \ldots, v_m \sim \mathbf{N}(0, Id_d)$

Random Matrix with Polynomial Entries: Graph Matrix

- General set-up:
  - Assume there is some **underlying input,** eg. $v_1, \ldots, v_m \sim \mathbf{N}(0, Id_d)$
  - A matrix **M**: each entry is a polynomial of inputs in $\{v_i\}$

Random Matrix with Polynomial Entries: Graph Matrix

- General set-up:
  - Assume there is some **underlying input,** eg. $v_1, ..., v_m \sim \mathbf{N}(0, Id_d)$
  - A matrix **M**: each entry is a polynomial of inputs in $\{v_i\}$
- A baby example:

Random Matrix with Polynomial Entries: Graph Matrix

- General set-up:
  - Assume there is some **underlying input,** eg. $v_1, \ldots, v_m \sim \mathbf{N}(0, Id_d)$
  - A matrix **M**: each entry is a polynomial of inputs in $\{v_i\}$
- A baby example:
  - $M \in R^{m \times d} : M[i, a] = v_i[a]$ is a linear function of $\{v_i\}$

Random Matrix with Polynomial Entries: Graph Matrix

- General set-up:
    - Assume there is some **underlying input,** eg. $v_1, \ldots, v_m \sim \mathbf{N}(0, Id_d)$
    - A matrix **M**: each entry is a polynomial of inputs in $\{v_i\}$
- A baby example:
    - $M \in R^{m \times d} : M[i,a] = v_i[a]$ is a linear function of $\{v_i\}$
- Backbone for the recent progress in average-case Sum-of-Squares Lower Bounds

Random Matrix with Polynomial Entries: Graph Matrix

- General set-up:
  - Assume there is some **underlying input,** eg. $v_1, \ldots, v_m \sim \mathbf{N}(0, Id_d)$
  - A matrix **M**: each entry is a polynomial of inputs in $\{v_i\}$
- A baby example:
  - $M \in R^{m \times d} : M[i,a] = v_i[a]$ is a linear function of $\{v_i\}$
- Backbone for the recent progress in average-case Sum-of-Squares Lower Bounds
  - eg. Planted Clique, Sparse Independent Set, Densest-k-Subgraph…

Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$

$a, b \in [m], i, j \in [d]$

Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$

$a, b \in [m], i, j \in [d]$

  - Each entry is a degree-4 polynomial of v_i

Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $$A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$$

$a, b \in [m], i, j \in [d]$

  - Each entry is a degree-4 polynomial of v_i
  - Represented by a "graph"

Zooming into $A_1 \in \mathbf{R}^{m\times m}$

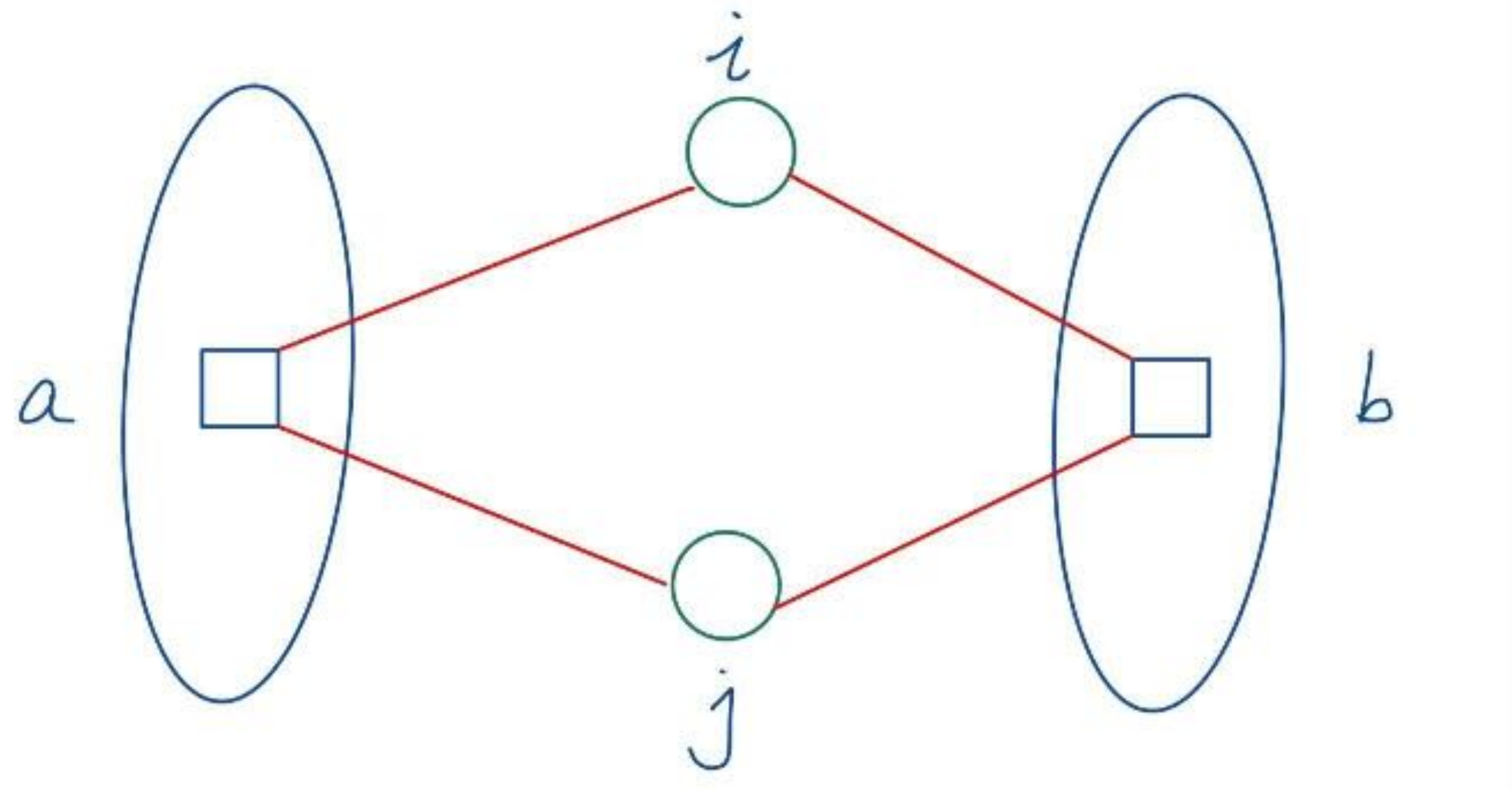
$M \in \mathbf{R}^{m\times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$

$a, b \in [m], i, j \in [d]$

- Each entry is a degree-4 polynomial of v_i
- Represented by a “graph”



Zooming into $A_1 \in \mathbf{R}^{m \times m}$

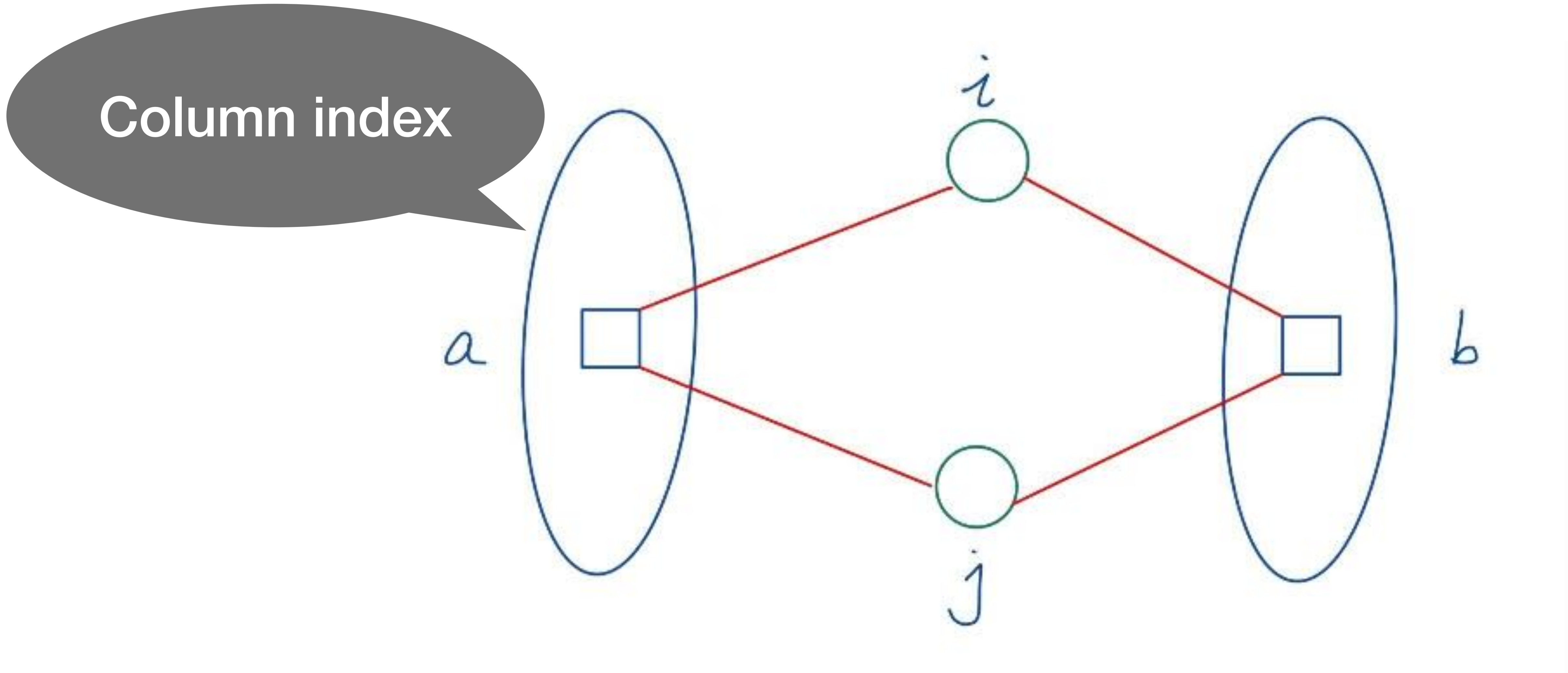
$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^{m} : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$

$a, b \in [m], i, j \in [d]$

- Each entry is a degree-4 polynomial of v_i
- Represented by a "graph"



Zooming into $A_1 \in \mathbf{R}^{m \times m}$

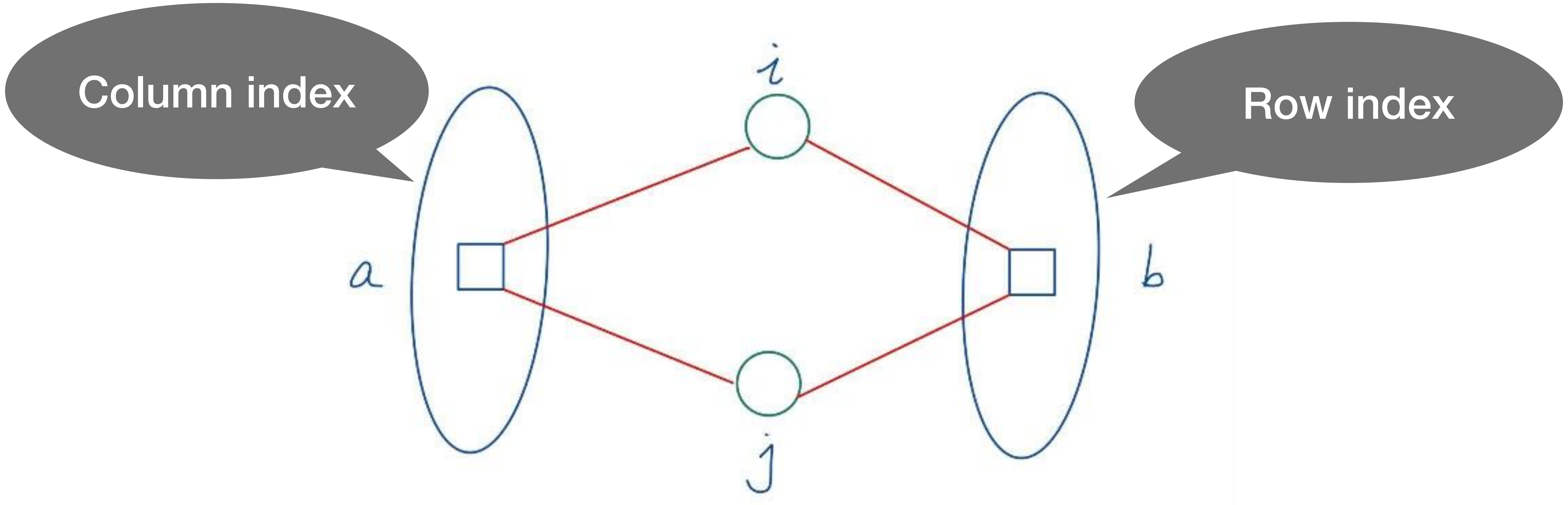
$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$

$a, b \in [m], i, j \in [d]$

- Each entry is a degree-4 polynomial of v_i
- Represented by a "graph"



Zooming into $A_1 \in \mathbf{R}^{m \times m}$

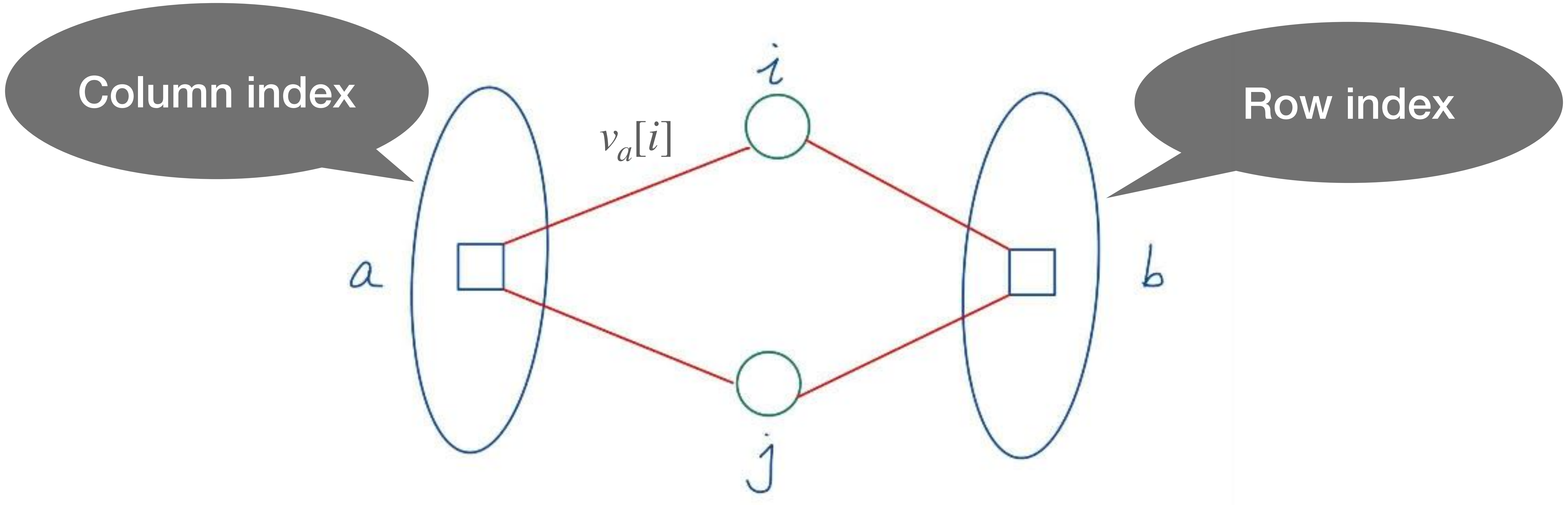
$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $$A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$$

$a, b \in [m], i, j \in [d]$

  - Each entry is a degree-4 polynomial of v_i
  - Represented by a "graph"



Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m \times m} : M[i, j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $$A_1[a, b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$$

    - Each entry is a degree-4 polynomial of v_i

Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$
    - Each entry is a degree-4 polynomial of v_i
    - Correlations across entries

Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$
    - Each entry is a degree-4 polynomial of v_i
    - Correlations across entries
        - We only have md-bits of randomness from $\{v_i\}$

Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $$A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$$
  - Each entry is a degree-4 polynomial of v_i
  - Correlations across entries
    - We only have md-bits of randomness from $\{v_i\}$
    - This is a matrix of $m^2 \gg md$ entries. **Correlation is inevitable!**

Zooming into $A_1 \in \mathbf{R}^{m \times m}$

$M \in \mathbf{R}^{m \times m} : M[i,j] = \langle v_i, v_j \rangle^2$

$\eta \in \mathbf{R}^m : \eta_i := \|v_i\|_2^2 - 1$

- $A_1[a,b] = \sum_{i \neq j \in [d]} v_a[i] \cdot v_b[i] \cdot v_a[j] \cdot v_b[j]$
  - Each entry is a degree-4 polynomial of v_i
  - Correlations across entries
    - We only have md-bits of randomness from $\{v_i\}$
    - This is a matrix of $m^2 \gg md$ entries. **Correlation is inevitable!**

Most results from RMT do not apply!

Spectral Norm Bounds for Graph Matrix

Spectral Norm Bounds for Graph Matrix

- Previous bounds are lossy in **polylog(d, m)** factors

Spectral Norm Bounds for Graph Matrix

- Previous bounds are lossy in **polylog(d, m)** factors
- Our main technical work

Spectral Norm Bounds for Graph Matrix

- Previous bounds are lossy in **polylog(d, m)** factors
- Our main technical work
    - For the matrices that arise in our analysis, we give norm bounds that are **tight up to an absolute constant**

Spectral Norm Bounds for Graph Matrix

- Previous bounds are lossy in **polylog(d, m)** factors
- Our main technical work
  - For the matrices that arise in our analysis, we give norm bounds that are **tight up to an absolute constant**
  - Therefore, we are able to improve previous bounds

Spectral Norm Bounds for Graph Matrix

- Previous bounds are lossy in **polylog(d, m)** factors
- Our main technical work
  - For the matrices that arise in our analysis, we give norm bounds that are **tight up to an absolute constant**
  - Therefore, we are able to improve previous bounds

$$m = O\left(\frac{d^2}{polylog(d)}\right)$$

Spectral Norm Bounds for Graph Matrix

- Previous bounds are lossy in **polylog(d, m)** factors
- Our main technical work
    - For the matrices that arise in our analysis, we give norm bounds that are **tight up to an absolute constant**
    - Therefore, we are able to improve previous bounds

$$m = O\left(\frac{d^2}{polylog(d)}\right)$$

Spectral Norm Bounds for Graph Matrix

- Previous bounds are lossy in **polylog(d, m)** factors
- Our main technical work
    - For the matrices that arise in our analysis, we give norm bounds that are **tight up to an absolute constant**
    - Therefore, we are able to improve previous bounds

$$m = O\left(\frac{d^2}{polylog(d)}\right) \quad \Rightarrow \quad m \leq cd^2$$

Norm Bound Strategy

Norm Bound Strategy

- High-level strategy from previous works

Norm Bound Strategy

- High-level strategy from previous works
- Trace-moment method: for any matrix M, and any $q \in \mathbf{N}$, $\mathbf{E}[\|M\|_{sp}^{2q}] \leq \mathbf{E}[Tr(MM^T)^q]$ where the expectation is taken over the underlying random input

Norm Bound Strategy

- High-level strategy from previous works
- Trace-moment method: for any matrix M, and any $q \in \mathbf{N}$, $\mathbf{E}[\|M\|_{sp}^{2q}] \leq \mathbf{E}[Tr(MM^T)^q]$ where the expectation is taken over the underlying random input
- Higher-moments to get a w.h.p. bound

Norm Bound Strategy

- High-level strategy from previous works
- Trace-moment method: for any matrix M, and any $q \in \mathbf{N}$, $\mathbf{E}[\|M\|_{sp}^{2q}] \leq \mathbf{E}[Tr(MM^T)^q]$ where the expectation is taken over the underlying random input
- Higher-moments to get a w.h.p. bound
- Applying for A_1:

Norm Bound Strategy

- High-level strategy from previous works
- Trace-moment method: for any matrix M, and any $q \in \mathbf{N}$, $\mathbf{E}[\|M\|_{sp}^{2q}] \leq \mathbf{E}[Tr(MM^T)^q]$ where the expectation is taken over the underlying random input
- Higher-moments to get a w.h.p. bound
- Applying for A_1:

$$\mathbf{E}[\|A_1\|_{sp}^{2q}] \leq \mathbf{E}[(A_1A_1^T)^q] = \mathbf{E}\left[\sum_{S_1, T_1, S_2, T_2, \ldots S_{q-1}, T_{q-1} \in [m[} A_1[S_1, T_1]A_1^T[T_1, S_2] \cdots A_1^T[T_{q-1}, S_1]\right]$$

Norm Bound Strategy

- High-level strategy from previous works
- Trace-moment method: for any matrix M, and any $q \in \mathbf{N}$, $\mathbf{E}[\|M\|_{sp}^{2q}] \leq \mathbf{E}[Tr(MM^T)^q]$ where the expectation is taken over the underlying random input
- Higher-moments to get a w.h.p. bound
- Applying for A_1:

$$\mathbf{E}[\|A_1\|_{sp}^{2q}] \leq \mathbf{E}[(A_1A_1^T)^q] = \mathbf{E}\left[\sum_{S_1, T_1, S_2, T_2, \ldots S_{q-1}, T_{q-1} \in [m[} A_1[S_1, T_1]A_1^T[T_1, S_2] \cdots A_1^T[T_{q-1}, S_1]\right]$$

where the expectation is taken over input $\{v_i\}$

Overview

1. Ellipsoid Fitting Conjecture
2. Constructing an Ellipsoid
3. Analysis via Graph Matrices
4. **A Local Machinery for Tight Norm Bounds**

Basics for Trace-Moment Method

Basics for Trace-Moment Method

- Crucial idea for trace-moment method:

Basics for Trace-Moment Method

- Crucial idea for trace-moment method:
    - each edge is independent -> the expectation **factorizes** over edges;

Basics for Trace-Moment Method

- Crucial idea for trace-moment method:
    - each edge is independent -> the expectation **factorizes** over edges;
    - each edge corresponds to a mean-0 r.v. -> it needs to appear at least **twice** in the walk (o.w. the expectation vanishes).

Basics for Trace-Moment Method

- Crucial idea for trace-moment method:
    - each edge is independent -> the expectation **factorizes** over edges;
    - each edge corresponds to a mean-0 r.v. -> it needs to appear at least **twice** in the walk (o.w. the expectation vanishes).
- Sufficient to get a rough norm bound (that loses polylog factors)

Basics for Trace-Moment Method

- Crucial idea for trace-moment method:
    - each edge is independent -> the expectation **factorizes** over edges;
    - each edge corresponds to a mean-0 r.v. -> it needs to appear at least **twice** in the walk (o.w. the expectation vanishes).
- Sufficient to get a rough norm bound (that loses polylog factors)
- We give a more fine-grained analysis to control the lower-order factors

Basics for Trace-Moment Method

- Crucial idea for trace-moment method:
    - each edge is independent -> the expectation **factorizes** over edges;
    - each edge corresponds to a mean-0 r.v. -> it needs to appear at least **twice** in the walk (o.w. the expectation vanishes).
- Sufficient to get a rough norm bound (that loses polylog factors)
- We give a more fine-grained analysis to control the lower-order factors
    - This is a walk of 2q-steps, we bound the “contribution” from each step

Basics for Trace-Moment Method

- Crucial idea for trace-moment method:
    - each edge is independent -> the expectation **factorizes** over edges;
    - each edge corresponds to a mean-0 r.v. -> it needs to appear at least **twice** in the walk (o.w. the expectation vanishes).
- Sufficient to get a rough norm bound (that loses polylog factors)
- We give a more fine-grained analysis to control the lower-order factors
    - This is a walk of 2q-steps, we bound the "contribution" from each step
    - **Warning: each step may contain multiple edges**

A Taste of Our Local Analysis

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,
    - **F** (forward): uses a new edge, and arrive at a new vertex

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,
  - **F** (forward): uses a new edge, and arrive at a new vertex
  - **S** (surprise): uses a new edge, and arrive at a "seen" vertex

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,
  - **F** (forward): uses a new edge, and arrive at a new vertex
  - **S** (surprise): uses a new edge, and arrive at a "seen" vertex
  - **H** (middle): uses an edge that is making middle appearance (neither first/last)

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,
    - **F** (forward): uses a new edge, and arrive at a new vertex
    - **S** (surprise): uses a new edge, and arrive at a "seen" vertex
    - **H** (middle): uses an edge that is making middle appearance (neither first/last)
    - **R** (return): uses an edge for the last time

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,
    - **F** (forward): uses a new edge, and arrive at a new vertex
    - **S** (surprise): uses a new edge, and arrive at a "seen" vertex
    - **H** (middle): uses an edge that is making middle appearance (neither first/last)
    - **R** (return): uses an edge for the last time
- Each edge gives two types of contributions

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,
  - **F** (forward): uses a new edge, and arrive at a new vertex
  - **S** (surprise): uses a new edge, and arrive at a "seen" vertex
  - **H** (middle): uses an edge that is making middle appearance (neither first/last)
  - **R** (return): uses an edge for the last time
- Each edge gives two types of contributions
  - **Vertex-factor** (of the vertex it leads to): the combinatorial choice from counting

A Taste of Our Local Analysis

- Each edge (random variable)'s appearance in the walk may be,
  - **F** (forward): uses a new edge, and arrive at a new vertex
  - **S** (surprise): uses a new edge, and arrive at a "seen" vertex
  - **H** (middle): uses an edge that is making middle appearance (neither first/last)
  - **R** (return): uses an edge for the last time
- Each edge gives two types of contributions
  - **Vertex-factor** (of the vertex it leads to): the combinatorial choice from counting
  - **Edge-factor**: the analytical factor from expectation of the random variable

A Taste of Our Local Analysis

A Taste of Our Local Analysis

- Assignment of **edge-factor**: for each random variable x,

A Taste of Our Local Analysis

- Assignment of **edge-factor**: for each random variable x,
  - Assign a factor of $\frac{1}{\sqrt{d}}$ for first (F) and last (R) appearance each;

A Taste of Our Local Analysis

- Assignment of **edge-factor**: for each random variable x,
  - Assign a factor of $\frac{1}{\sqrt{d}}$ for first (F) and last (R) appearance each;
  - if appears exactly twice, gives value $\mathbf{E}[x^2] = \frac{1}{d}$;

A Taste of Our Local Analysis

- Assignment of **edge-factor**: for each random variable x,
  - Assign a factor of $\frac{1}{\sqrt{d}}$ for first (F) and last (R) appearance each;
  - if appears exactly twice, gives value $\mathbf{E}[x^2] = \frac{1}{d}$;
  - If appears $t > 2$ times, gives a value at most $\mathbf{E}[x^t] \leq \frac{t^{t/2}}{d^{t/2}}$;

A Taste of Our Local Analysis

- Assignment of **edge-factor**: for each random variable x,
  - Assign a factor of $\frac{1}{\sqrt{d}}$ for first (F) and last (R) appearance each;
  - if appears exactly twice, gives value $\mathbf{E}[x^2] = \frac{1}{d}$;
  - If appears $t > 2$ times, gives a value at most $\mathbf{E}[x^t] \leq \frac{t^{t/2}}{d^{t/2}}$;
  - Assign a factor of $\frac{\sqrt{t}}{\sqrt{d}} \leq \frac{\sqrt{q}}{\sqrt{d}}$ for each middle (non first/last) appearance.

A Taste of Our Local Analysis

A Taste of Our Local Analysis

- Assignment of **vertex-factor**

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
    - Each vertex, when it first appears, contributes a value of d (or m)

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
  - Each vertex, when it first appears, contributes a value of d (or m)
  - We split the factor among its first and last appearance

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
  - Each vertex, when it first appears, contributes a value of d (or m)
  - We split the factor among its first and last appearance
    - That is a factor of $\sqrt{d}$ (or $\sqrt{m}$)

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
    - Each vertex, when it first appears, contributes a value of d (or m)
    - We split the factor among its first and last appearance
        - That is a factor of $\sqrt{d}$ (or $\sqrt{m}$)
    - (Main lemma) For any vertex's subsequent appearance,

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
  - Each vertex, when it first appears, contributes a value of d (or m)
  - We split the factor among its first and last appearance
    - That is a factor of $\sqrt{d}$ (or $\sqrt{m}$)
  - (Main lemma) For any vertex's subsequent appearance,

If it is arrived via an H/S-edge, a cost of q^3 suffices ;

If it is arrived via an R-edge, a cost of 1 suffices.

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
  - Each vertex, when it first appears, contributes a value of d (or m)
  - We split the factor among its first and last appearance
    - That is a factor of $\sqrt{d}$ (or $\sqrt{m}$)
  - (Main lemma) For any vertex's subsequent appearance,

If it is arrived via an H/S-edge, a cost of q^3 suffices ;

If it is arrived via an R-edge, a cost of 1 suffices.

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
  - Each vertex, when it first appears, contributes a value of d (or m)
  - We split the factor among its first and last appearance
    - That is a factor of $\sqrt{d}$ (or $\sqrt{m}$)
  - (Main lemma) For any vertex's subsequent appearance,

If it is arrived via an H/S-edge, a cost of q^3 suffices ;

If it is arrived via an R-edge, a cost of 1 suffices.

Similar lemmas were known before for G.O.E. and adjacency matrix only

Sketch of Main Lemma

Sketch of Main Lemma

- Assume the walk is tree-like (no **S** edge), and there is also no **H** edge

Sketch of Main Lemma

- Assume the walk is tree-like (no **S** edge), and there is also no **H** edge
    - Each time we are at a vertex, it is incident to only **1 F** edge

Sketch of Main Lemma

- Assume the walk is tree-like (no **S** edge), and there is also no **H** edge
    - Each time we are at a vertex, it is incident to only **1 F** edge
    - Hence, we **do not** need any cost to specify where an R edge leads to.

Sketch of Main Lemma

- Assume the walk is tree-like (no **S** edge), and there is also no **H** edge
    - Each time we are at a vertex, it is incident to only **1 F** edge
    - Hence, we **do not** need any cost to specify where an R edge leads to.
- When the walk is not tree-like

Sketch of Main Lemma

- Assume the walk is tree-like (no **S** edge), and there is also no **H** edge
    - Each time we are at a vertex, it is incident to only **1 F** edge
    - Hence, we **do not** need any cost to specify where an R edge leads to.
- When the walk is not tree-like
    - Each S/H edge may incur confusion as a vertex may now be incident to multiple F edges

Sketch of Main Lemma

- Assume the walk is tree-like (no **S** edge), and there is also no **H** edge
  - Each time we are at a vertex, it is incident to only **1 F** edge
  - Hence, we **do not** need any cost to specify where an R edge leads to.
- When the walk is not tree-like
  - Each S/H edge may incur confusion as a vertex may now be incident to multiple F edges
  - Observation: each comes with a **gap** in the vertex-factor, and we can use this gap to encode auxiliary information

Sketch of Main Lemma

- Assume the walk is tree-like (no **S** edge), and there is also no **H** edge
    - Each time we are at a vertex, it is incident to only **1 F** edge
    - Hence, we **do not** need any cost to specify where an R edge leads to.
- When the walk is not tree-like
    - Each S/H edge may incur confusion as a vertex may now be incident to multiple F edges
    - Observation: each comes with a **gap** in the vertex-factor, and we can use this gap to encode auxiliary information
    - Gap: as opposed to going to a new vertex that gives $\sqrt{d}$, a label in q suffices now

A Taste of Our Local Analysis

- Assignment of **vertex-factor**
  - Each vertex, when it first appears, contributes a value of d (or m)
  - We split the factor among its first and last appearance
    - That is a factor of $\sqrt{d}$ (or $\sqrt{m}$)
  - (Main lemma) For any vertex's subsequent appearance,

If it is arrived via an H/S-edge, a cost of q^3 suffices ;

If it is arrived via an R-edge, a cost of 1 suffices.

A Taste of Our Local Analysis

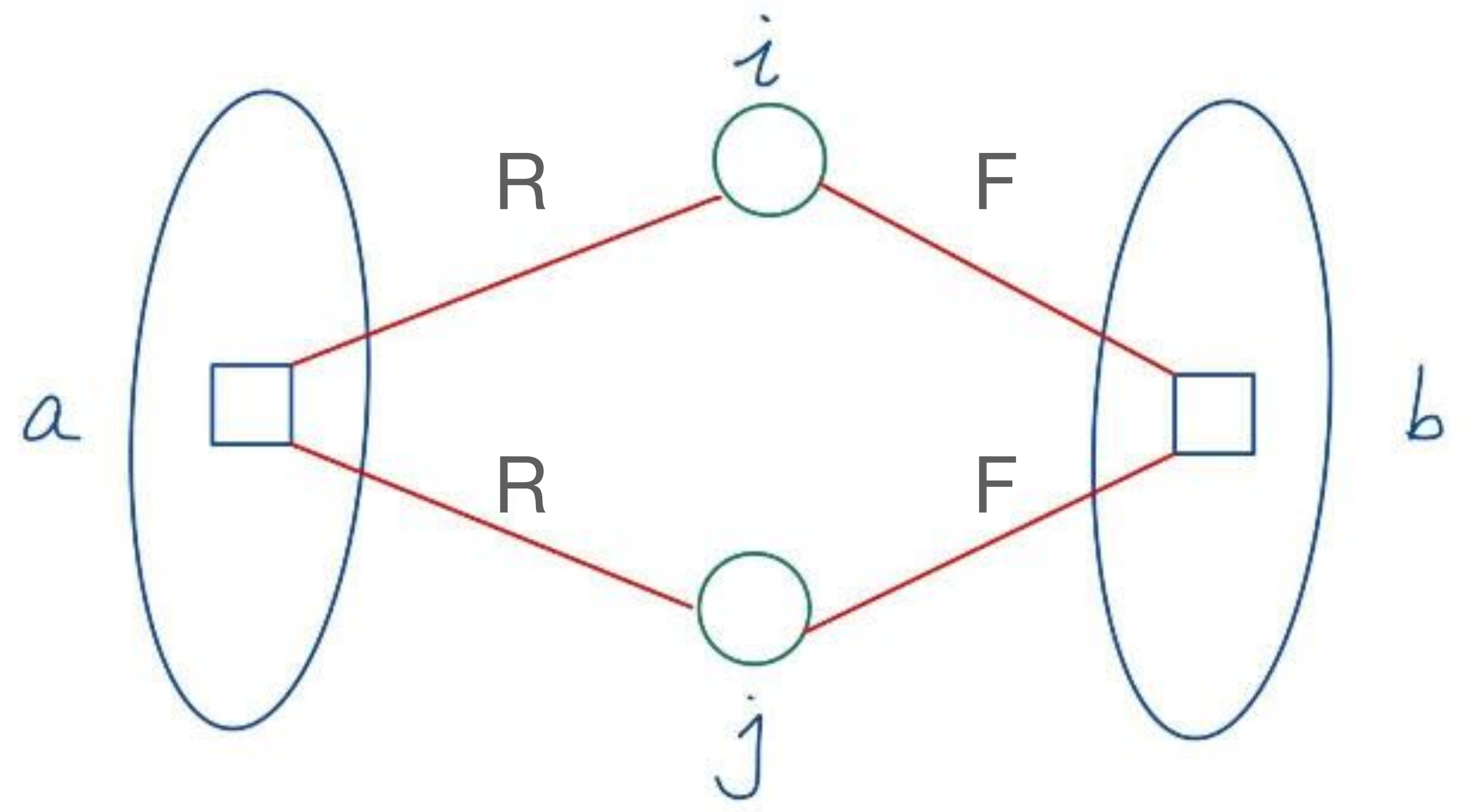
- Assignment of **vertex-factor**
  - Each vertex, when it first appears, contributes a value of d (or m)
  - We split the factor among its first and last appearance
    - That is a factor of $\sqrt{d}$ (or $\sqrt{m}$)
  - (Main lemma) For any vertex's subsequent appearance,

If it is arrived via an H/S-edge, a cost of q^3 suffices ;

If it is arrived via an R-edge, a cost of 1 suffices.

  - Each vertex on the vertex boundary appears in both steps

A Taste of Our Local Analysis



Edge-factors

F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

A Taste of Our Local Analysis

Edge-factors

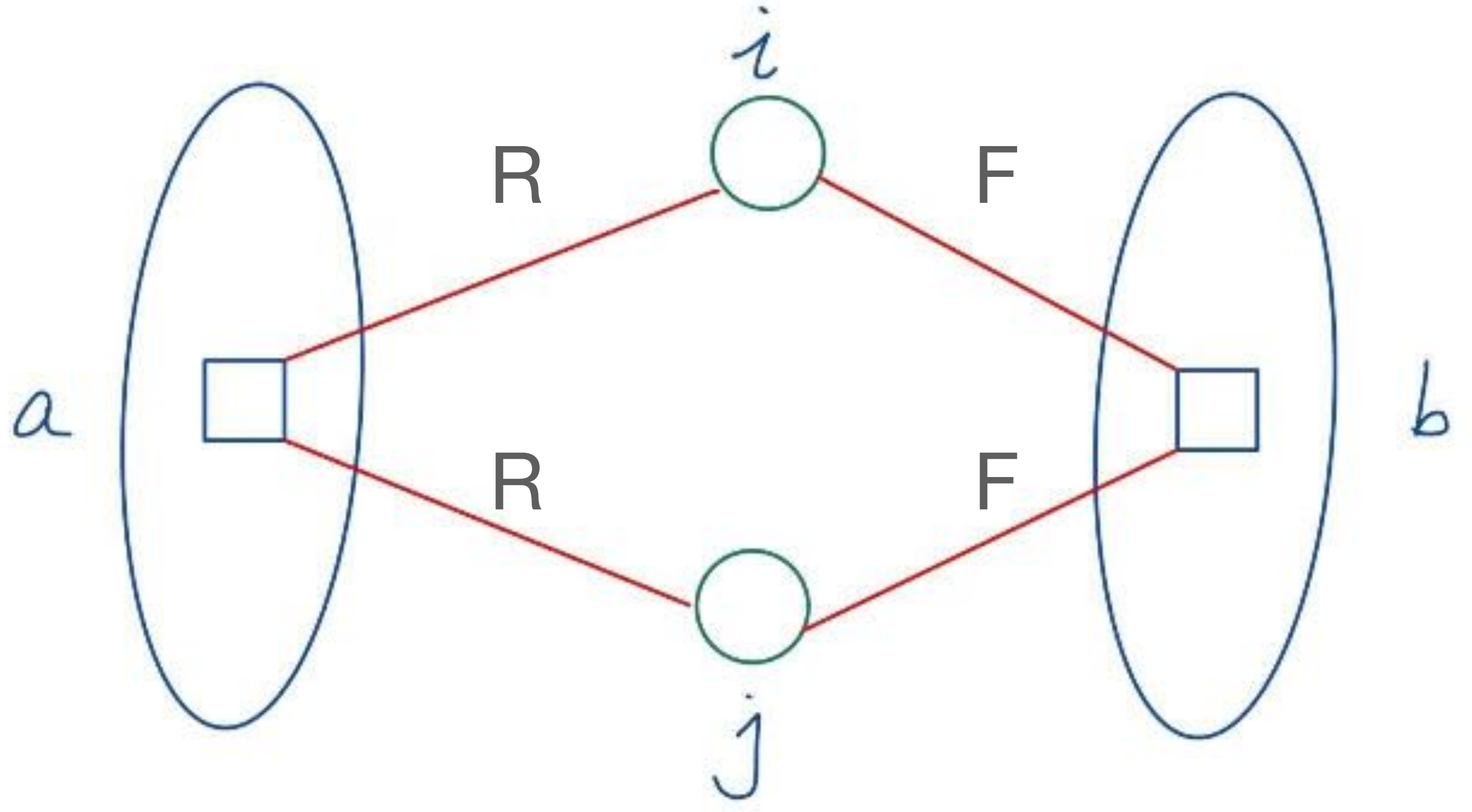
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$

A Taste of Our Local Analysis

Edge-factors

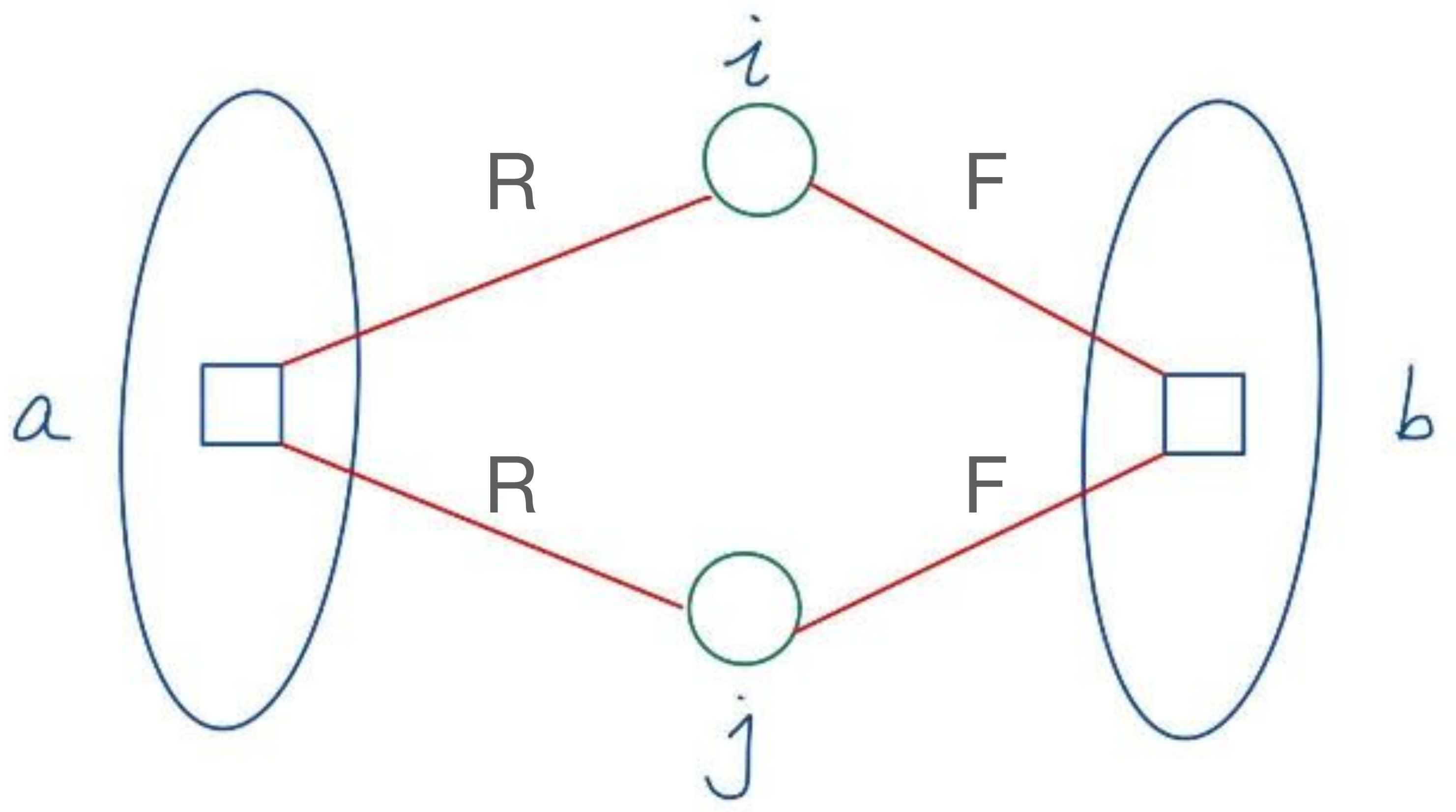
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor

A Taste of Our Local Analysis

Edge-factors

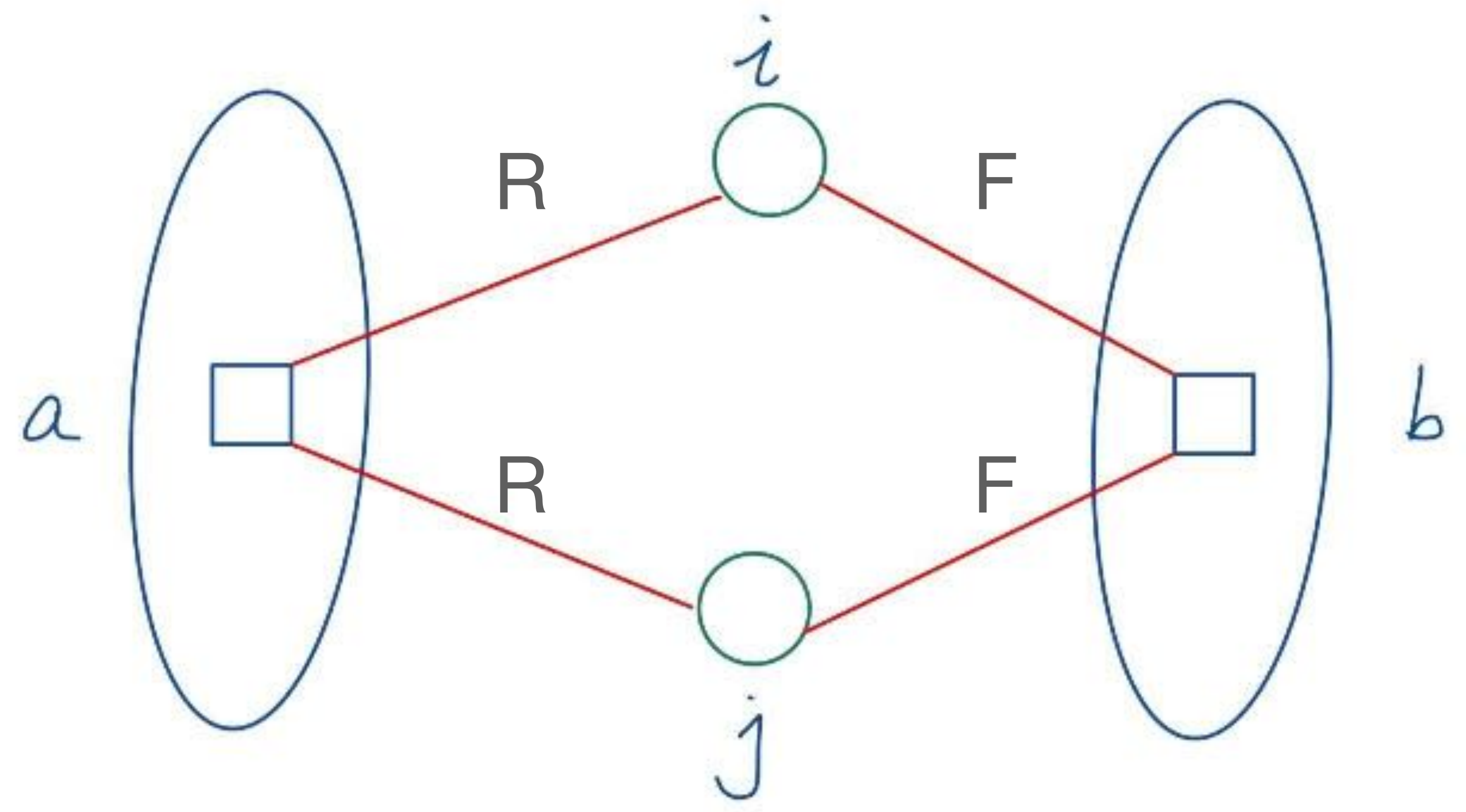
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor
- j is making “middle” appearance, hence no factor

A Taste of Our Local Analysis

Edge-factors

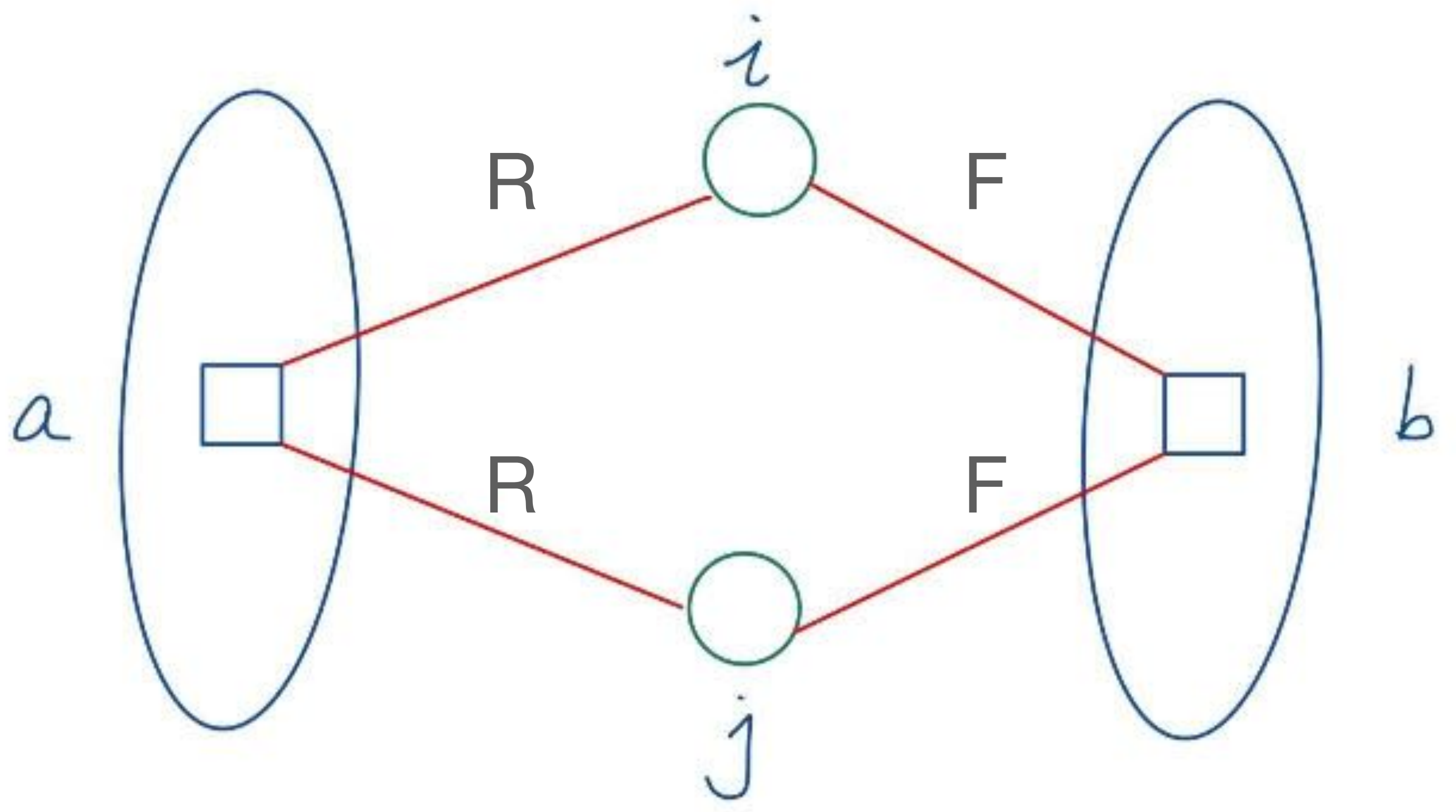
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor
- j is making “middle” appearance, hence no factor

Why middle?

A Taste of Our Local Analysis

Edge-factors

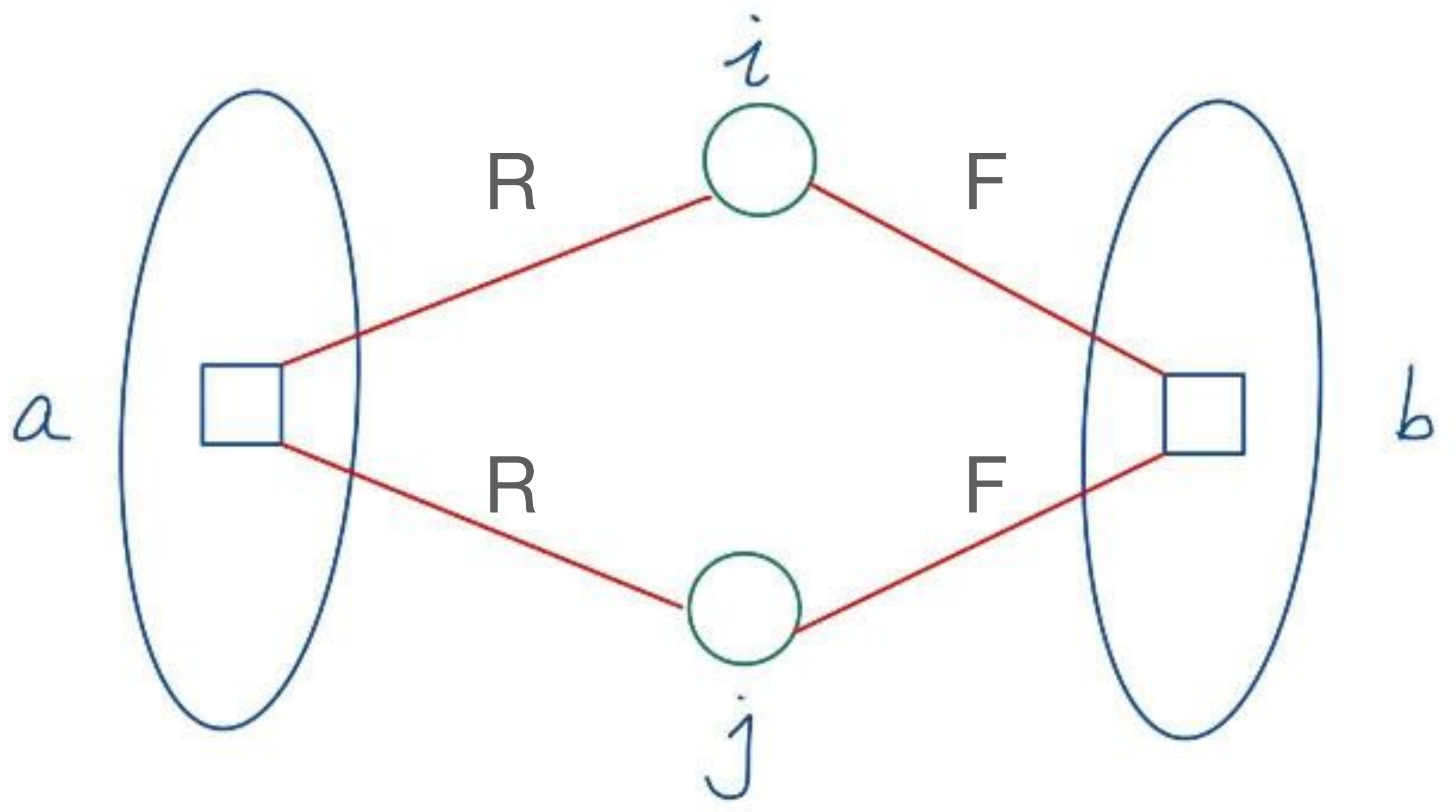
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor
- j is making “middle” appearance, hence no factor

Why middle?

Each edge appears at least two times

A Taste of Our Local Analysis

Edge-factors

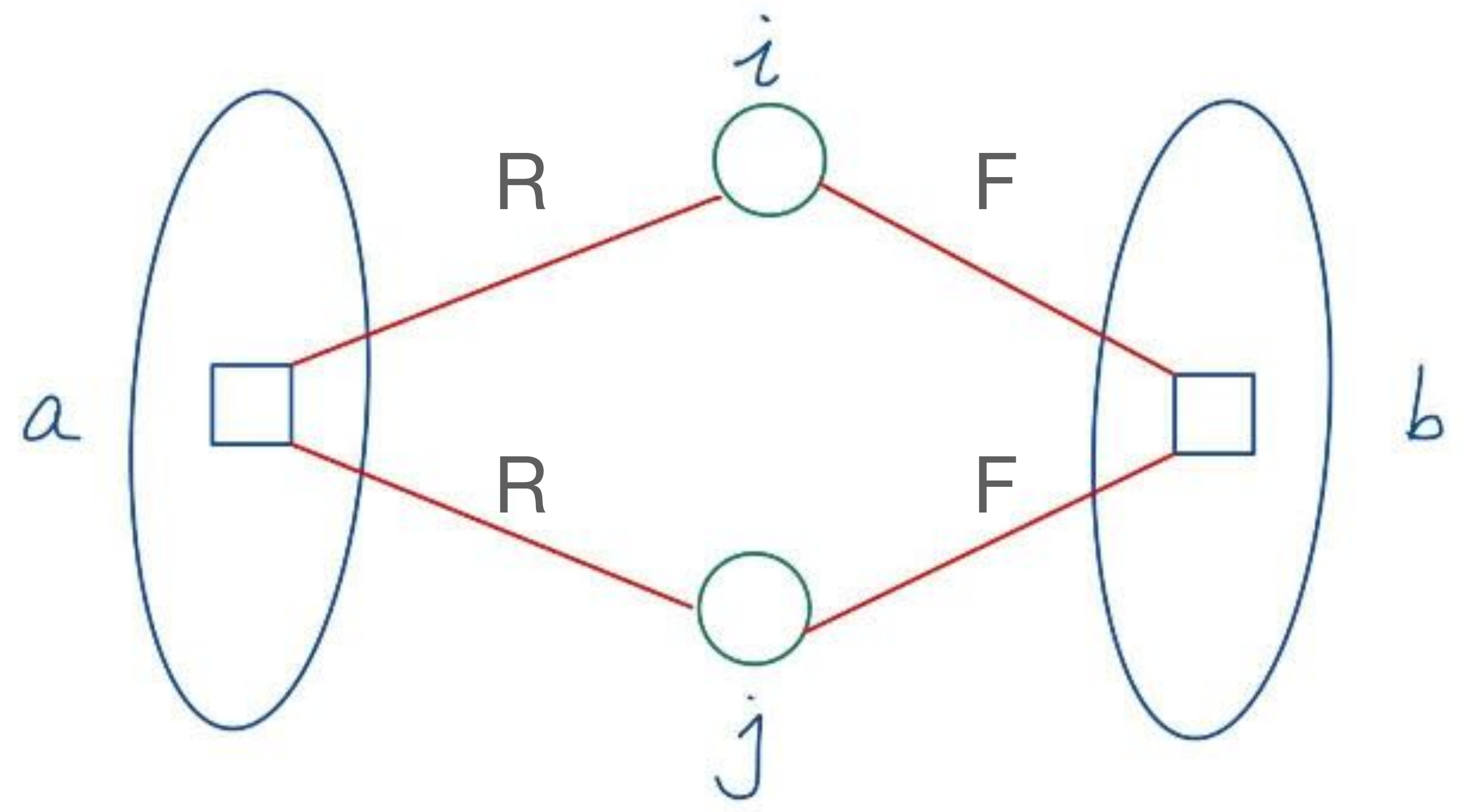
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor
- j is making “middle” appearance, hence no factor
- b is making “first” appearance, but not “last”, hence $\sqrt{m}$

Why middle?

Each edge appears at least two times

A Taste of Our Local Analysis

Edge-factors

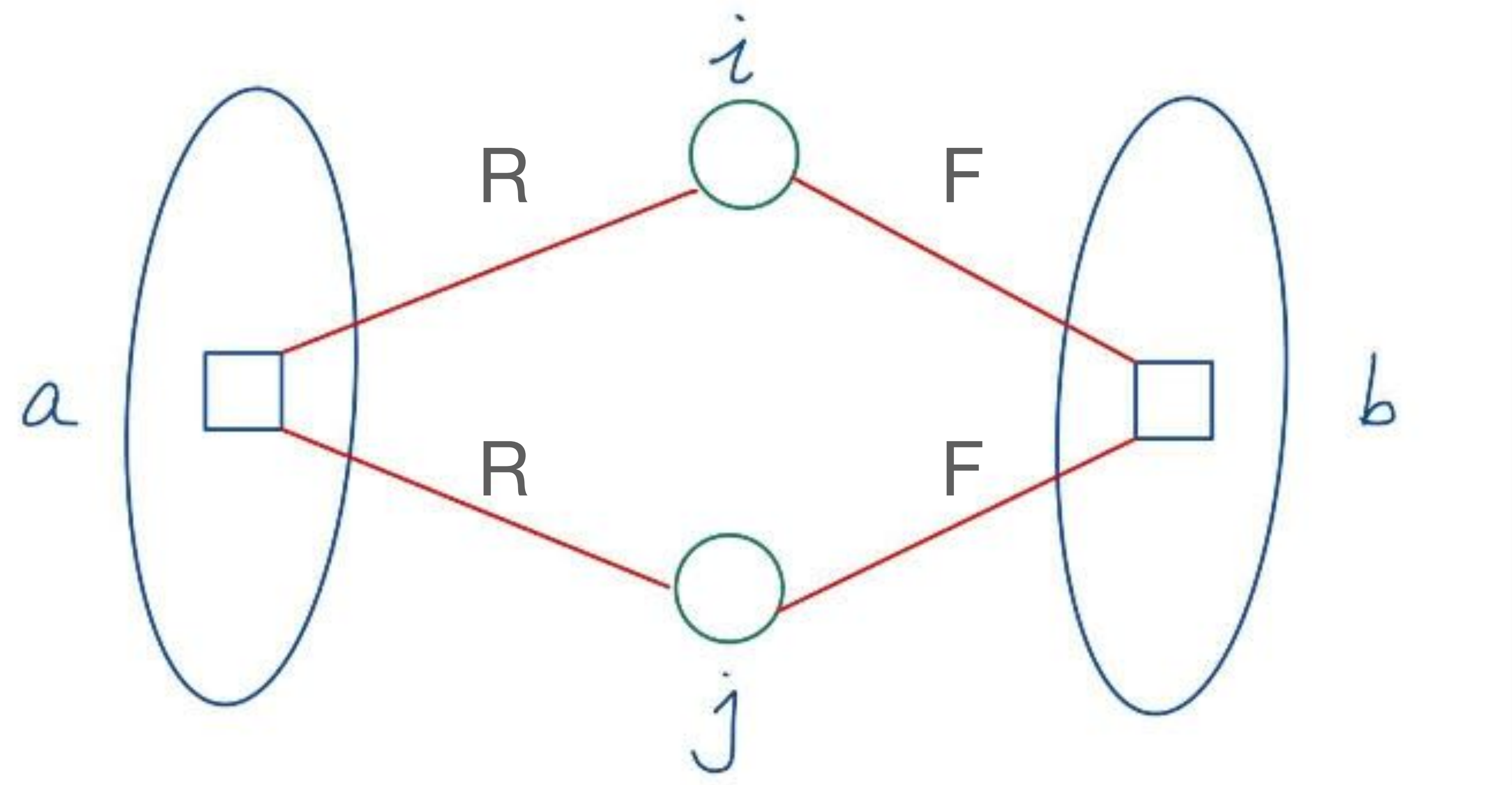
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor
- j is making “middle” appearance, hence no factor
- b is making “first” appearance, but not “last”, hence $\sqrt{m}$
- Edge-value $1/d^2$

Why middle?

Each edge appears at least two times

A Taste of Our Local Analysis

Edge-factors

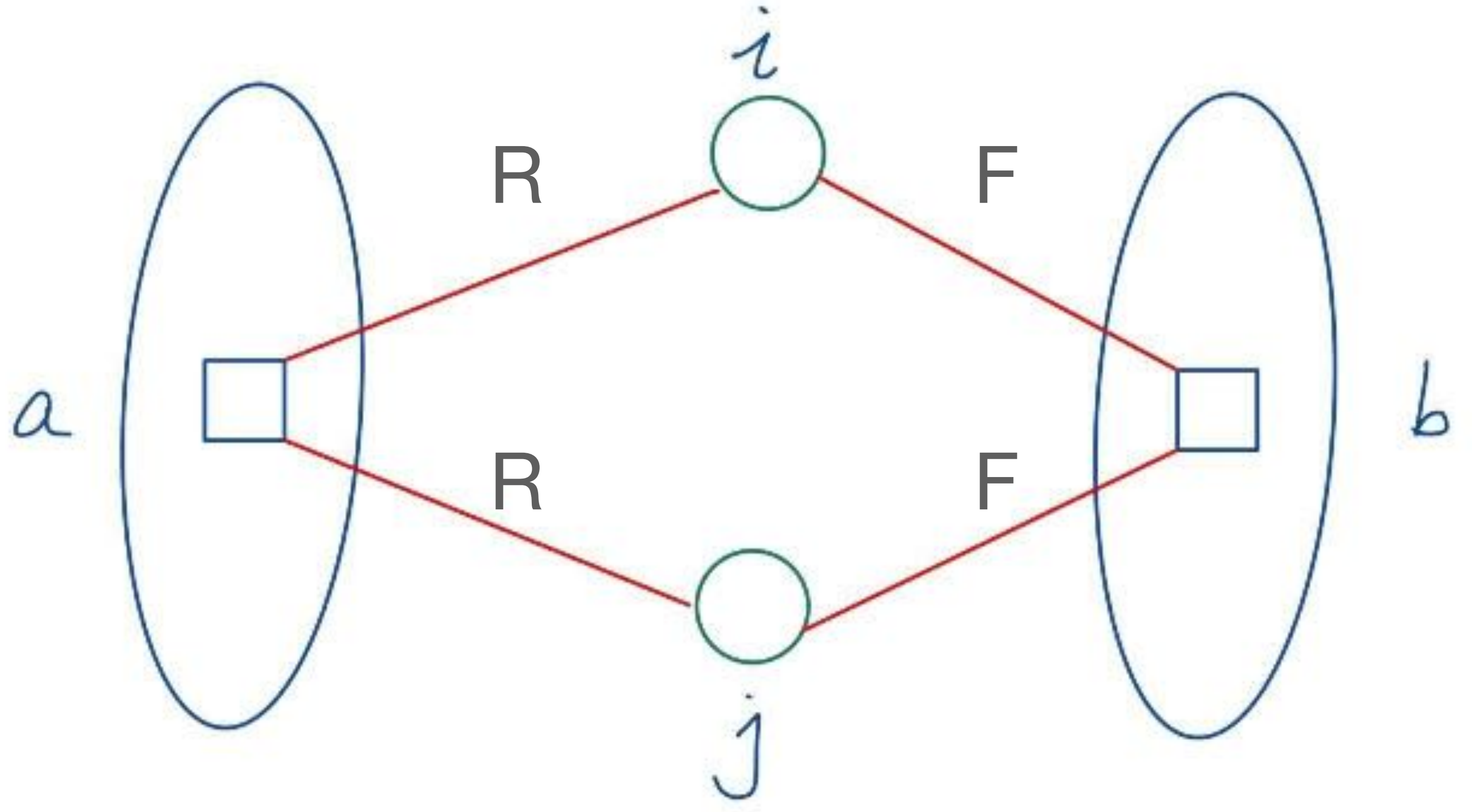
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



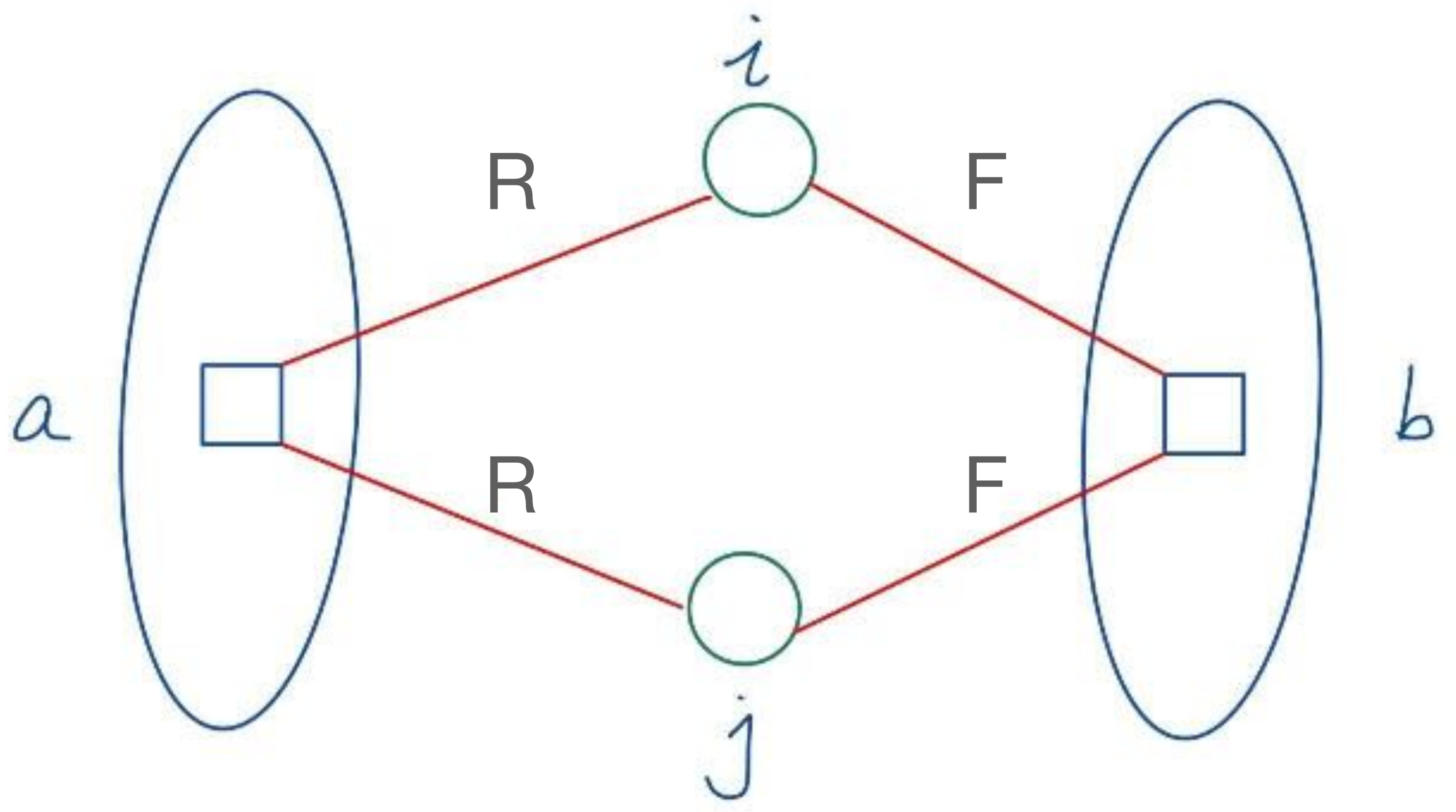
- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor
- j is making “middle” appearance, hence no factor
- b is making “first” appearance, but not “last”, hence $\sqrt{m}$
- Edge-value $1/d^2$

Why middle?

Each edge appears at least two times

A Taste of Our Local Analysis

Edge-factors		Vertex-factors	
F/S/R:	$1/\sqrt{d}$	First/Last:	$\sqrt{d}$ (or $\sqrt{m}$)
H:	$\sqrt{q}/\sqrt{d}$	Arrival via S/H:	q^3



- a can be making “last" appearance, but not “first”, hence $\sqrt{m}$
- i is making “middle” appearance, hence no factor
- j is making “middle” appearance, hence no factor
- b is making “first” appearance, but not “last”, hence $\sqrt{m}$
- Edge-value $1/d^2$

$\Rightarrow$ $m/d^2 \le c^2$

Why middle?

Each edge appears at least two times

A Taste of Our Local Analysis

Edge-factors

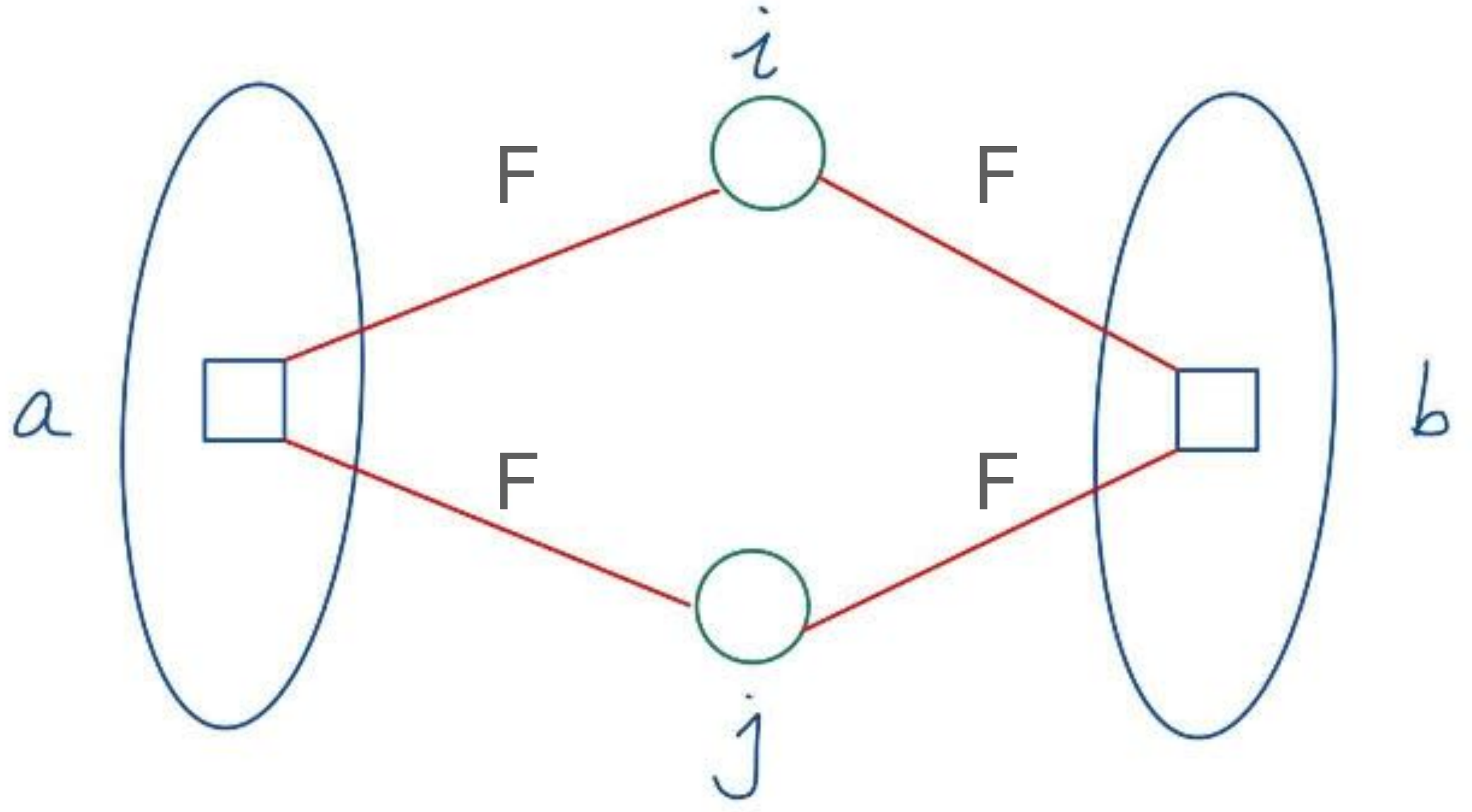
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



A Taste of Our Local Analysis

Edge-factors

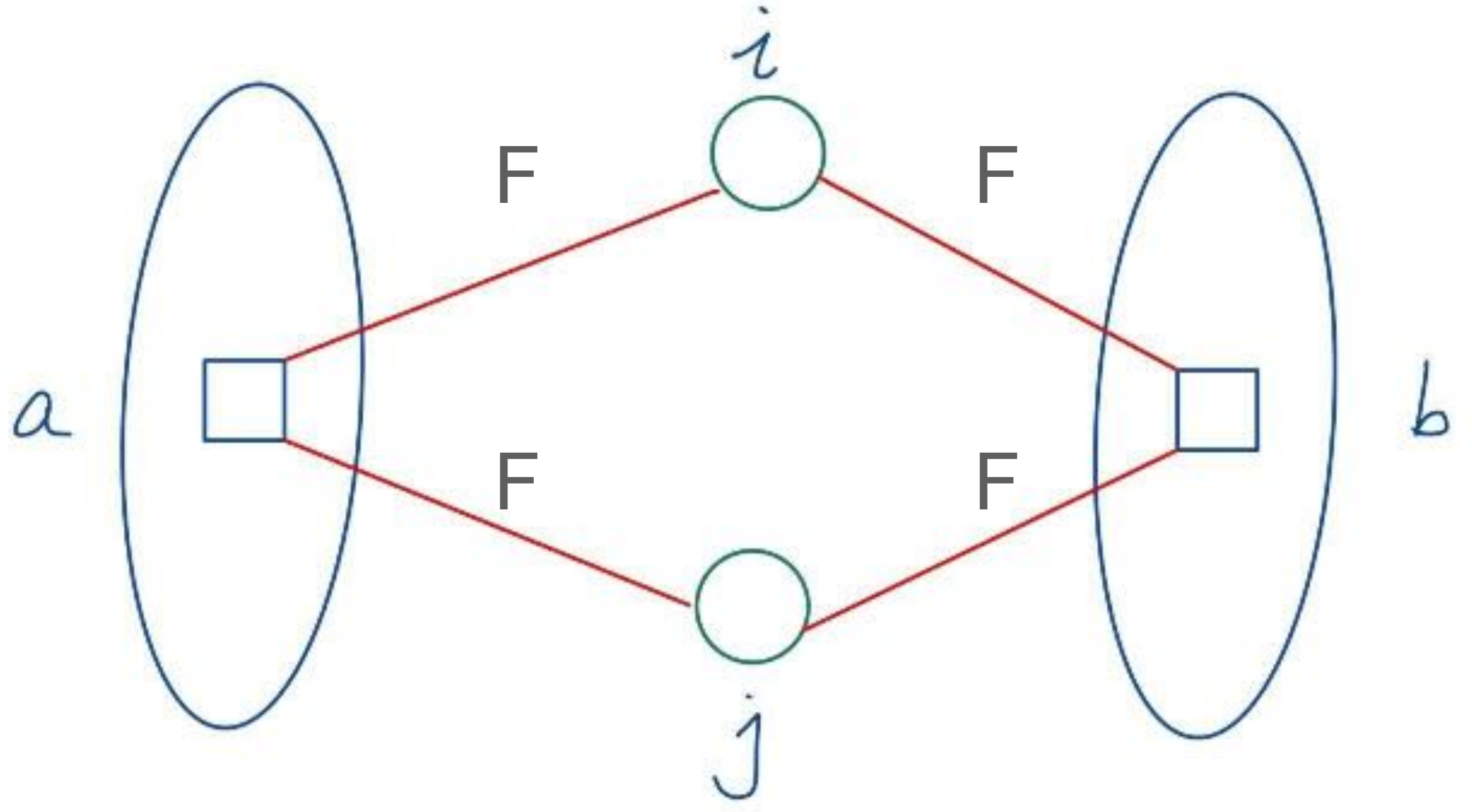
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making “middle” appearance, hence no factor

A Taste of Our Local Analysis

Edge-factors

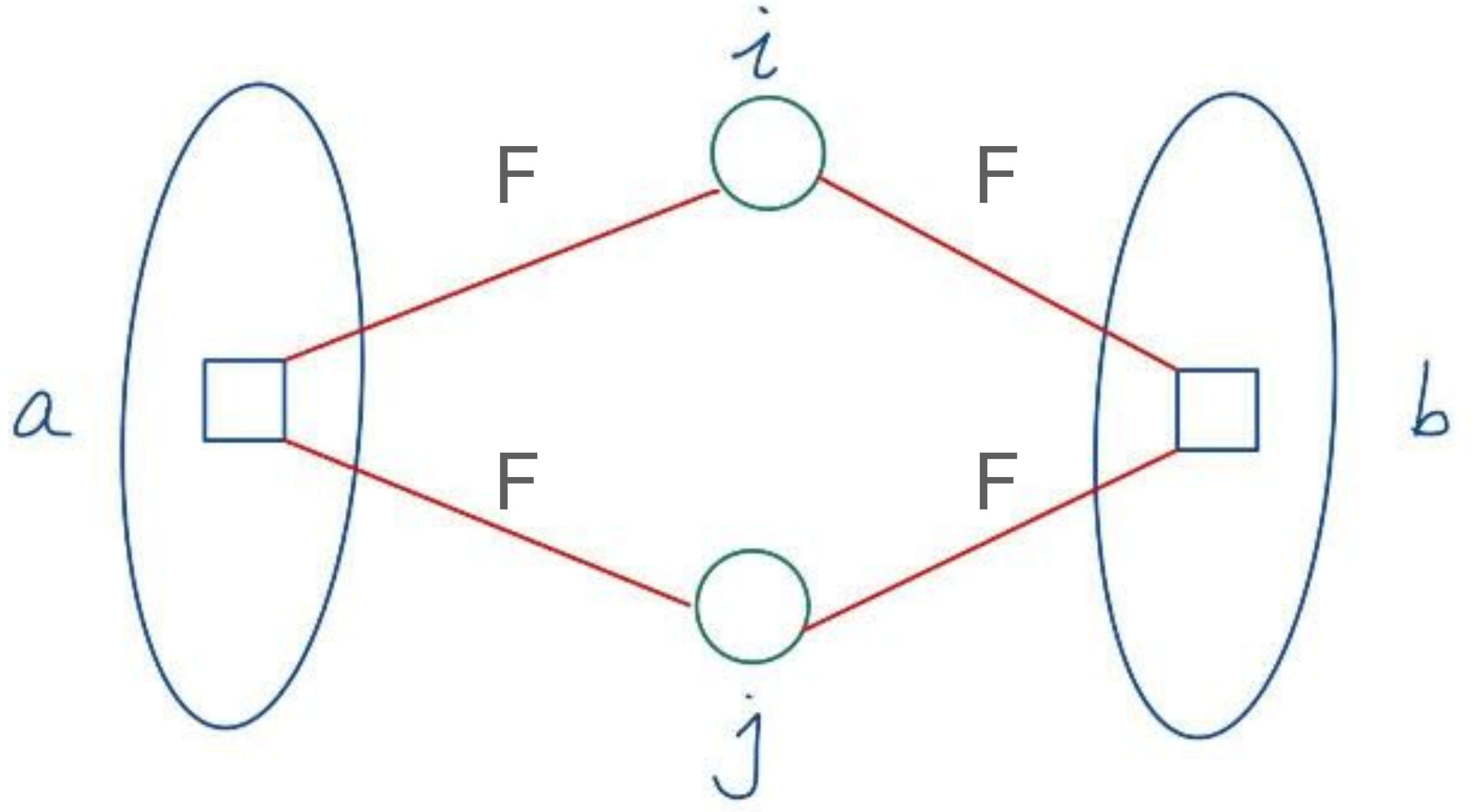
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is making "first" appearance, but not "last", hence $\sqrt{d}$

A Taste of Our Local Analysis

Edge-factors

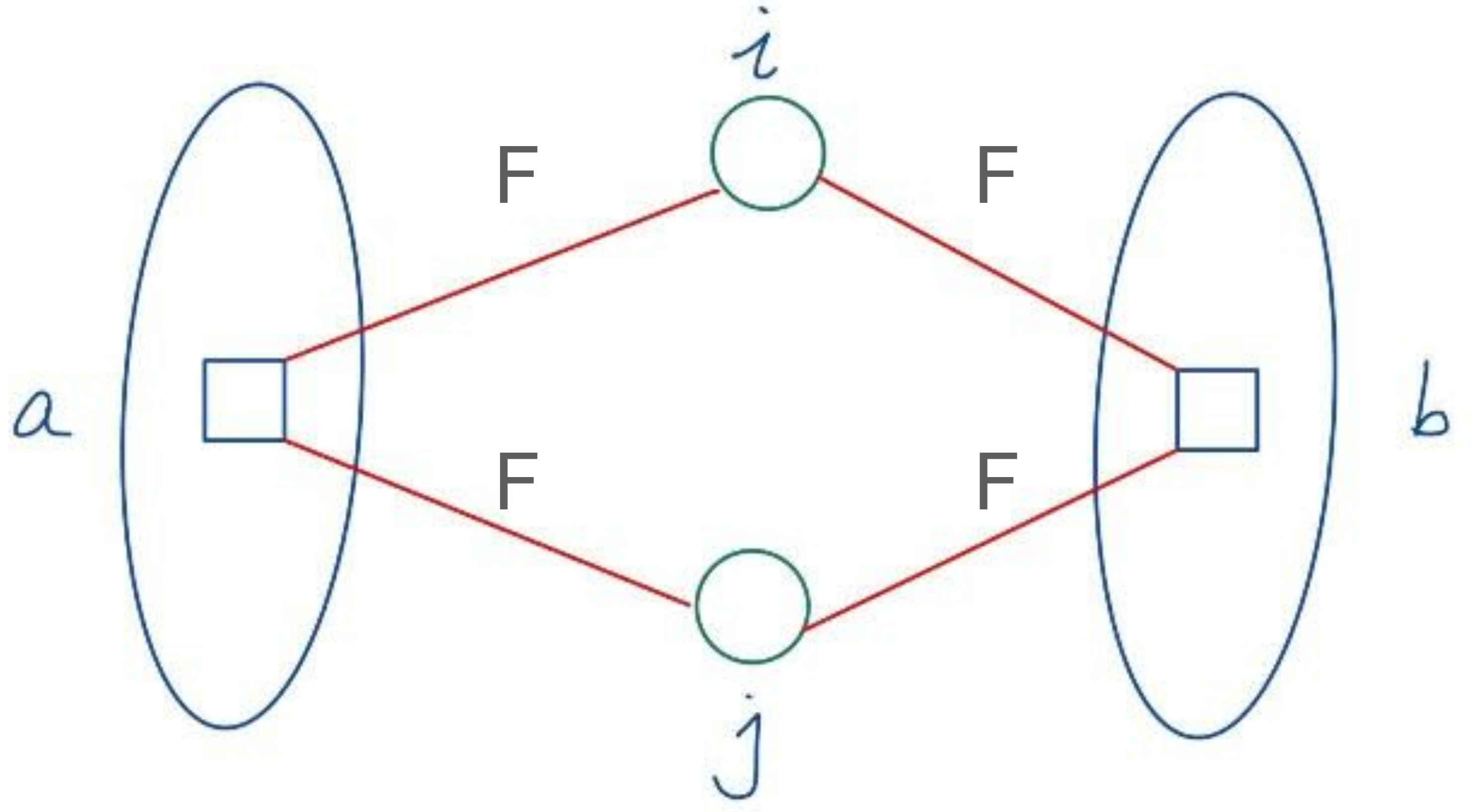
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

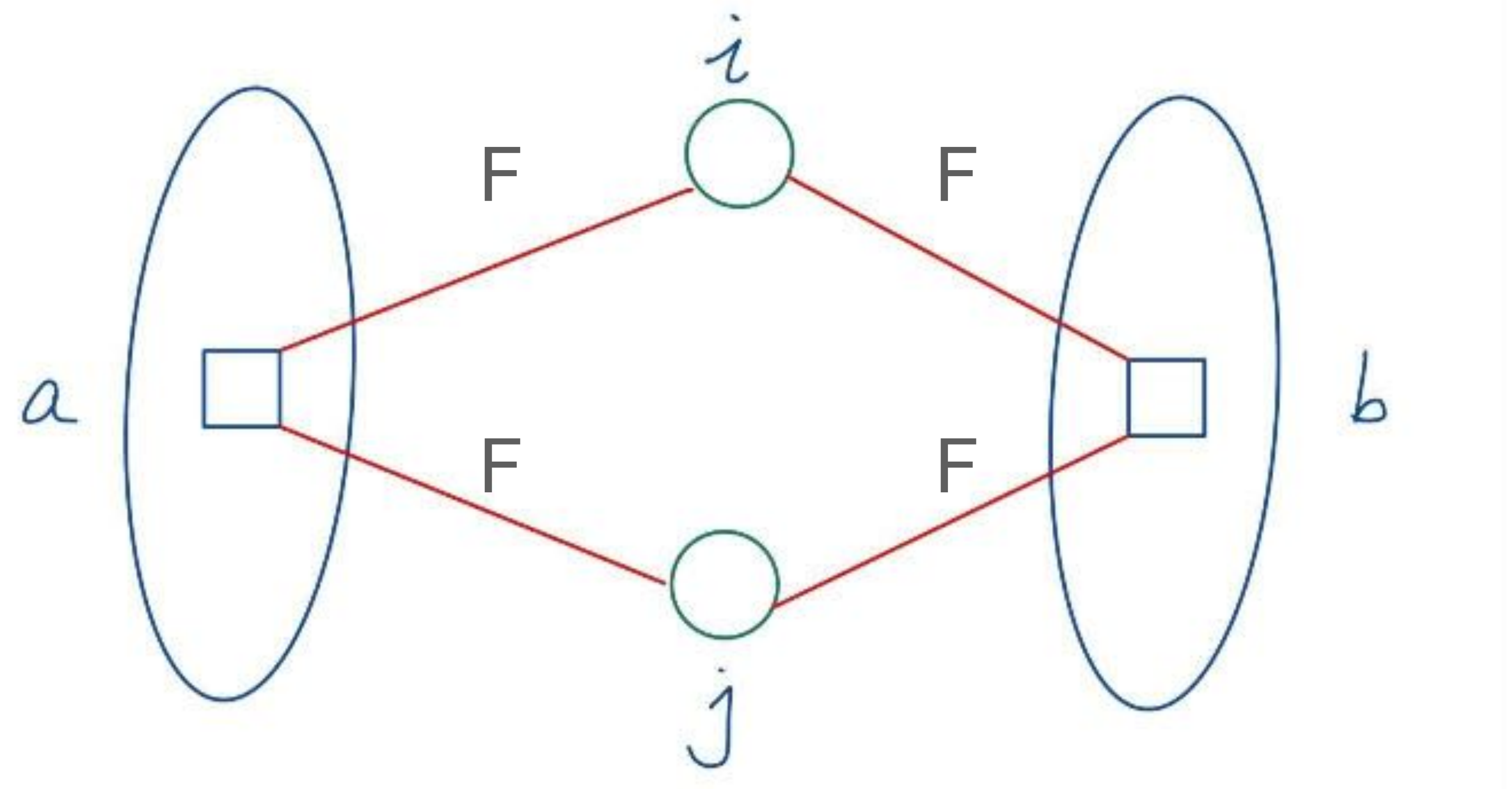


- a is making "middle" appearance, hence no factor
- i is making "first" appearance, but not "last", hence $\sqrt{d}$
- j is making "first" appearance, but not "last", hence $\sqrt{d}$

A Taste of Our Local Analysis

Edge-factors	
F/S/R:	$1/\sqrt{d}$
H:	$\sqrt{q}/\sqrt{d}$

Vertex-factors	
First/Last:	$\sqrt{d}$ (or $\sqrt{m}$)
Arrival via S/H:	q^3



- a is making "middle" appearance, hence no factor
- i is making "first" appearance, but not "last", hence $\sqrt{d}$
- j is making "first" appearance, but not "last", hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$

A Taste of Our Local Analysis

Edge-factors

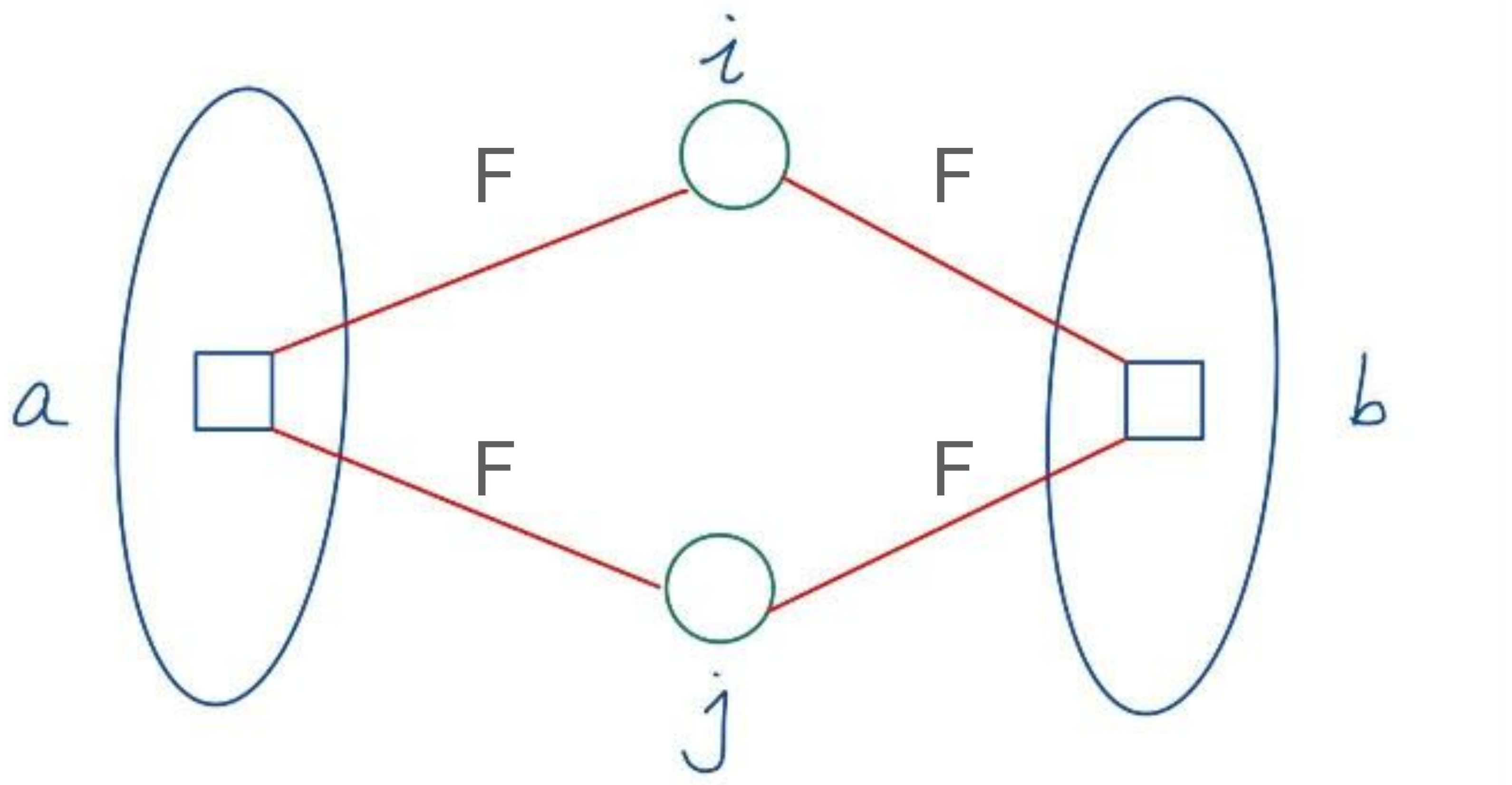
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

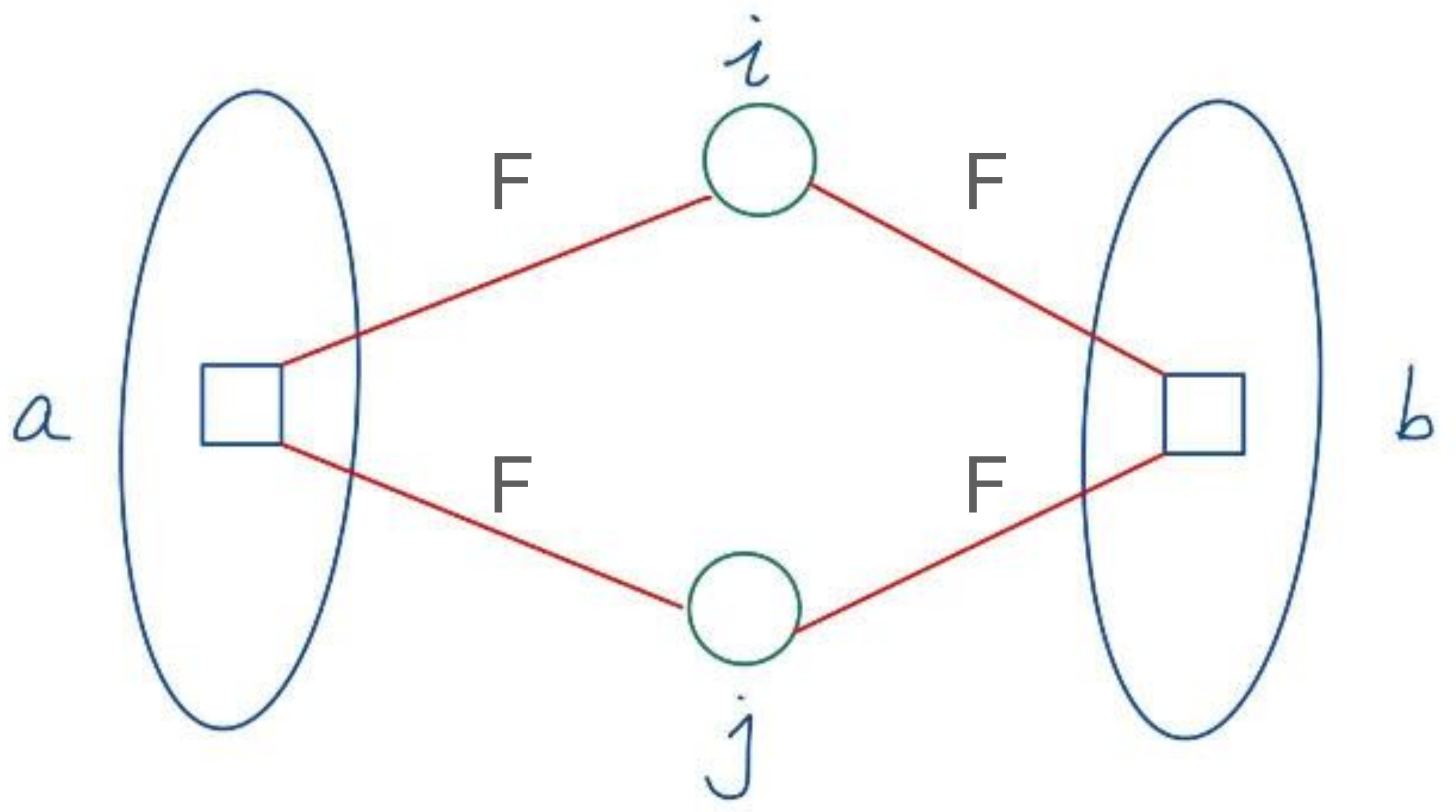
Arrival via S/H: q^3



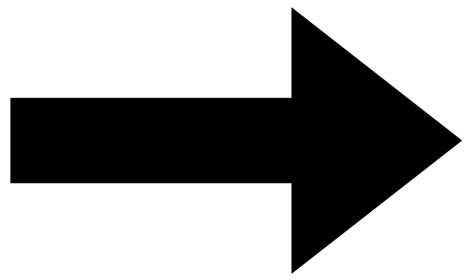
- a is making "middle" appearance, hence no factor
- i is making "first" appearance, but not "last", hence $\sqrt{d}$
- j is making "first" appearance, but not "last", hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$
- Edge-value $1/d^2$

A Taste of Our Local Analysis

Edge-factors	Vertex-factors
F/S/R: $1/\sqrt{d}$	First/Last: $\sqrt{d}$ (or $\sqrt{m}$)
H: $\sqrt{q}/\sqrt{d}$	Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is making "first" appearance, but not "last", hence $\sqrt{d}$
- j is making "first" appearance, but not "last", hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$
- Edge-value $1/d^2$



A Taste of Our Local Analysis

Edge-factors

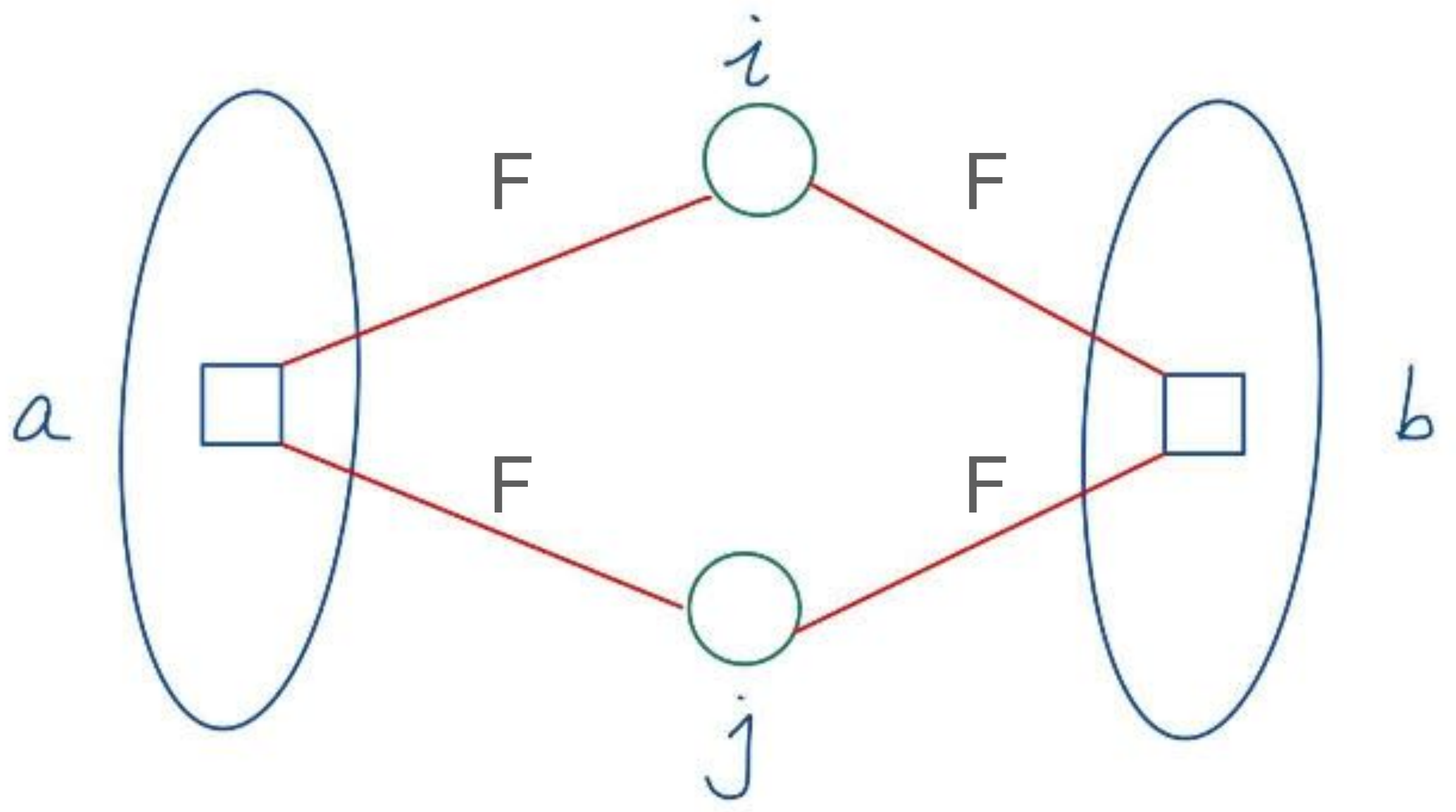
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

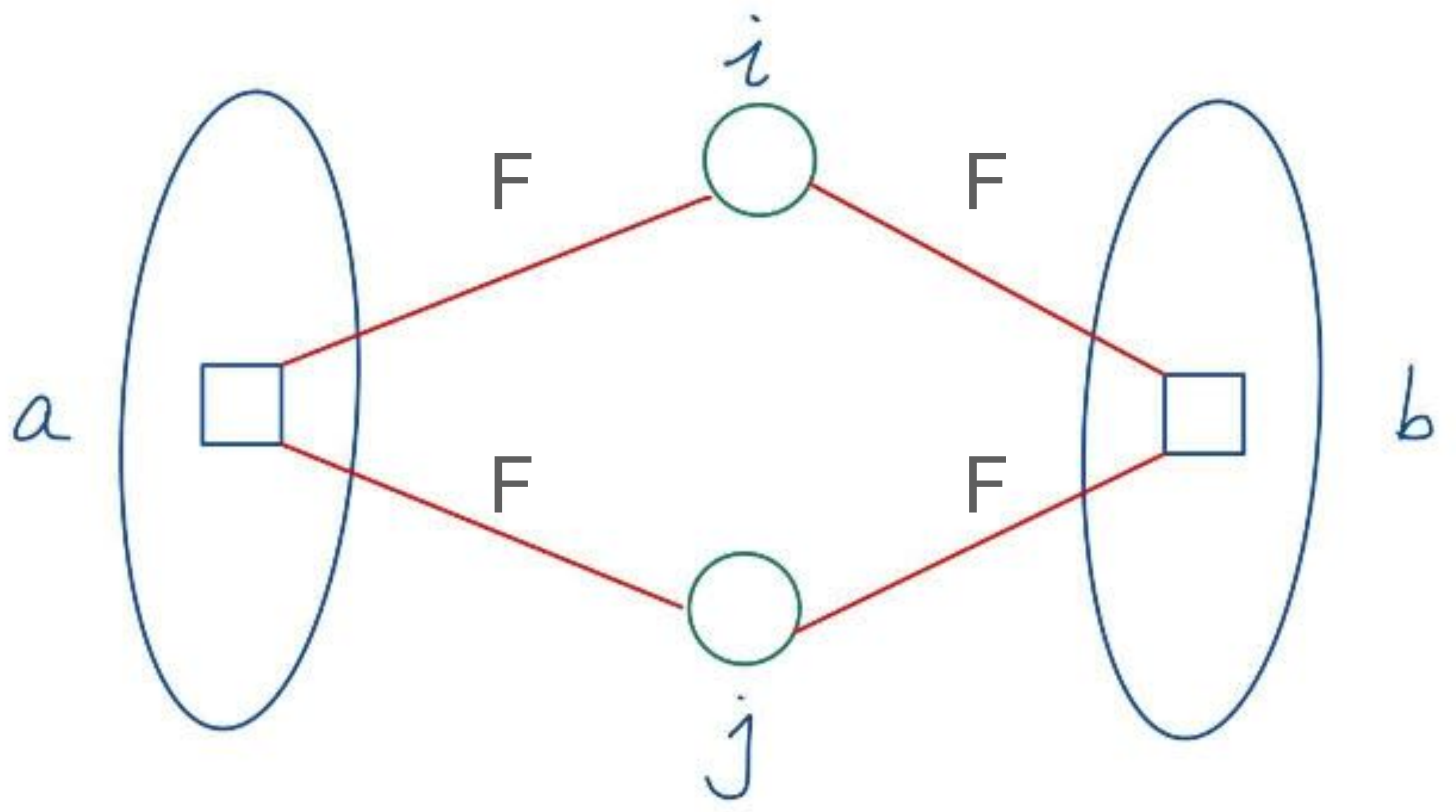


- a is making "middle" appearance, hence no factor
- i is making "first" appearance, but not "last", hence $\sqrt{d}$
- j is making "first" appearance, but not "last", hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$
- Edge-value $1/d^2$

$$\sqrt{m}d/d^2 \leq \sqrt{c}$$

A Taste of Our Local Analysis

Edge-factors		Vertex-factors	
F/S/R:	$1/\sqrt{d}$	First/Last:	$\sqrt{d}$ (or $\sqrt{m}$)
H:	$\sqrt{q}/\sqrt{d}$	Arrival via S/H:	q^3

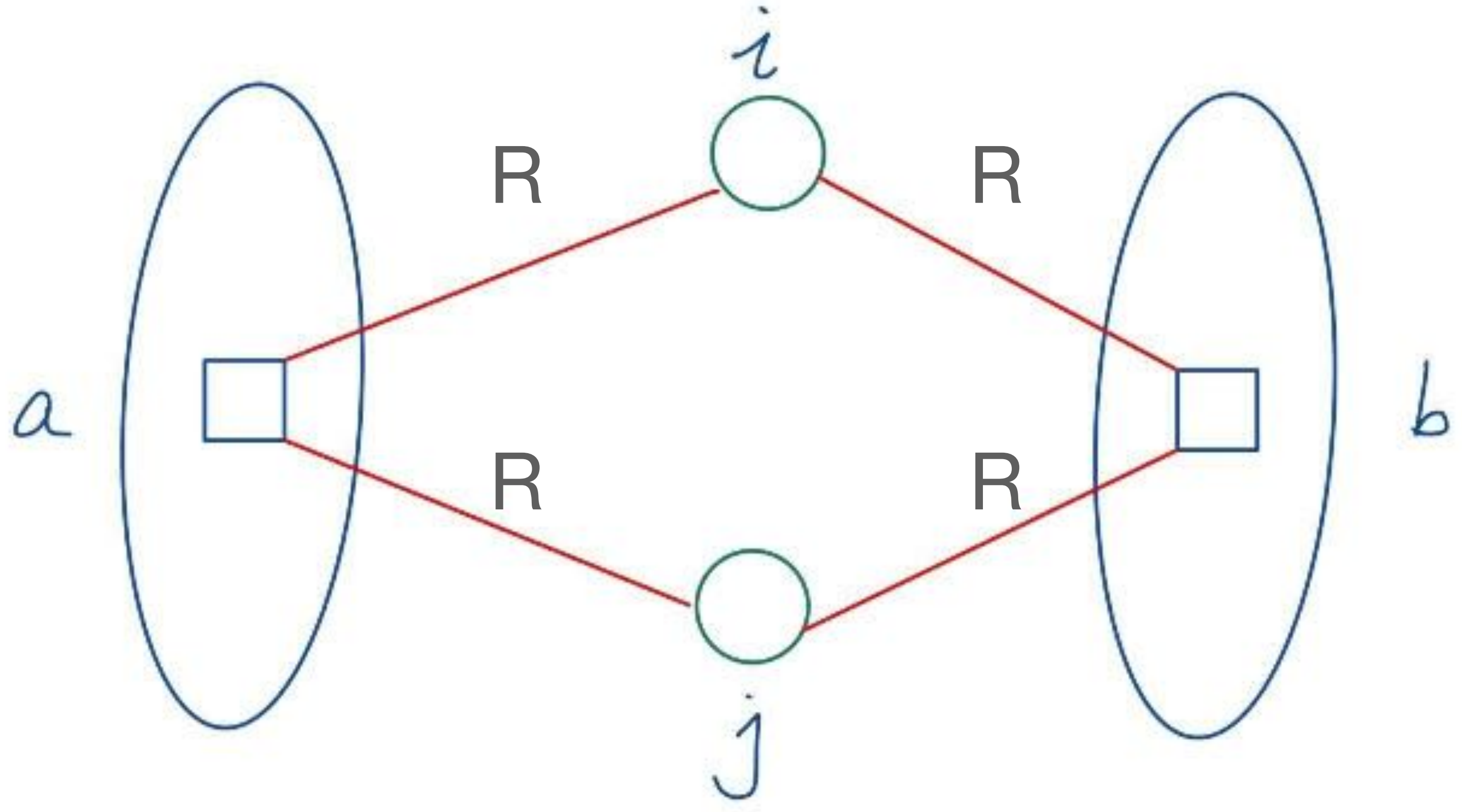


- a is making "middle" appearance, hence no factor
- i is making "first" appearance, but not "last", hence $\sqrt{d}$
- j is making "first" appearance, but not "last", hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$
- Edge-value $1/d^2$

$$\sqrt{md}/d^2 \leq \sqrt{c}$$

Dominant term!

A Taste of Our Local Analysis



Edge-factors

F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

A Taste of Our Local Analysis

Edge-factors

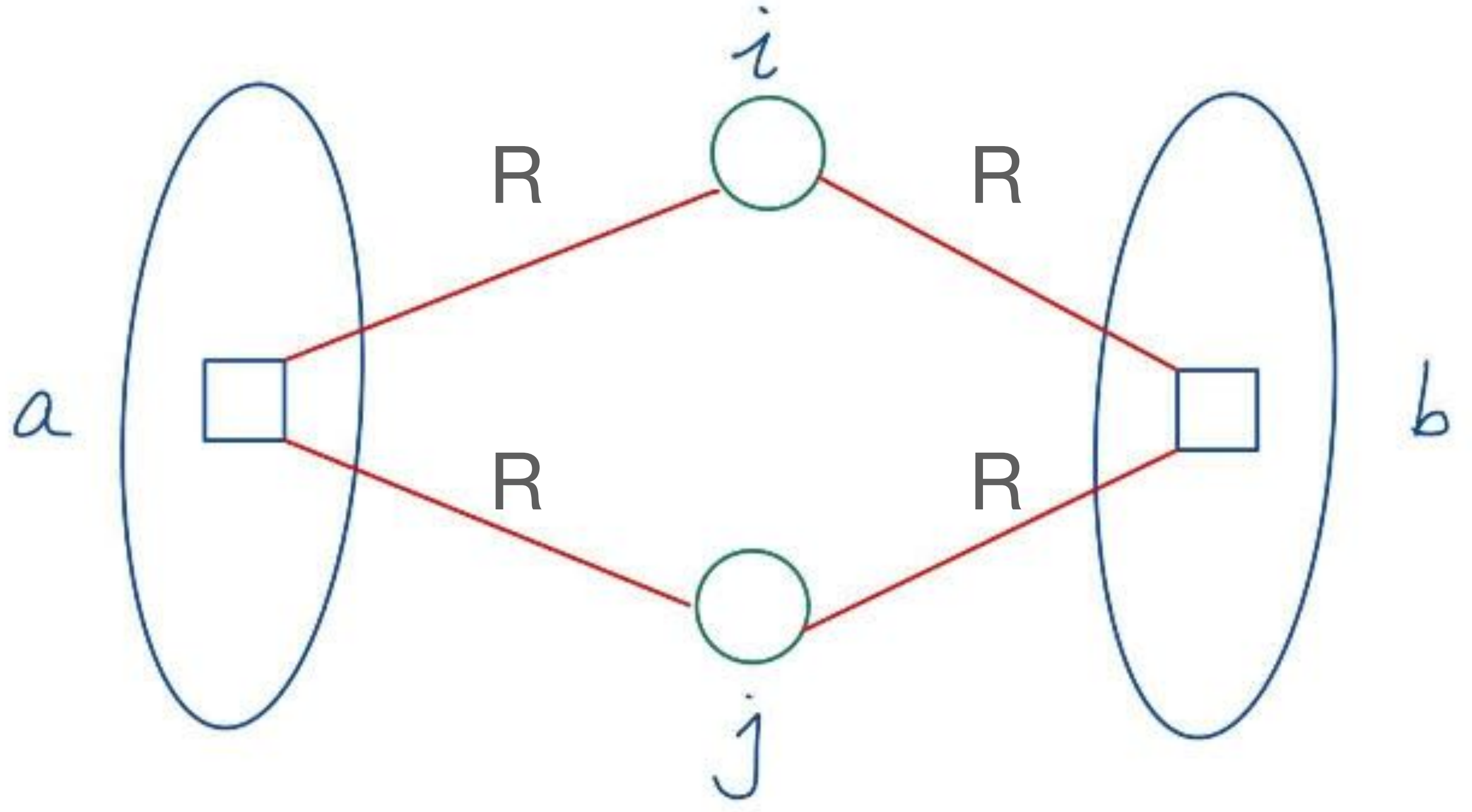
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor

A Taste of Our Local Analysis

Edge-factors

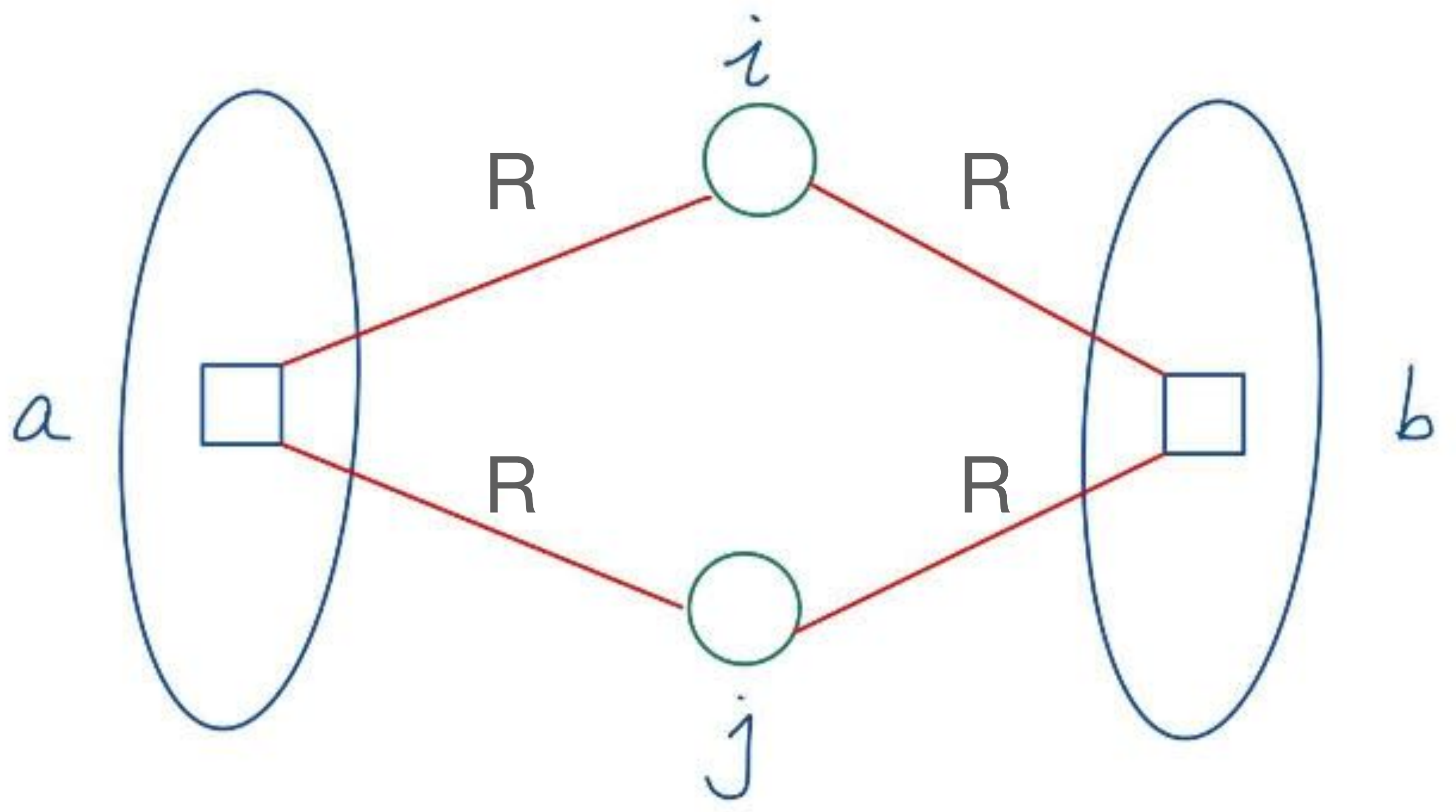
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making “middle” appearance, hence no factor
- i is making “last” appearance, but not “first”, hence $\sqrt{d}$

A Taste of Our Local Analysis

Edge-factors

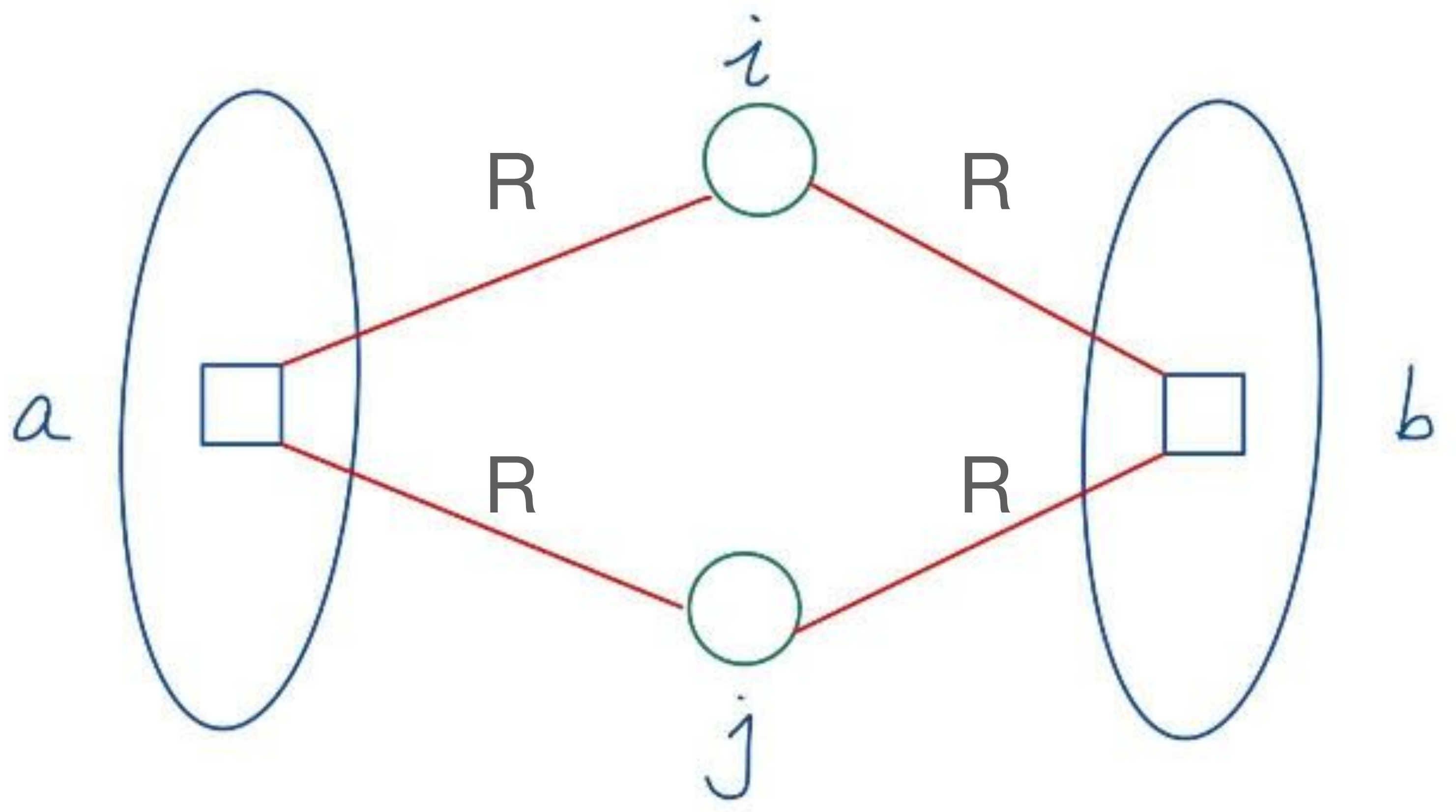
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

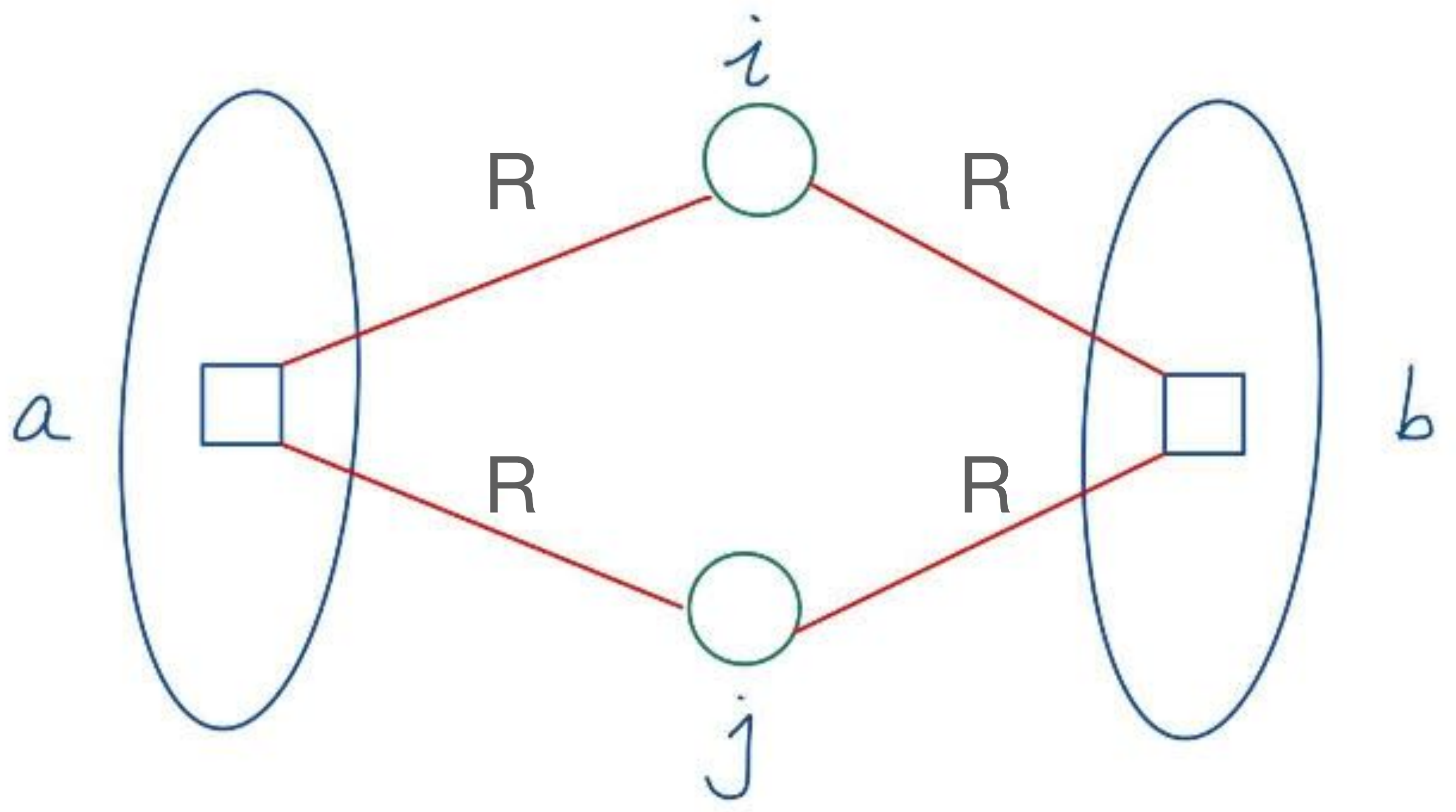
Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is making "last" appearance, but not "first", hence $\sqrt{d}$
- j is making "last" appearance, but not "first", hence $\sqrt{d}$

A Taste of Our Local Analysis

Edge-factors	Vertex-factors
F/S/R: $1/\sqrt{d}$	First/Last: $\sqrt{d}$ (or $\sqrt{m}$)
H: $\sqrt{q}/\sqrt{d}$	Arrival via S/H: q^3



- a is making “middle” appearance, hence no factor
- i is making “last” appearance, but not “first”, hence $\sqrt{d}$
- j is making “last” appearance, but not “first”, hence $\sqrt{d}$
- b is making “last” appearance, but not “first”, hence $\sqrt{m}$

A Taste of Our Local Analysis

Edge-factors

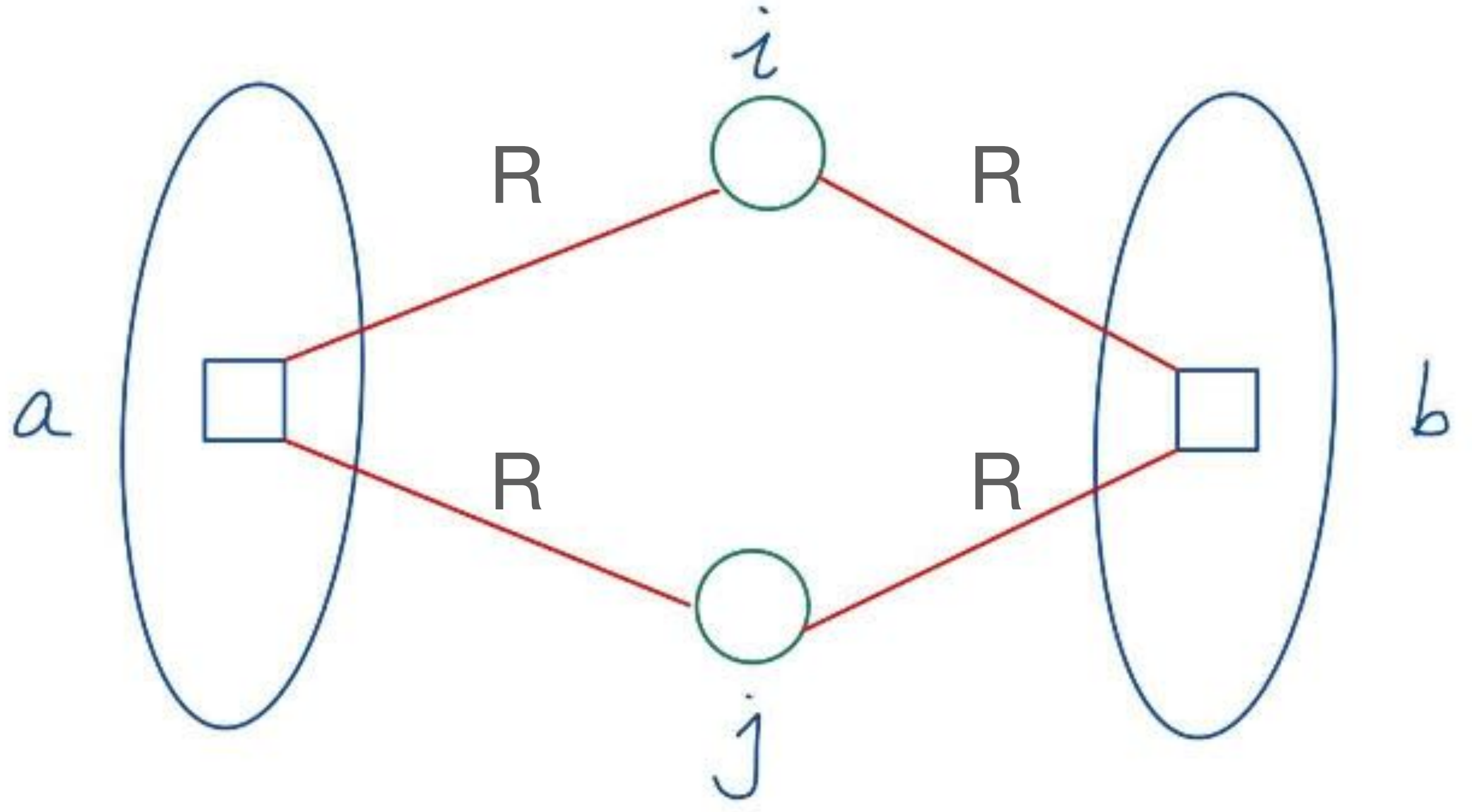
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is making "last" appearance, but not "first", hence $\sqrt{d}$
- j is making "last" appearance, but not "first", hence $\sqrt{d}$
- b is making "last" appearance, but not "first", hence $\sqrt{m}$
- Edge-value $1/d^2$

A Taste of Our Local Analysis

Edge-factors

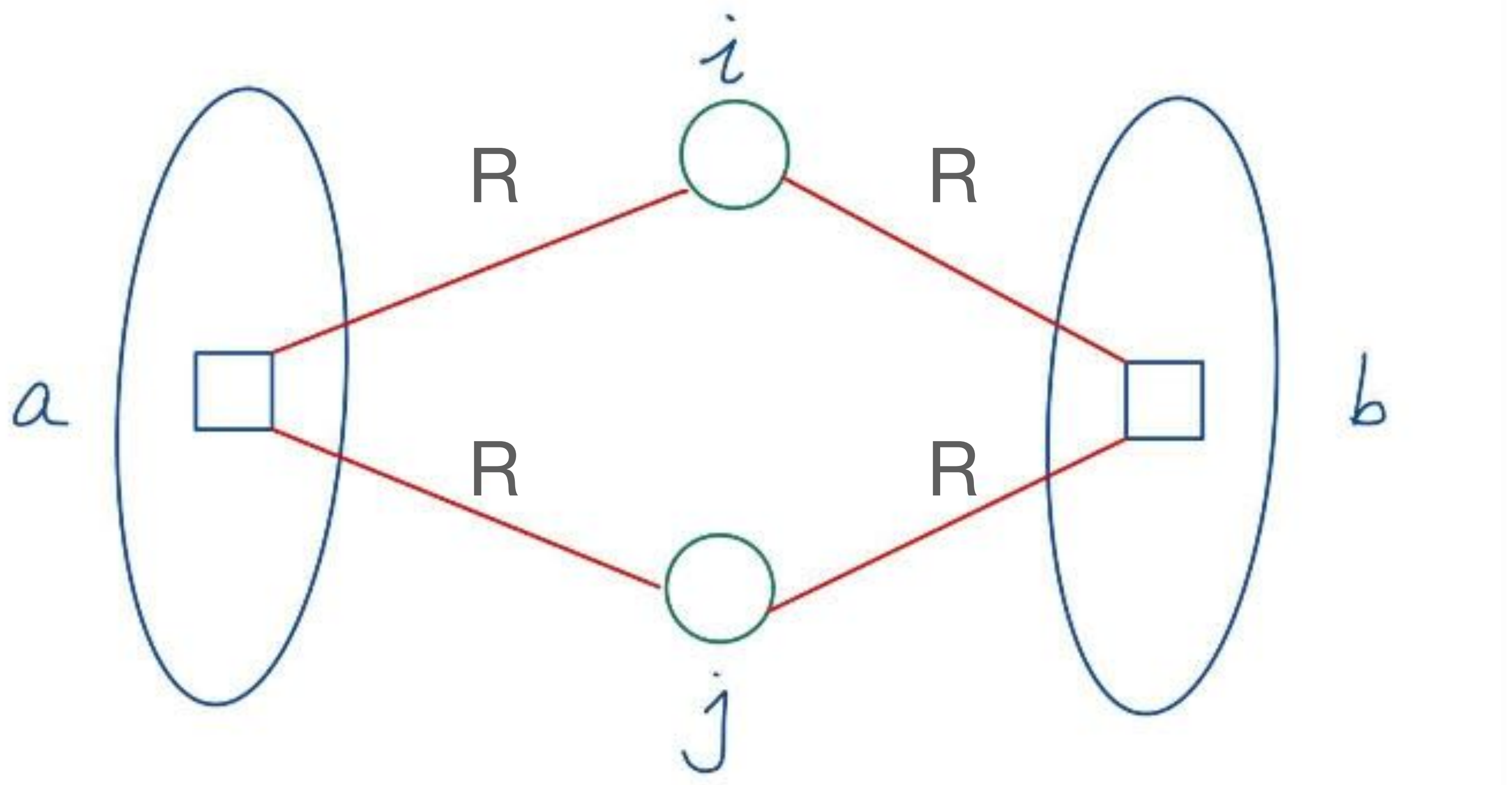
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making “middle” appearance, hence no factor
- i is making “last” appearance, but not “first”, hence $\sqrt{d}$
- j is making “last” appearance, but not “first”, hence $\sqrt{d}$
- b is making “last” appearance, but not “first”, hence $\sqrt{m}$
- Edge-value $1/d^2$

A Taste of Our Local Analysis

Edge-factors

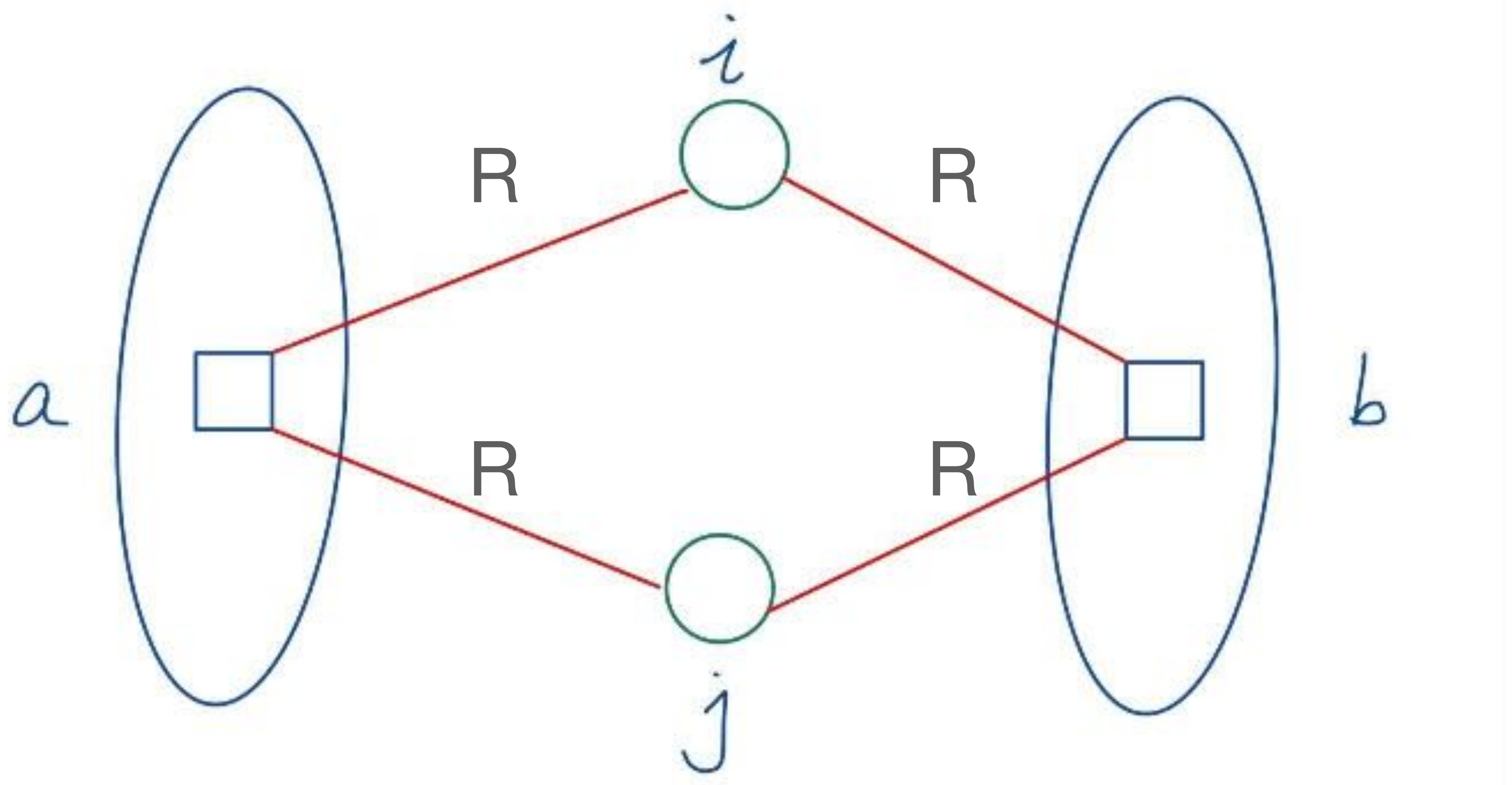
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

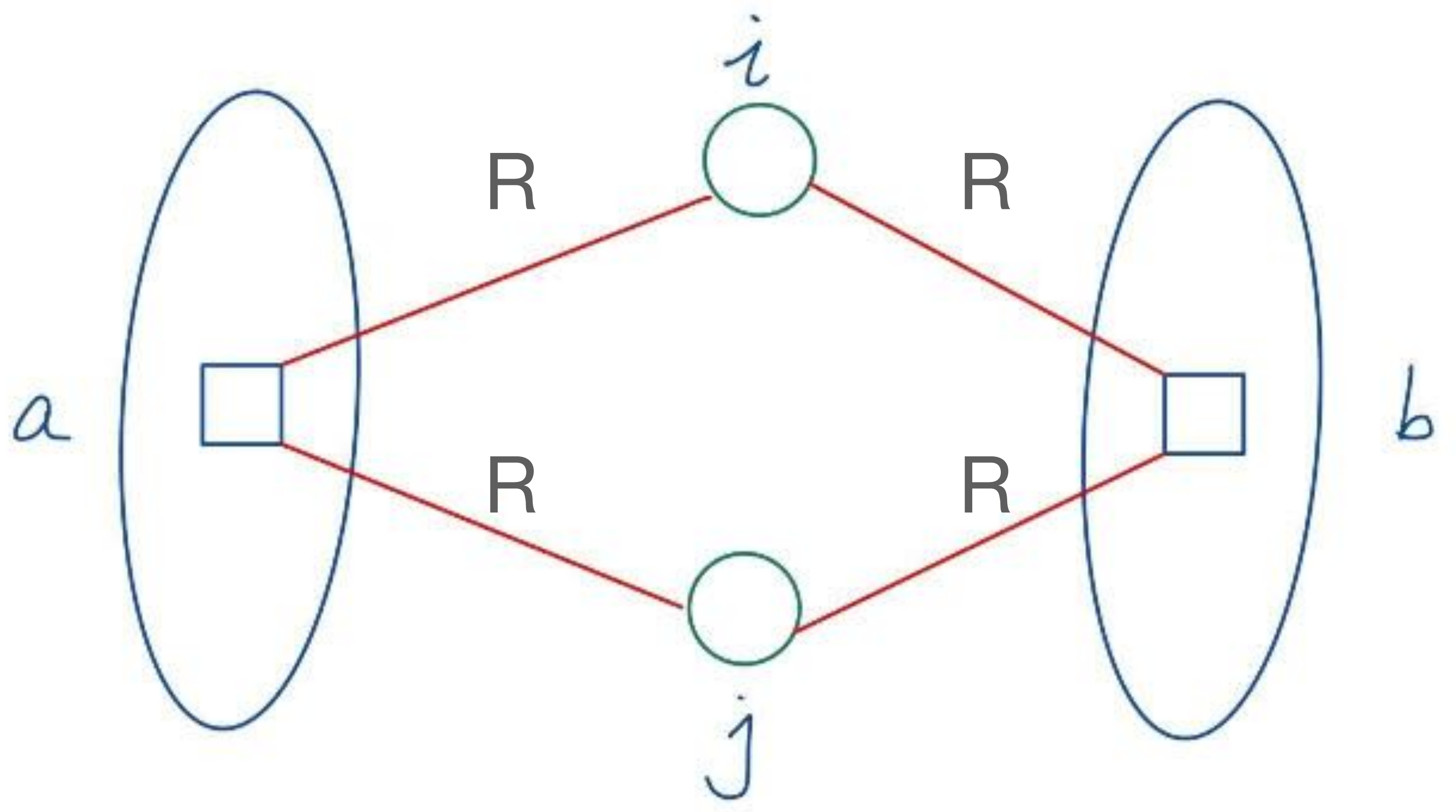


- a is making “middle” appearance, hence no factor
- i is making “last” appearance, but not “first”, hence $\sqrt{d}$
- j is making “last” appearance, but not “first”, hence $\sqrt{d}$
- b is making “last” appearance, but not “first”, hence $\sqrt{m}$
- Edge-value $1/d^2$

$$\Rightarrow \quad \sqrt{m}d/d^2 \leq \sqrt{c}$$

A Taste of Our Local Analysis

Edge-factors	Vertex-factors
F/S/R: $1/\sqrt{d}$	First/Last: $\sqrt{d}$ (or $\sqrt{m}$)
H: $\sqrt{q}/\sqrt{d}$	Arrival via S/H: q^3

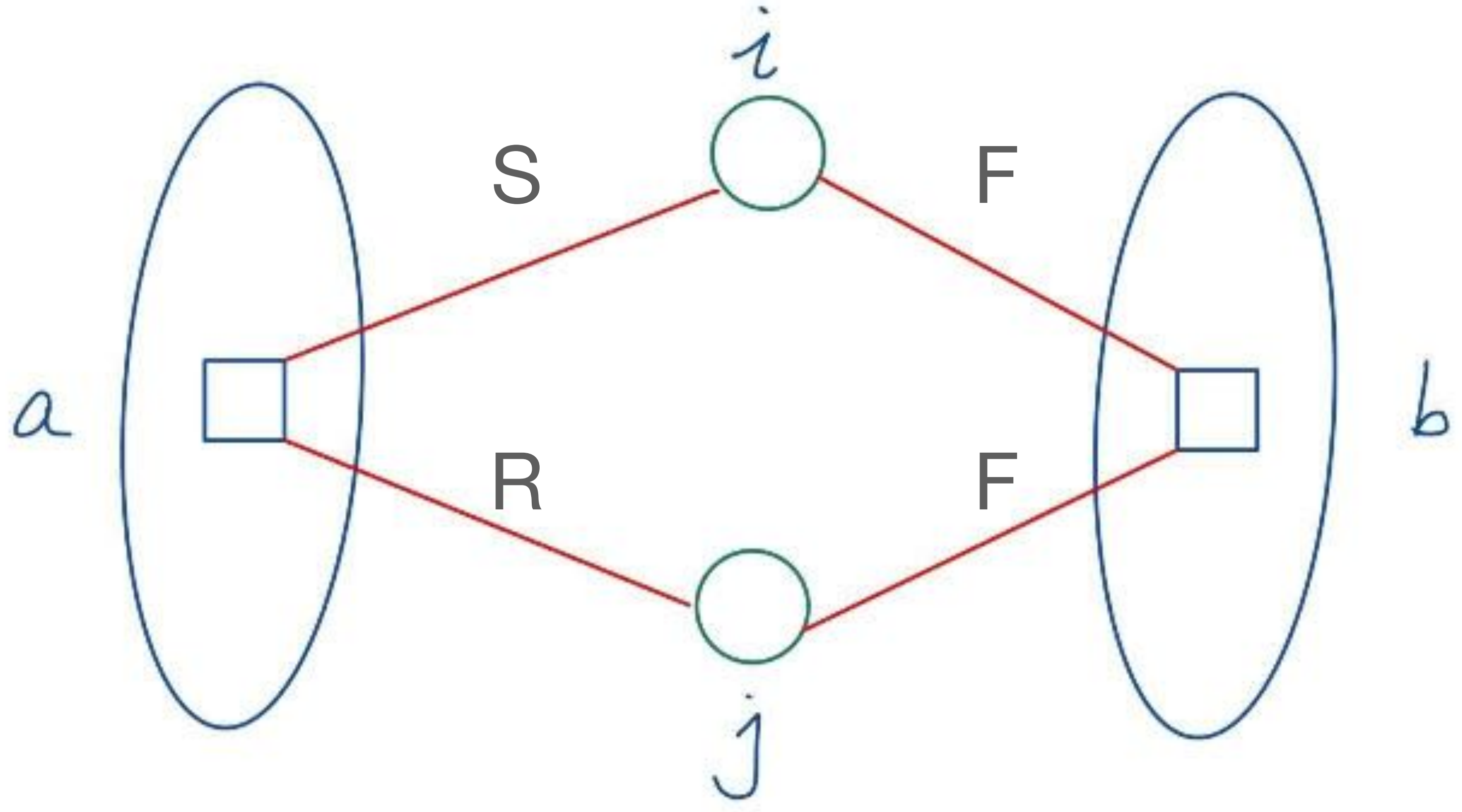


- a is making "middle" appearance, hence no factor
- i is making "last" appearance, but not "first", hence $\sqrt{d}$
- j is making "last" appearance, but not "first", hence $\sqrt{d}$
- b is making "last" appearance, but not "first", hence $\sqrt{m}$
- Edge-value $1/d^2$

$$\sqrt{m}d/d^2 \leq \sqrt{c}$$

Dominant term!

A Taste of Our Local Analysis



Edge-factors

F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

A Taste of Our Local Analysis

Edge-factors

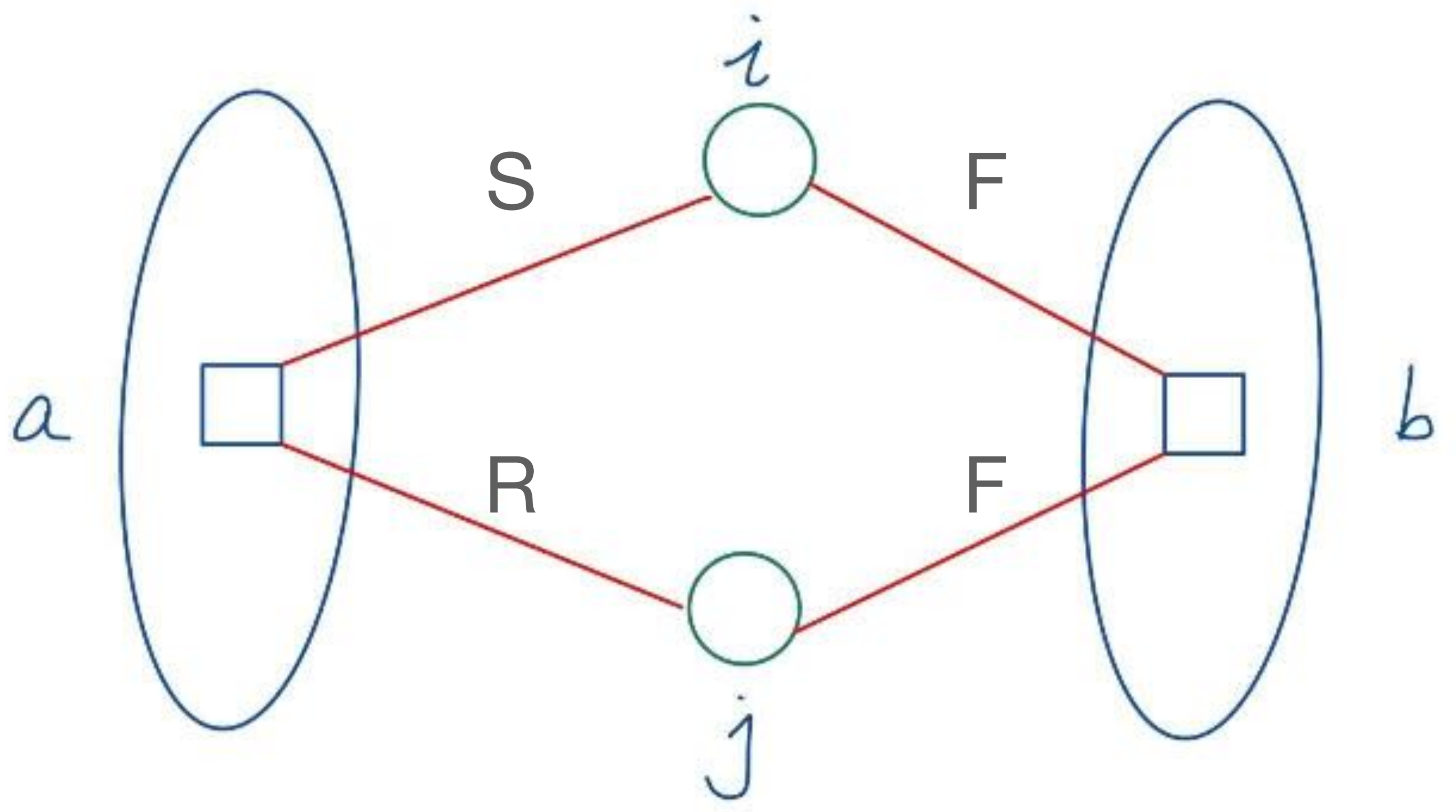
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor

A Taste of Our Local Analysis

Edge-factors

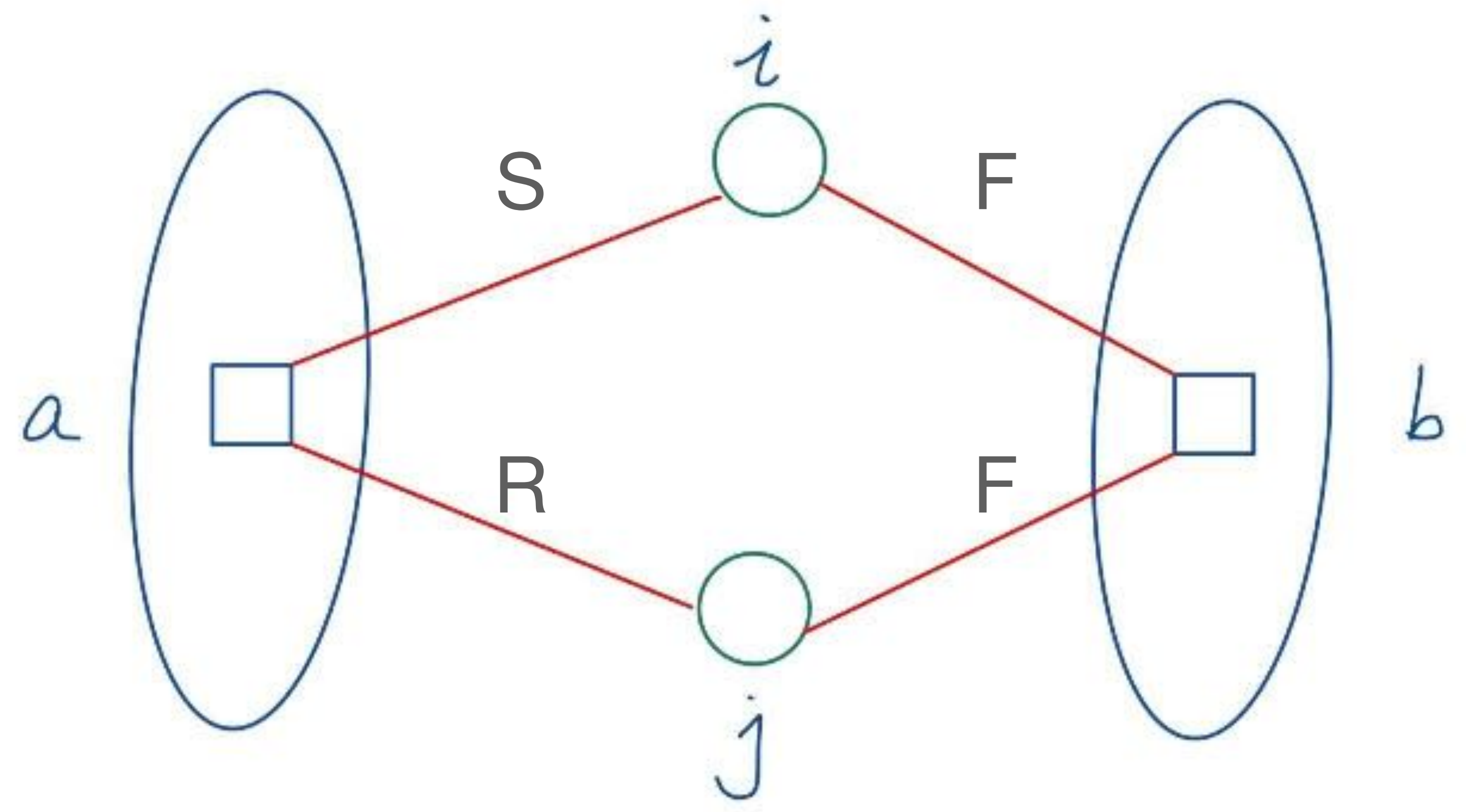
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is not "first" nor "last", yet it is arrived using S, hence q^3

A Taste of Our Local Analysis

Edge-factors

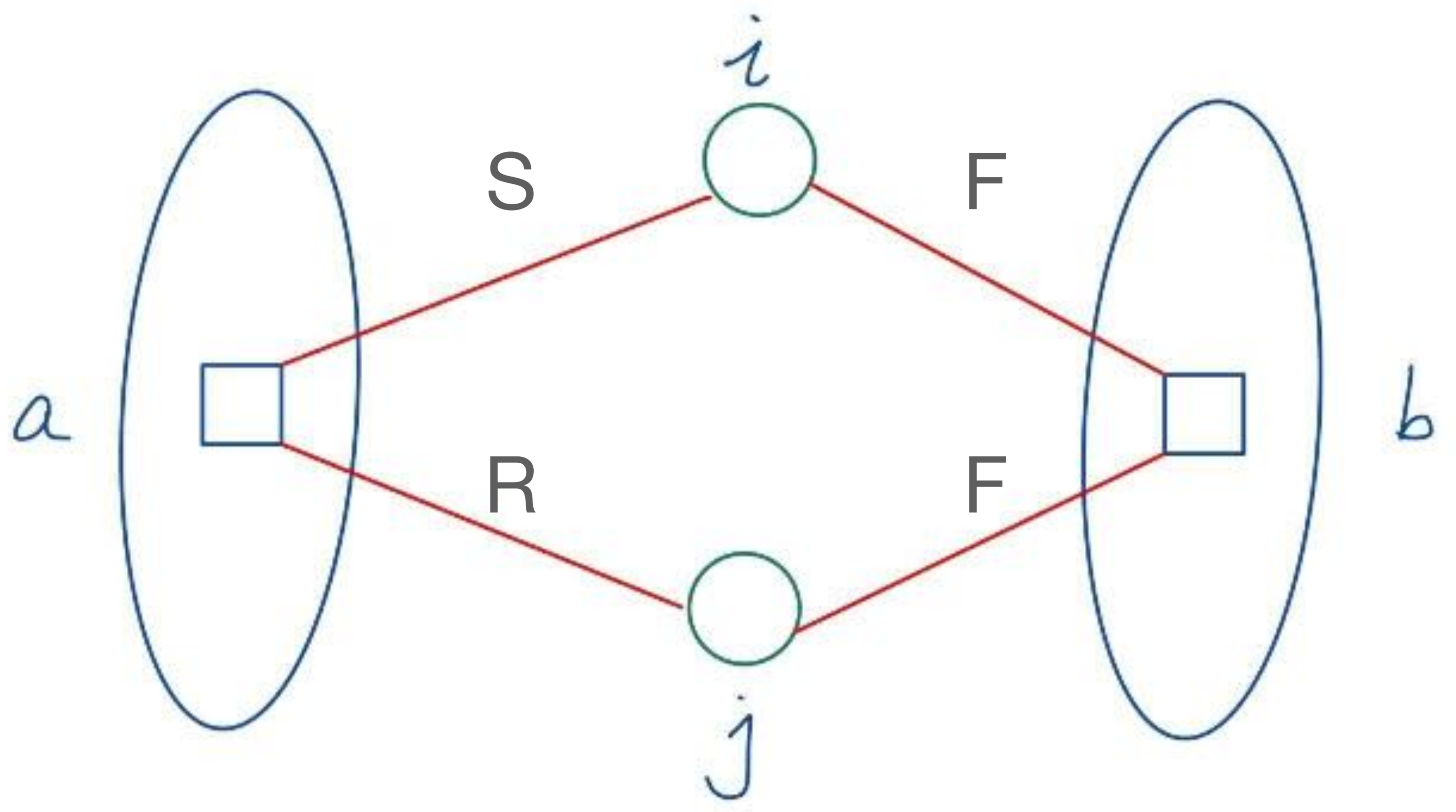
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is not "first" nor "last", yet it is arrived using S, hence q^3
- j is making "middle" appearance, hence $\sqrt{d}$

A Taste of Our Local Analysis

Edge-factors

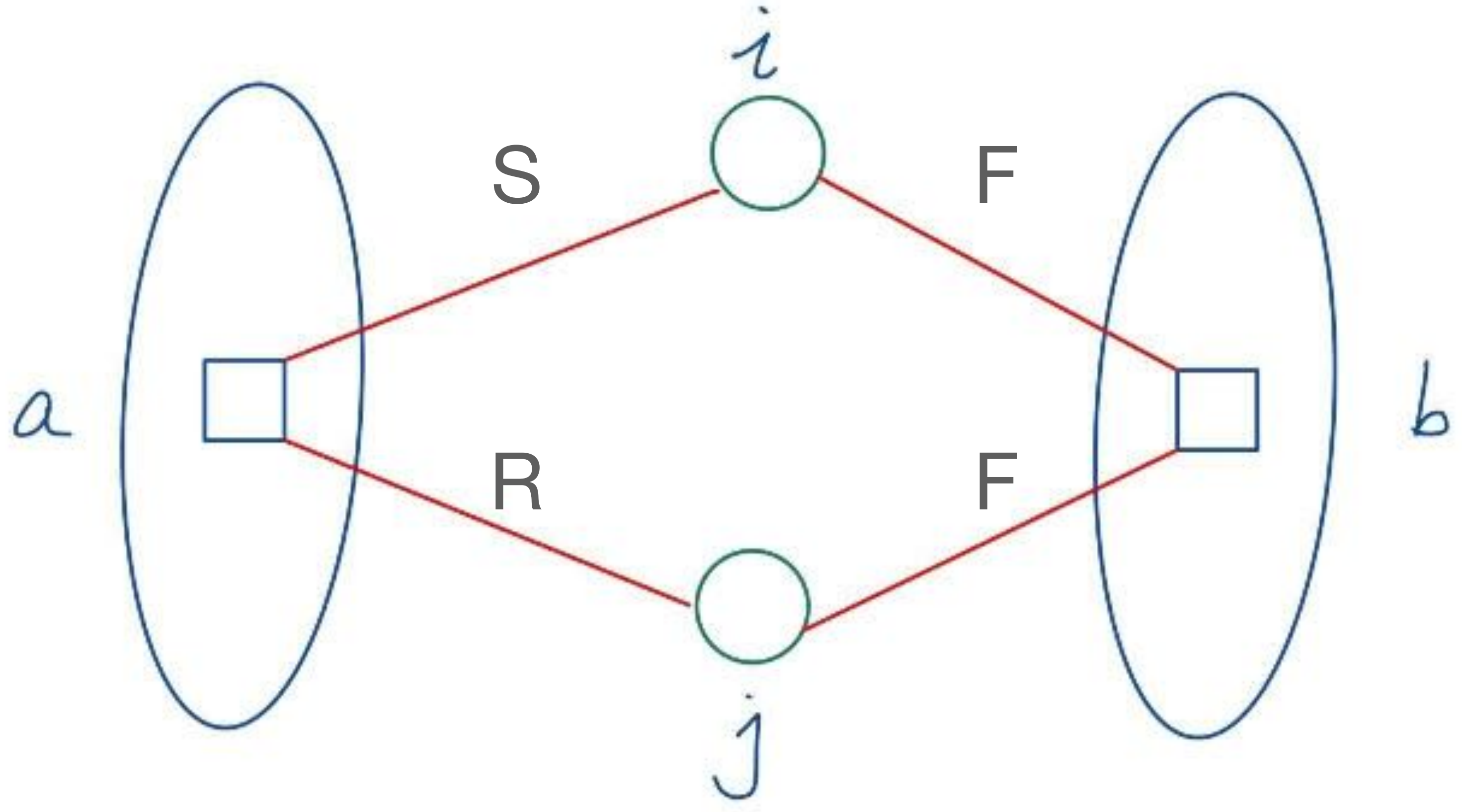
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making “middle” appearance, hence no factor
- i is not “first” nor “last”, yet it is arrived using S, hence q^3
- j is making “middle” appearance, hence $\sqrt{d}$
- b is making “first” appearance, but not “last”, hence $\sqrt{m}$

A Taste of Our Local Analysis

Edge-factors

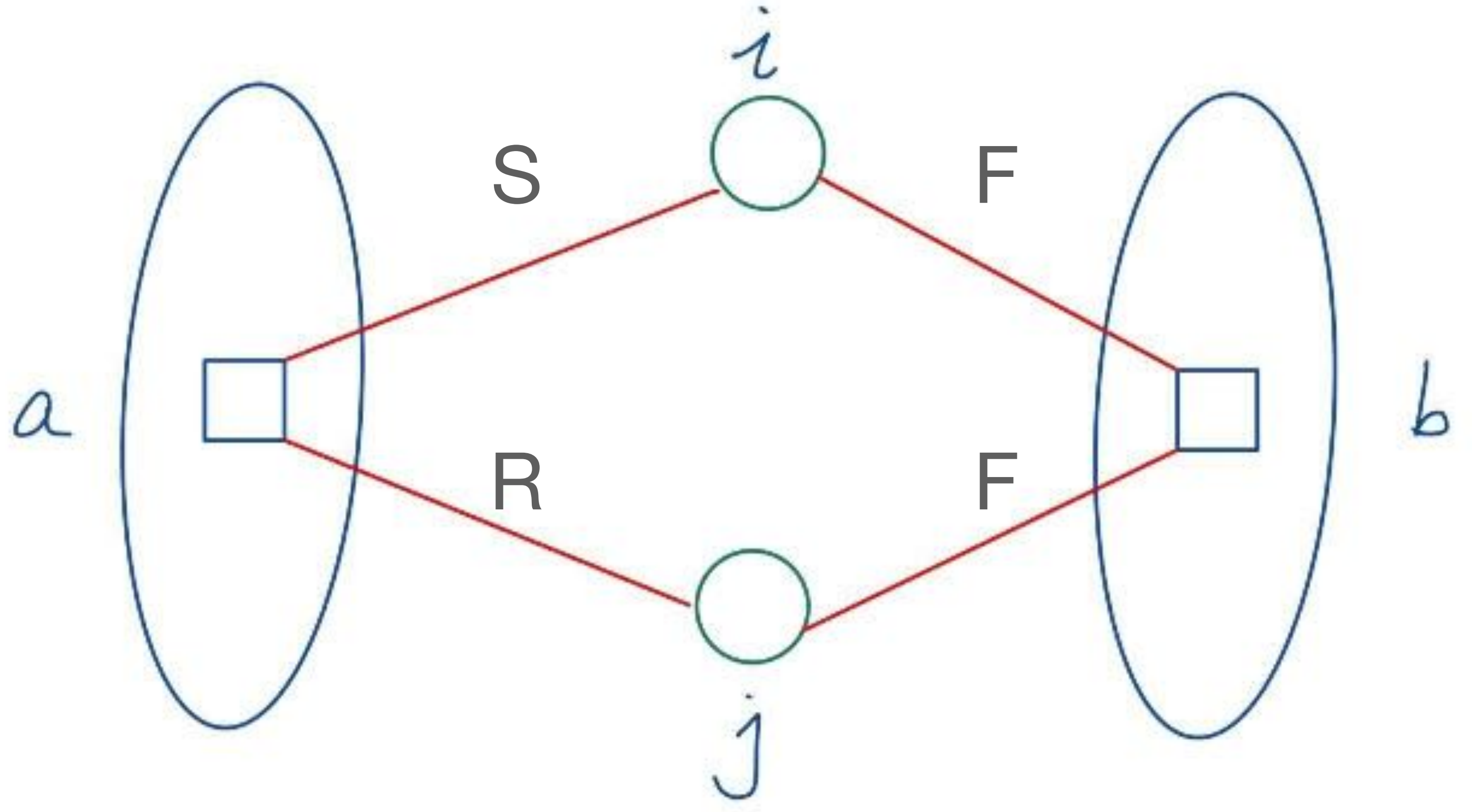
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is not "first" nor "last", yet it is arrived using S, hence q^3
- j is making "middle" appearance, hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$
- Edge-value $1/d^2$

A Taste of Our Local Analysis

Edge-factors

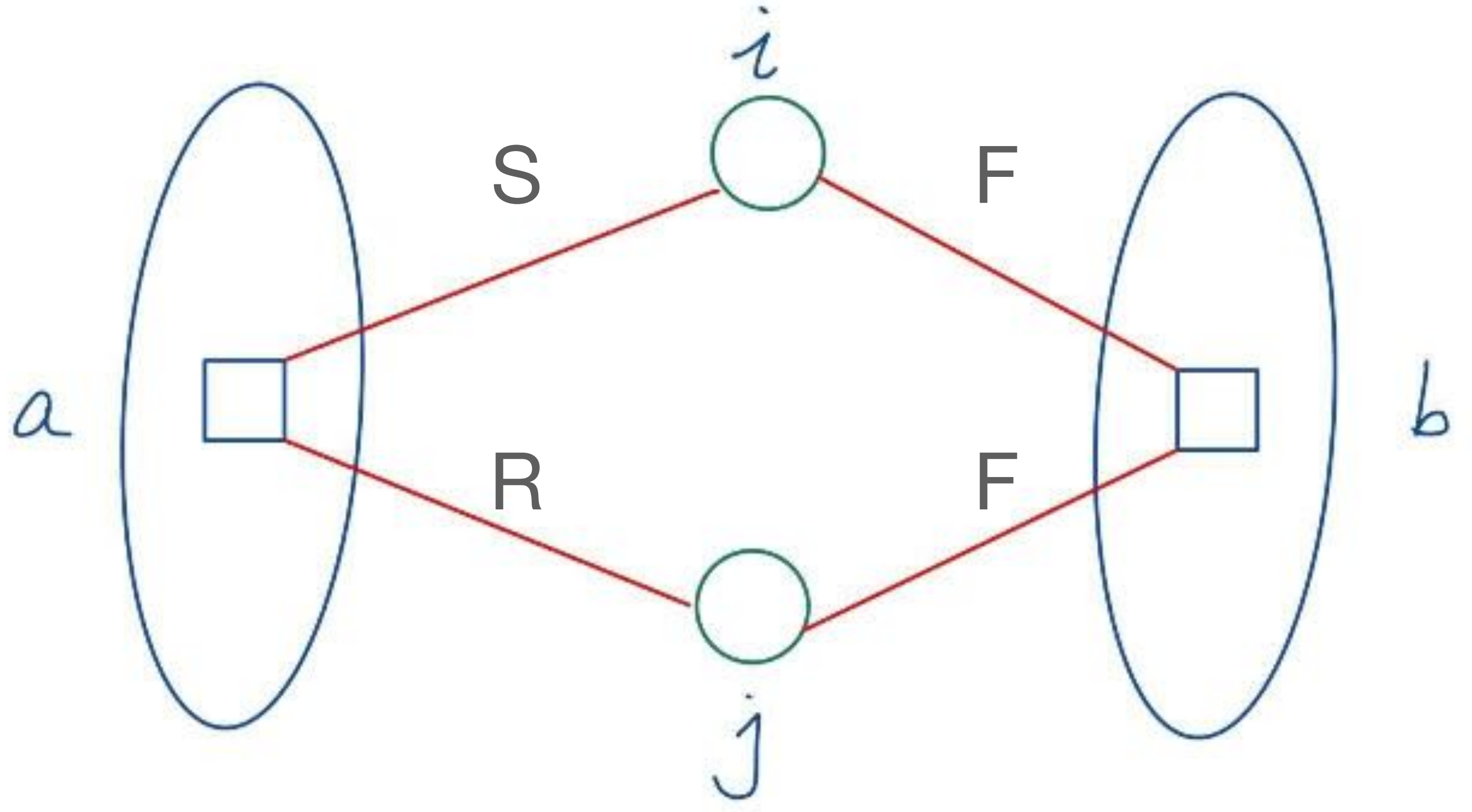
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

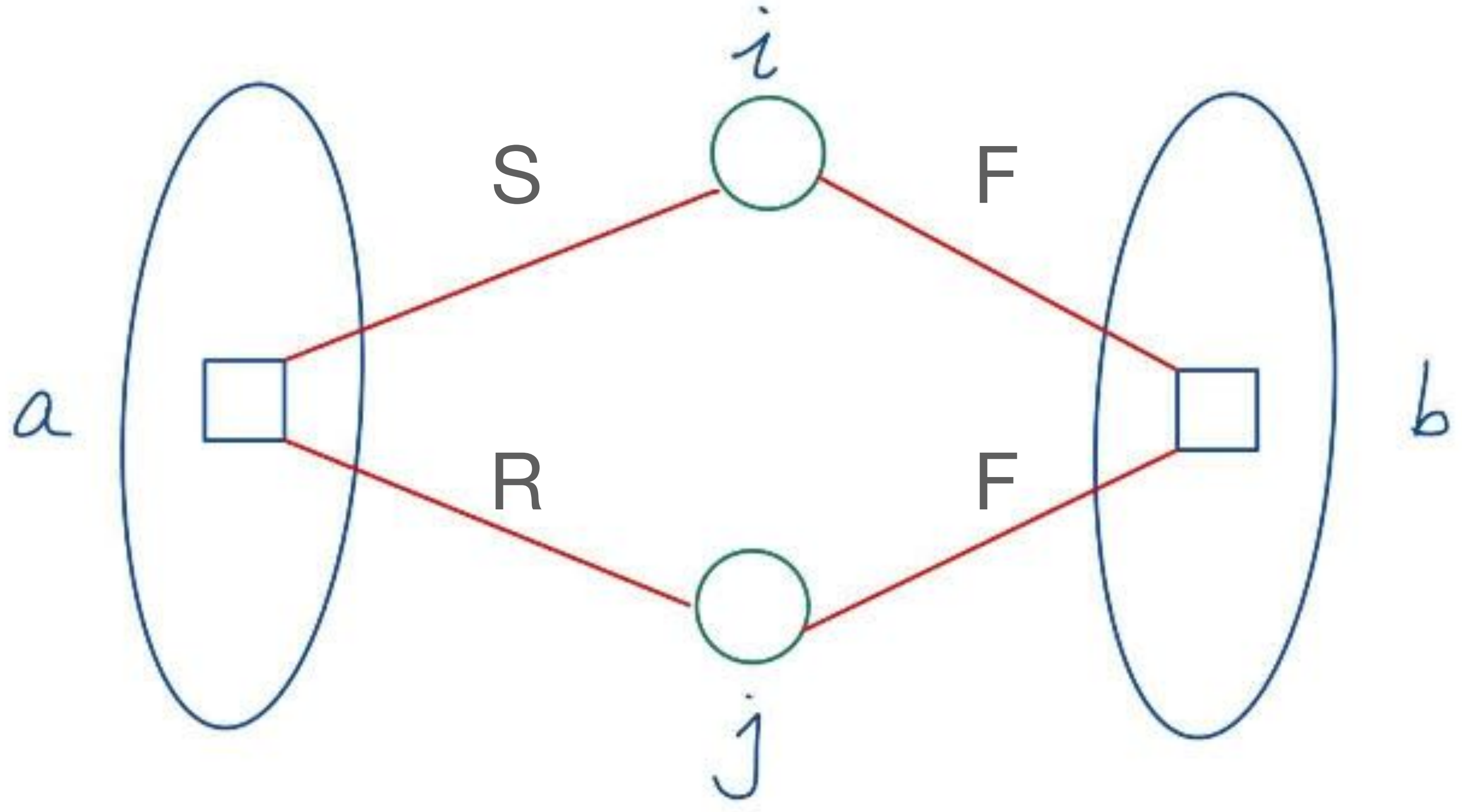
Arrival via S/H: q^3



- a is making "middle" appearance, hence no factor
- i is not "first" nor "last", yet it is arrived using S, hence q^3
- j is making "middle" appearance, hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$
- Edge-value $1/d^2$

A Taste of Our Local Analysis

Edge-factors		Vertex-factors	
F/S/R:	$1/\sqrt{d}$	First/Last:	$\sqrt{d}$ (or $\sqrt{m}$)
H:	$\sqrt{q}/\sqrt{d}$	Arrival via S/H:	q^3



- a is making "middle" appearance, hence no factor
- i is not "first" nor "last", yet it is arrived using S, hence q^3
- j is making "middle" appearance, hence $\sqrt{d}$
- b is making "first" appearance, but not "last", hence $\sqrt{m}$
- Edge-value $1/d^2$

$$\Rightarrow \sqrt{md}q^3/d^2 \leq o_d(1)$$

A Taste of Our Local Analysis

Edge-factors

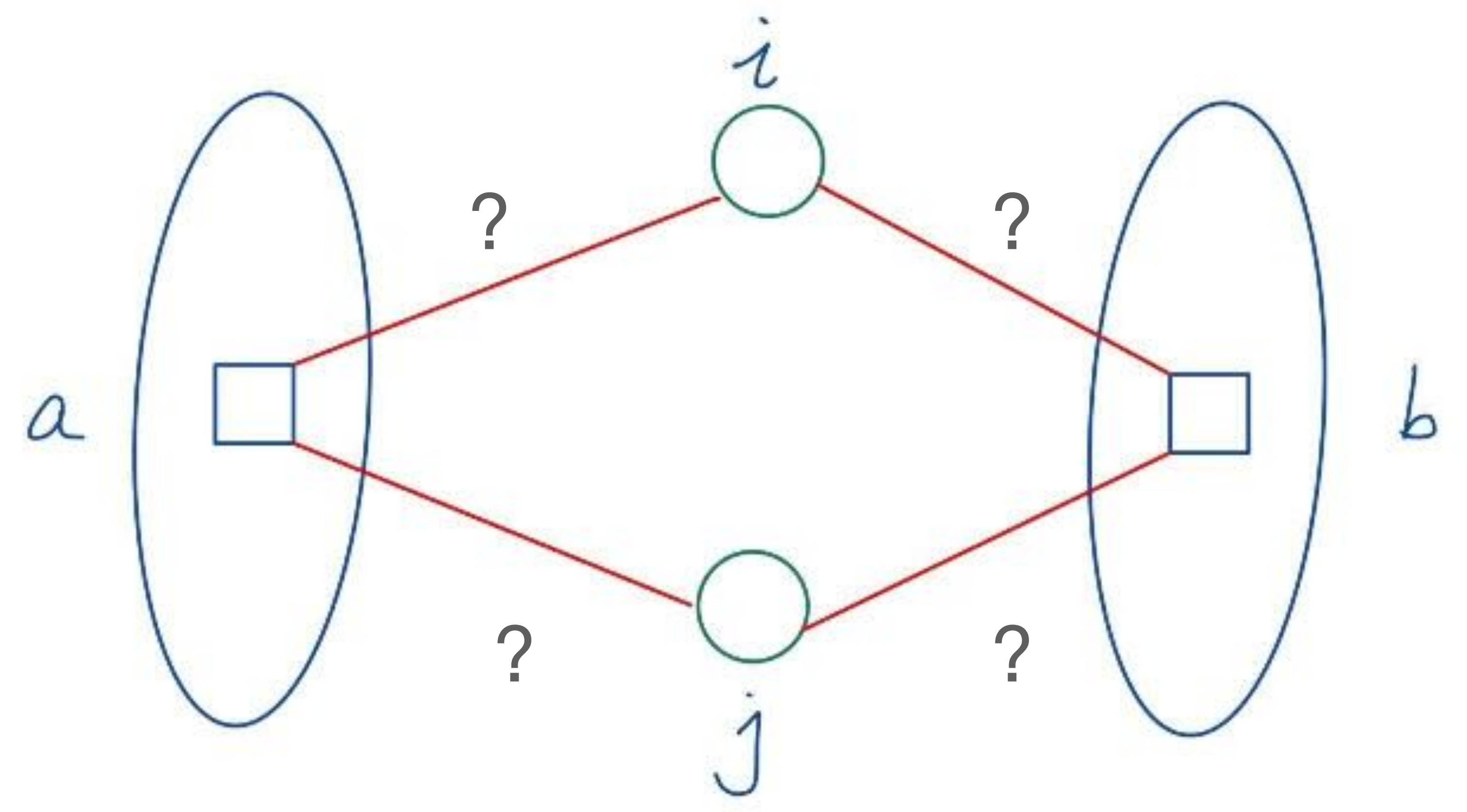
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



A Taste of Our Local Analysis

Edge-factors

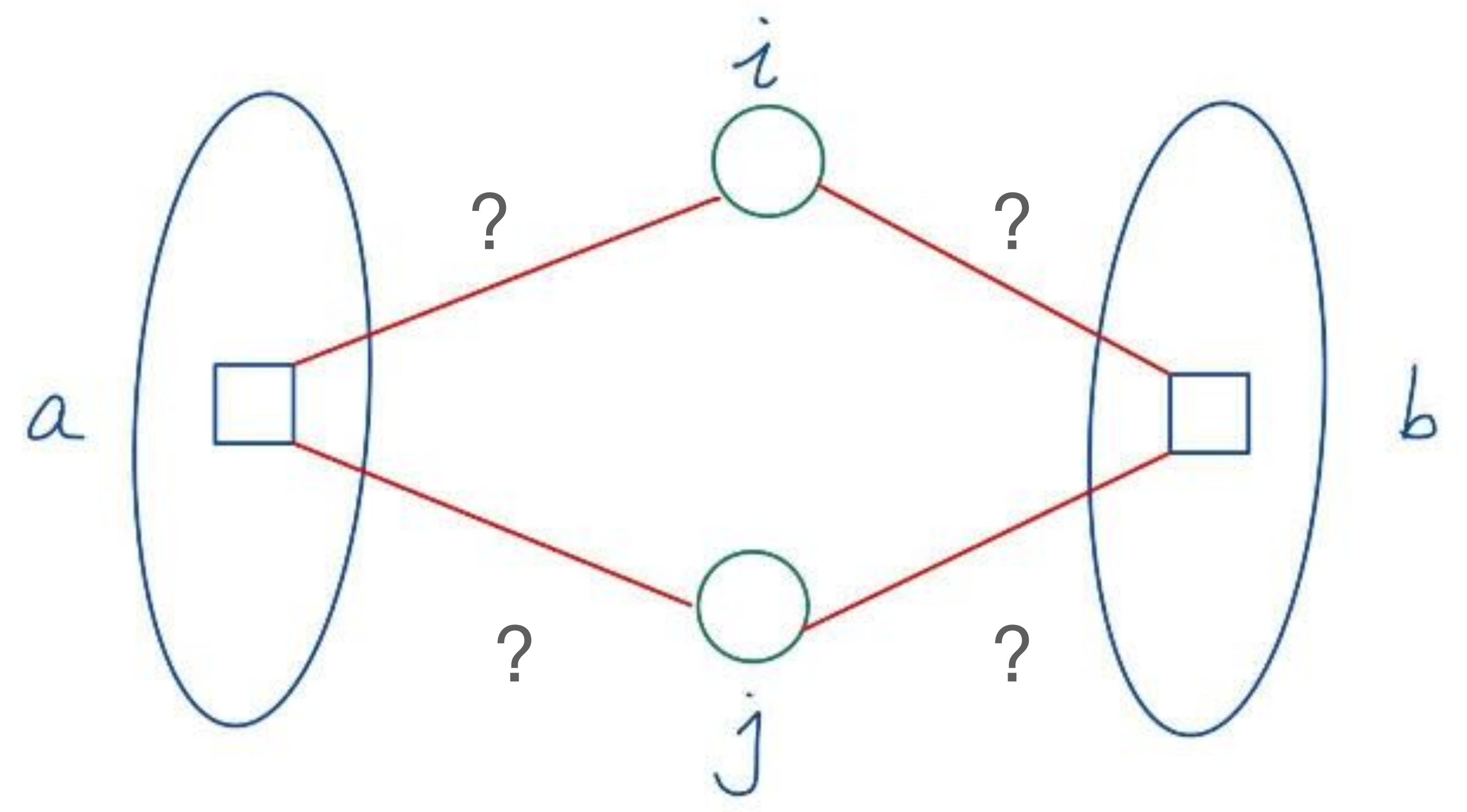
F/S/R: $1/\sqrt{d}$

H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3



A Taste of Our Local Analysis

Edge-factors

F/S/R: $1/\sqrt{d}$

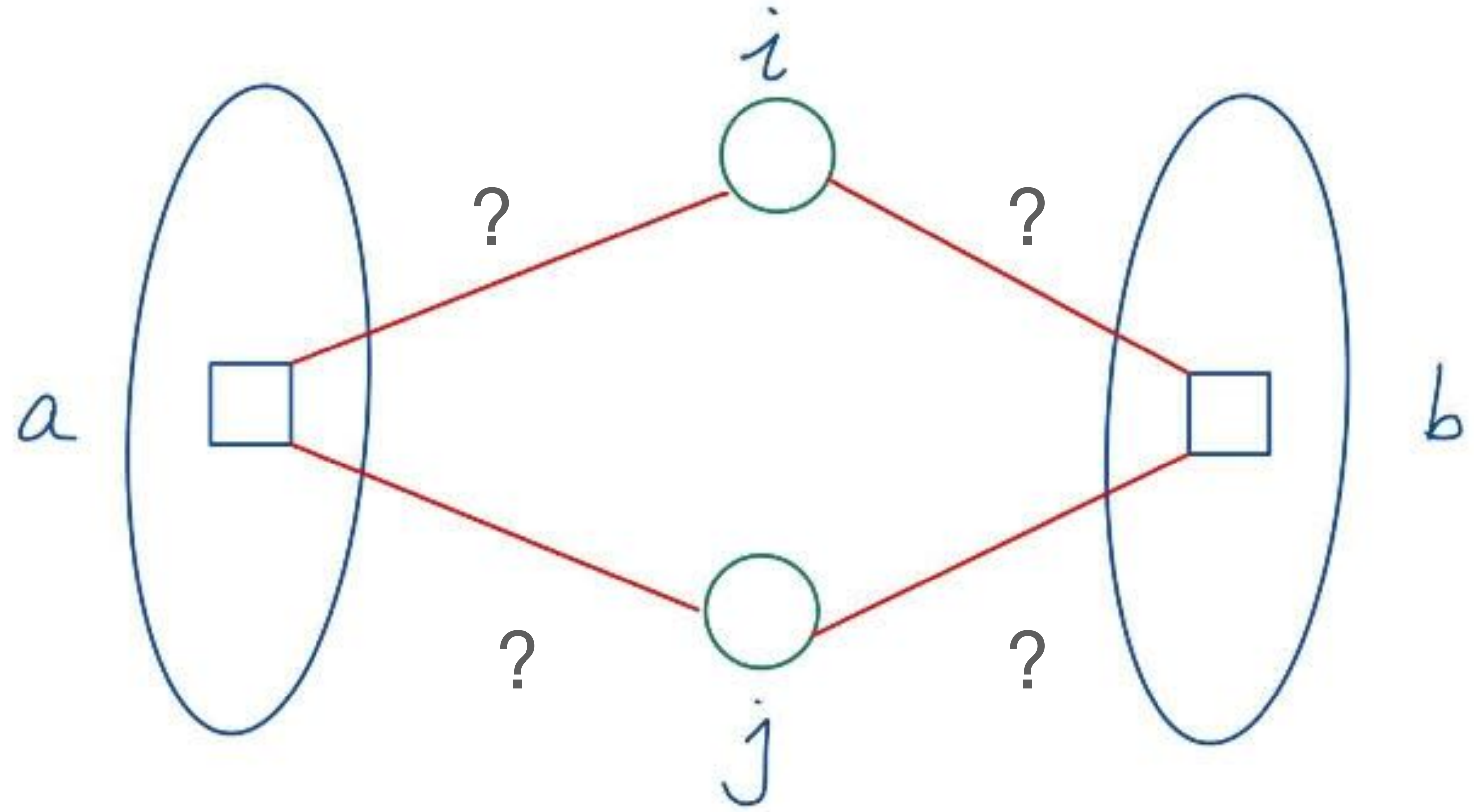
H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

- Norm bounds in 2-steps



A Taste of Our Local Analysis

Edge-factors

F/S/R: $1/\sqrt{d}$

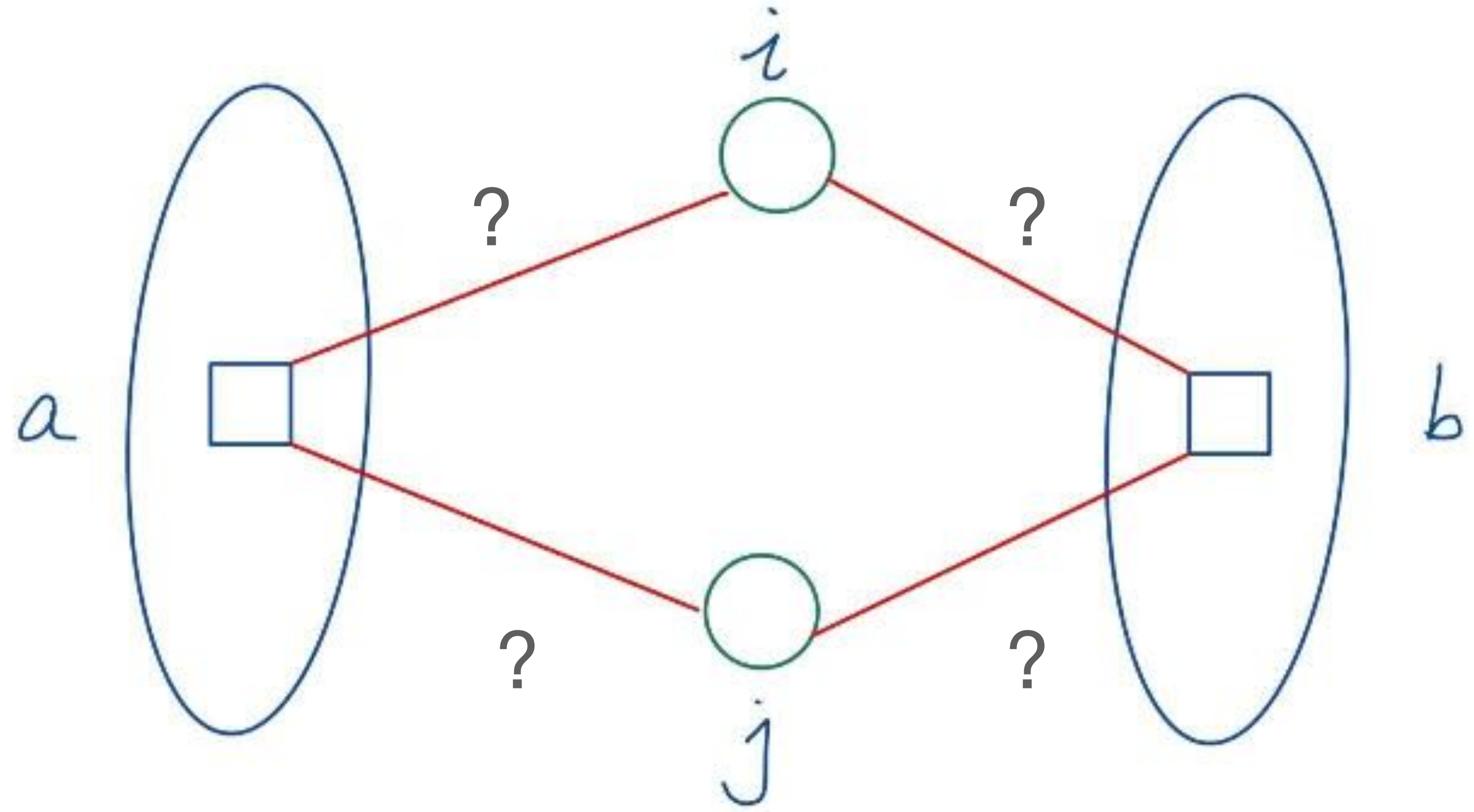
H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

- Norm bounds in 2-steps
  - Bound the local-value of each edge-labeling



A Taste of Our Local Analysis

Edge-factors

F/S/R: $1/\sqrt{d}$

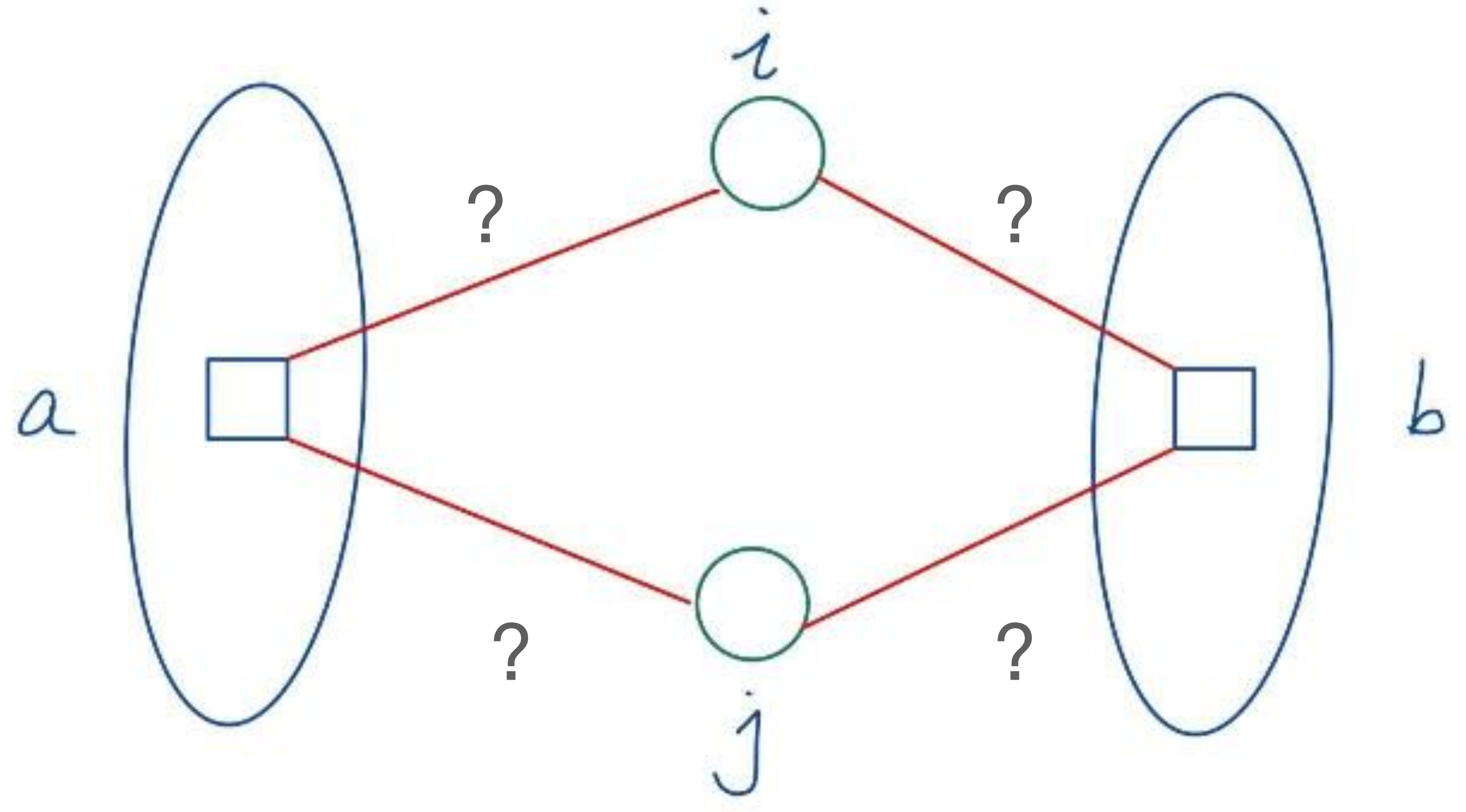
H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

- Norm bounds in 2-steps
  - Bound the local-value of each edge-labeling
  - And then sum over all F/R/S/H-edge-labeling of a given shape



A Taste of Our Local Analysis

Edge-factors

F/S/R: $1/\sqrt{d}$

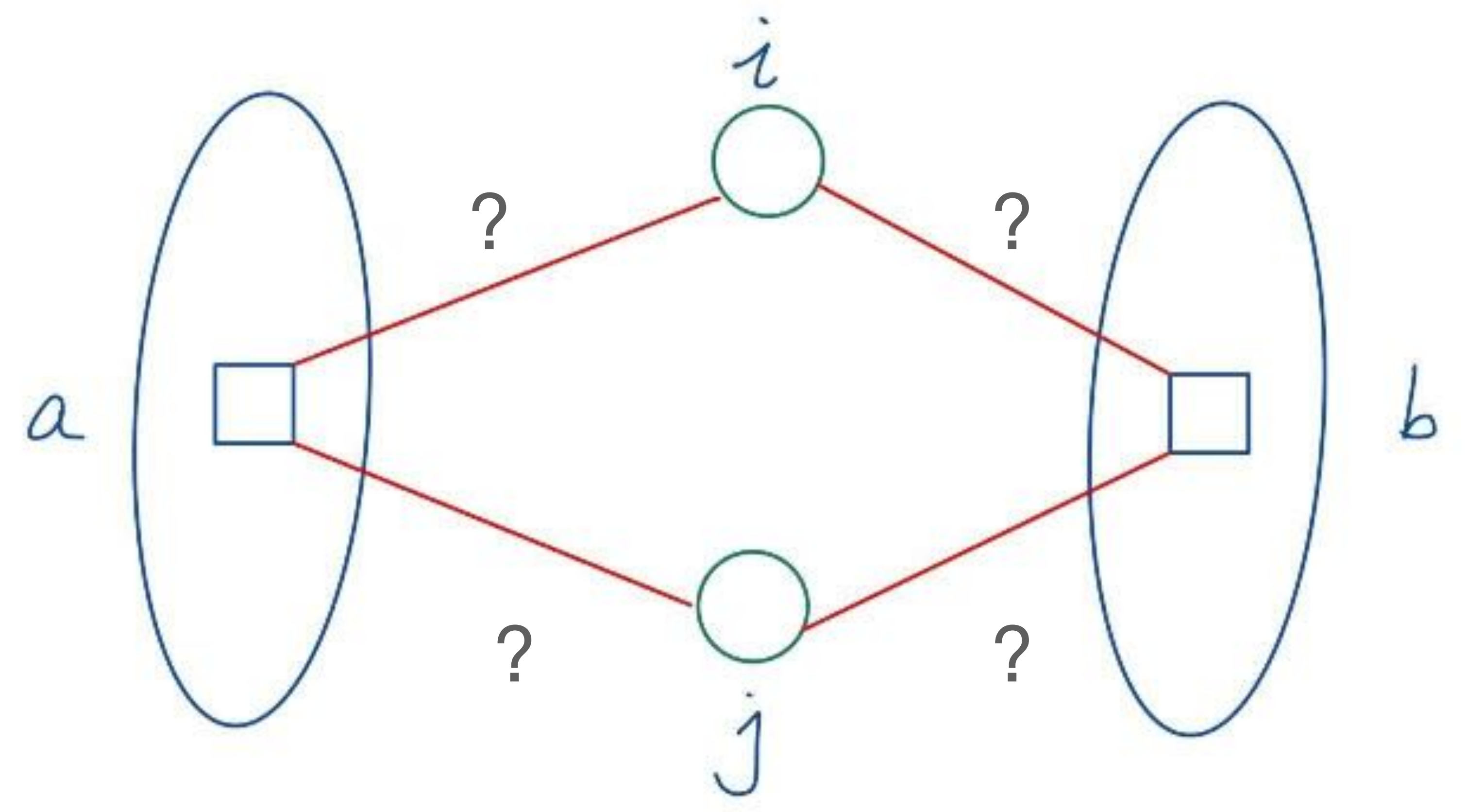
H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

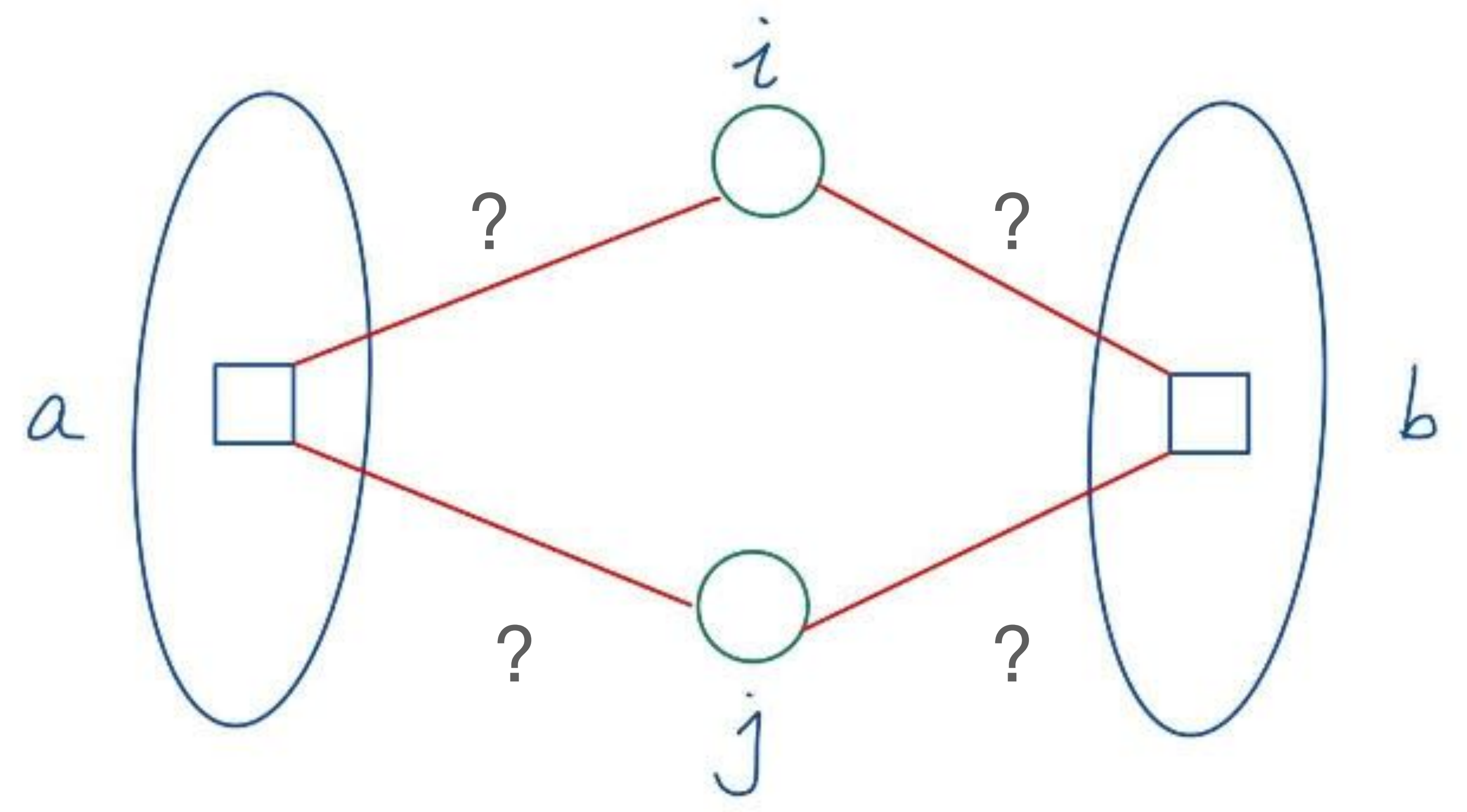
- Norm bounds in 2-steps
  - Bound the local-value of each edge-labeling
  - And then sum over all F/R/S/H-edge-labeling of a given shape
- Summing over 4^4-edge labelings gives a bound of $O(\sqrt{c})$ for A^{-1}



A Taste of Our Local Analysis

Edge-factors		Vertex-factors	
F/S/R:	$1/\sqrt{d}$	First/Last:	$\sqrt{d}$ (or $\sqrt{m}$)
H:	$\sqrt{q}/\sqrt{d}$	Arrival via S/H:	q^3

- Norm bounds in 2-steps
  - Bound the local-value of each edge-labeling
  - And then sum over all F/R/S/H-edge-labeling of a given shape
- Summing over 4^4-edge labelings gives a bound of $O(\sqrt{c})$ for A^{-1}
- **Exponentially** many more matrices that come up —> develop a **systematic** analysis



A Taste of Our Local Analysis

Edge-factors		Vertex-factors	
F/S/R:	$1/\sqrt{d}$	First/Last:	$\sqrt{d}$ (or $\sqrt{m}$)
H:	$\sqrt{q}/\sqrt{d}$	Arrival via S/H:	q^3

- Norm bounds in 2-steps
  - Bound the local-value of each edge-labeling
  - And then sum over all F/R/S/H-edge-labeling of a given shape
- Summing over 4^4-edge labelings gives a bound of $O(\sqrt{c})$ for A^{-1}
- **Exponentially** many more matrices that come up —> develop a **systematic** analysis

A Taste of Our Local Analysis

Edge-factors

F/S/R: $1/\sqrt{d}$

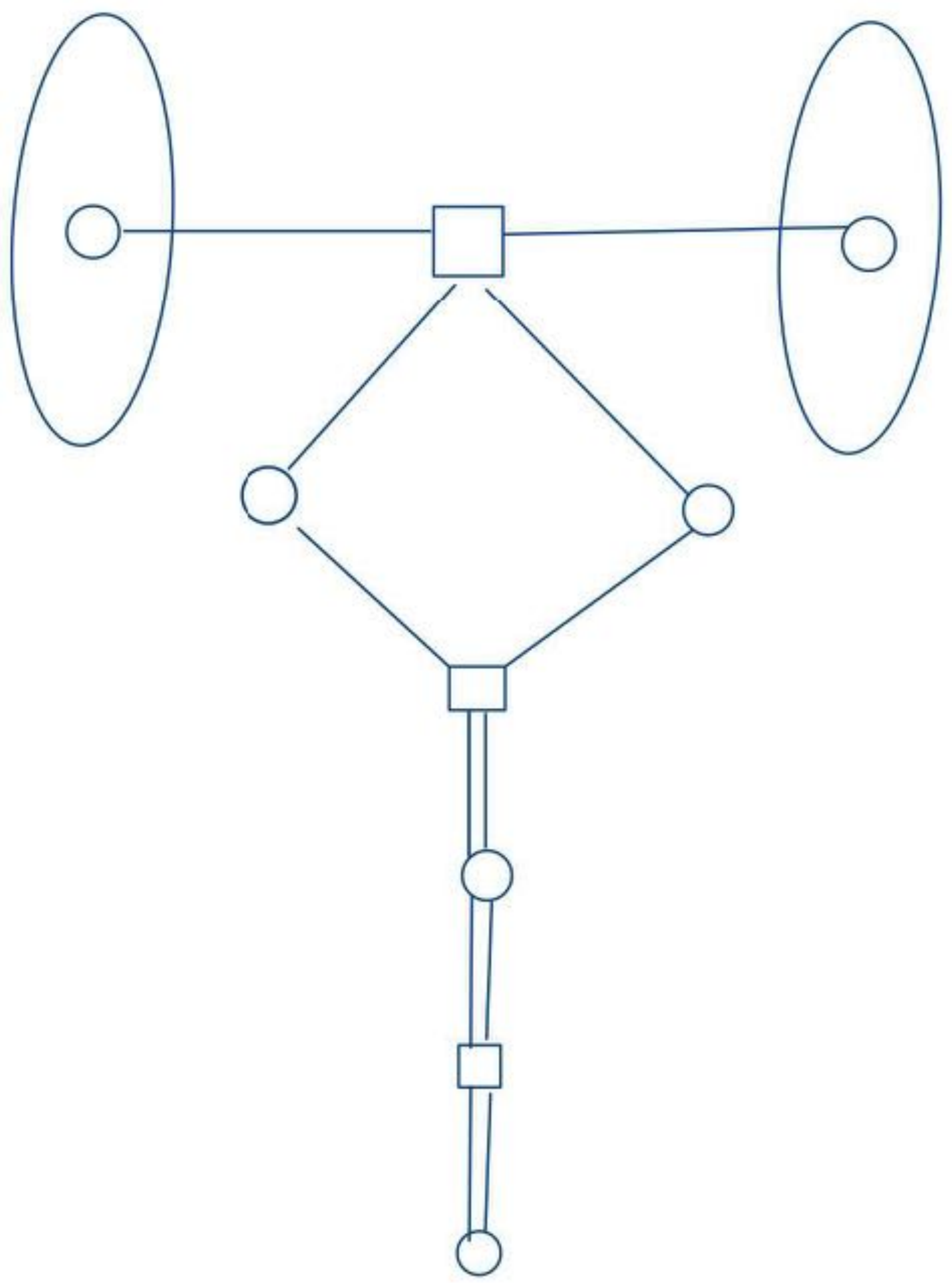
H: $\sqrt{q}/\sqrt{d}$

Vertex-factors

First/Last: $\sqrt{d}$ (or $\sqrt{m}$)

Arrival via S/H: q^3

- Norm bounds in 2-steps
  - Bound the local-value of each edge-labeling
  - And then sum over all F/R/S/H-edge-labeling of a given shape
- Summing over 4^4-edge labelings gives a bound of $O(\sqrt{c})$ for A^{-1}
- **Exponentially** many more matrices that come up —> develop a **systematic** analysis



Wrapping up

1. Ellipsoid Fitting Conjecture
2. Constructing an Ellipsoid
3. Analysis via Graph Matrices
4. **A Local Machinery for Tight Norm Bounds**

Thank you!

Open Question [SCPW Conjecture] For all $\epsilon > 0$, and for sufficiently large d,

- **(Positive)** If $m \leq (1-\epsilon)\frac{d^2}{4}$, there exists such an ellipsoid w.h.p.
    - Our construction experimentally fails [PTVW '22]
- **(Negative)** If $m \geq (1+\epsilon)\frac{d^2}{4}$, there does not exist such an ellipsoid w.h.p.

References

[GJJ+20] Mrinalkanti Ghosh, Fernando Granha Jeronimo, Chris Jones, Aaron Potechin, and Goutham Rajendran. *Sum-of-squares lower bounds for Sherrington-Kirkpatrick via planted affine planes.*

[KD22] Daniel M Kane and Ilias Diakonikolas. *A Nearly Tight Bound for Fitting an Ellipsoid to Gaussian Random Points*. arXiv preprint arXiv:2212.11221, 2022.

[PTVW22] Aaron Potechin, Paxton Turner, Prayaag Venkat, and Alexander S Wein. *Near-optimal fitting of ellipsoids to random points*. arXiv preprint arXiv:2208.09493, 2022.

[SCPW12] James Saunderson, Venkat Chandrasekaran, Pablo A Parrilo, and Alan S Willsky. Diagonal and low-rank matrix decompositions, correlation matrices, and ellipsoid fitting. SIAM Journal on Matrix Analysis and Applications, 33(4):1395–1416, 2012.

[SPW13] James Saunderson, Pablo A Parrilo, and Alan S Willsky. Diagonal and low-rank decompositions and fitting ellipsoids to random points. In 52nd IEEE Conference on Decision and Control, pages 6031–6036. IEEE, 2013.