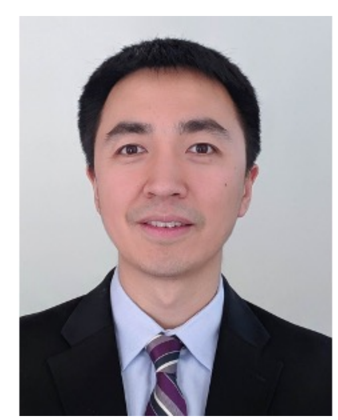
# Optimality of Random Regular Graphs for Sparse Network Designs

Jiaming Xu

The Fuqua School of Business
Duke University

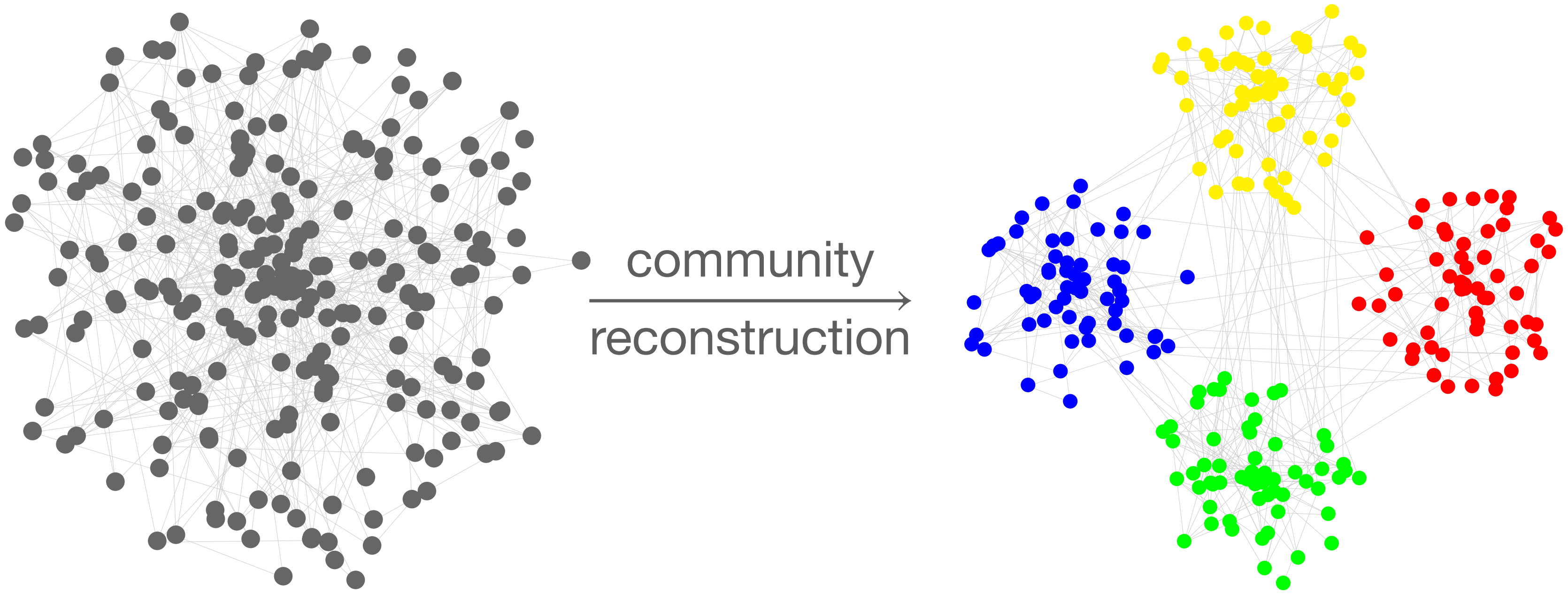
Joint work with
Weijia Li (Tsinghua), Xiaochun (Nora) Niu (Duke), and Yehua Wei (Duke)





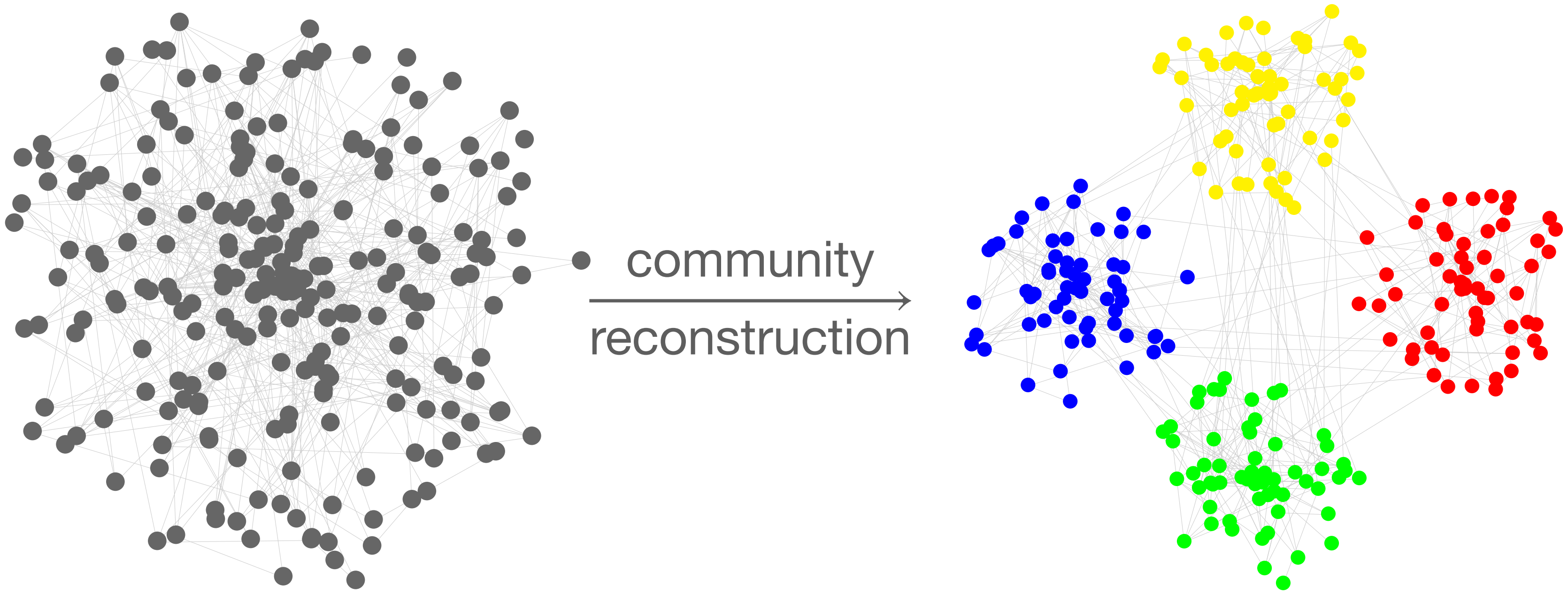
# Network Theory

- Networked systems and network data are everywhere
  - Statistical network analysis: Observe network data $\longrightarrow$ Infer hidden structures



# Network Theory

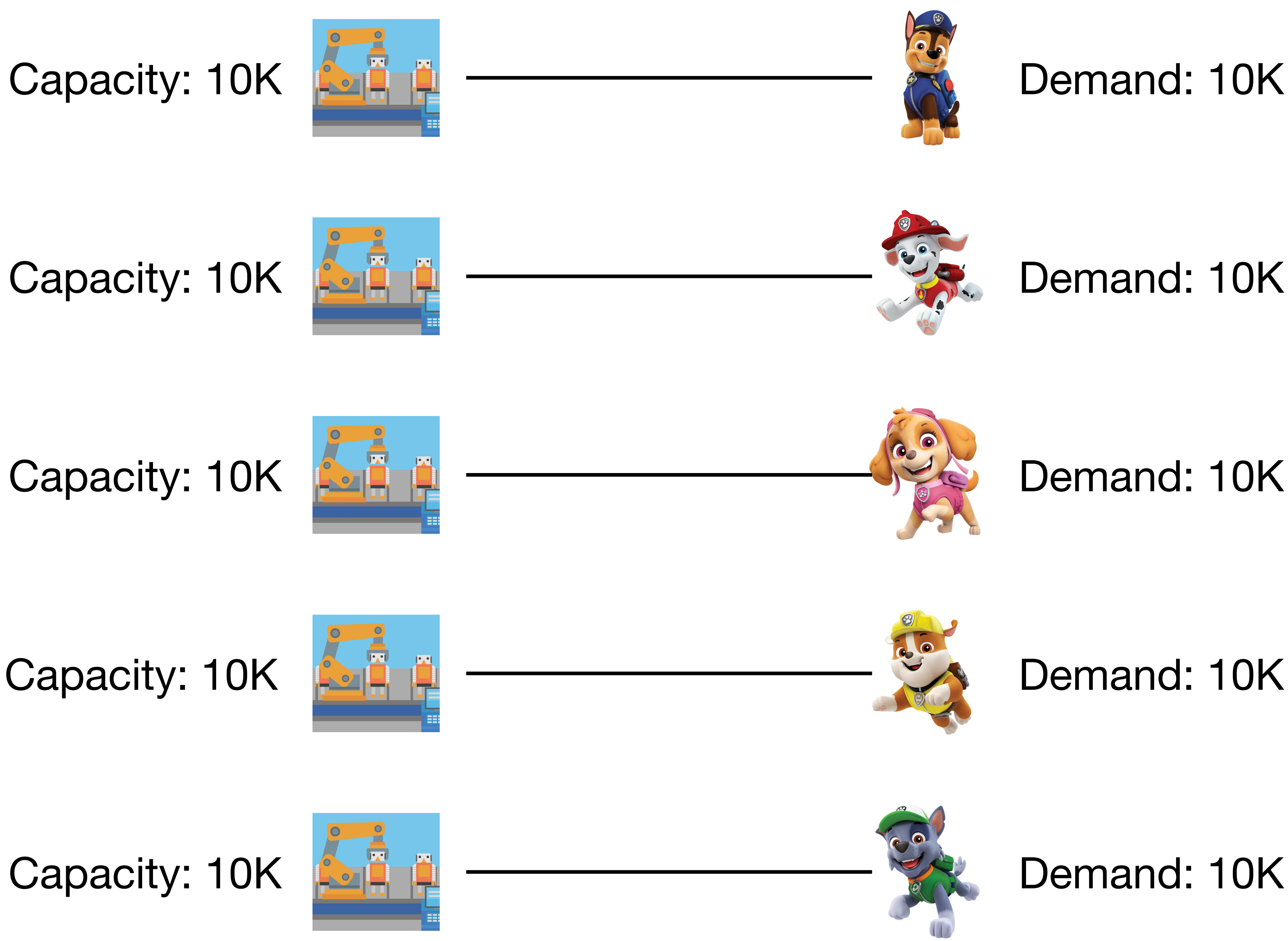
- Networked systems and network data are everywhere
  - Statistical network analysis: Observe network data $\longrightarrow$ Infer hidden structures




  - Network design: Design network structure $\longrightarrow$ Enable efficient operations (This Talk)

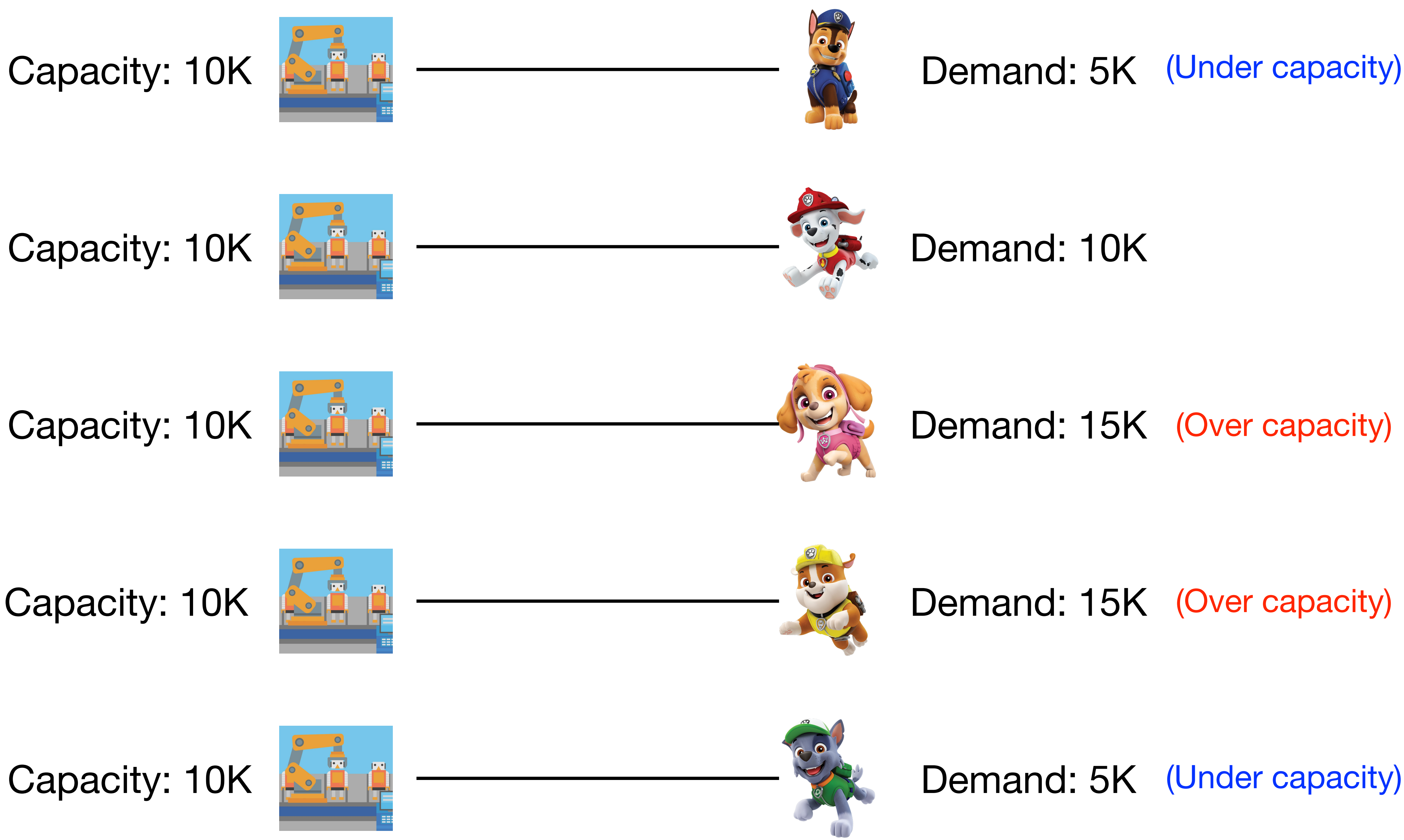
# Flexibility Network Design Under Uncertainty

A Toy Example: Ideal setting



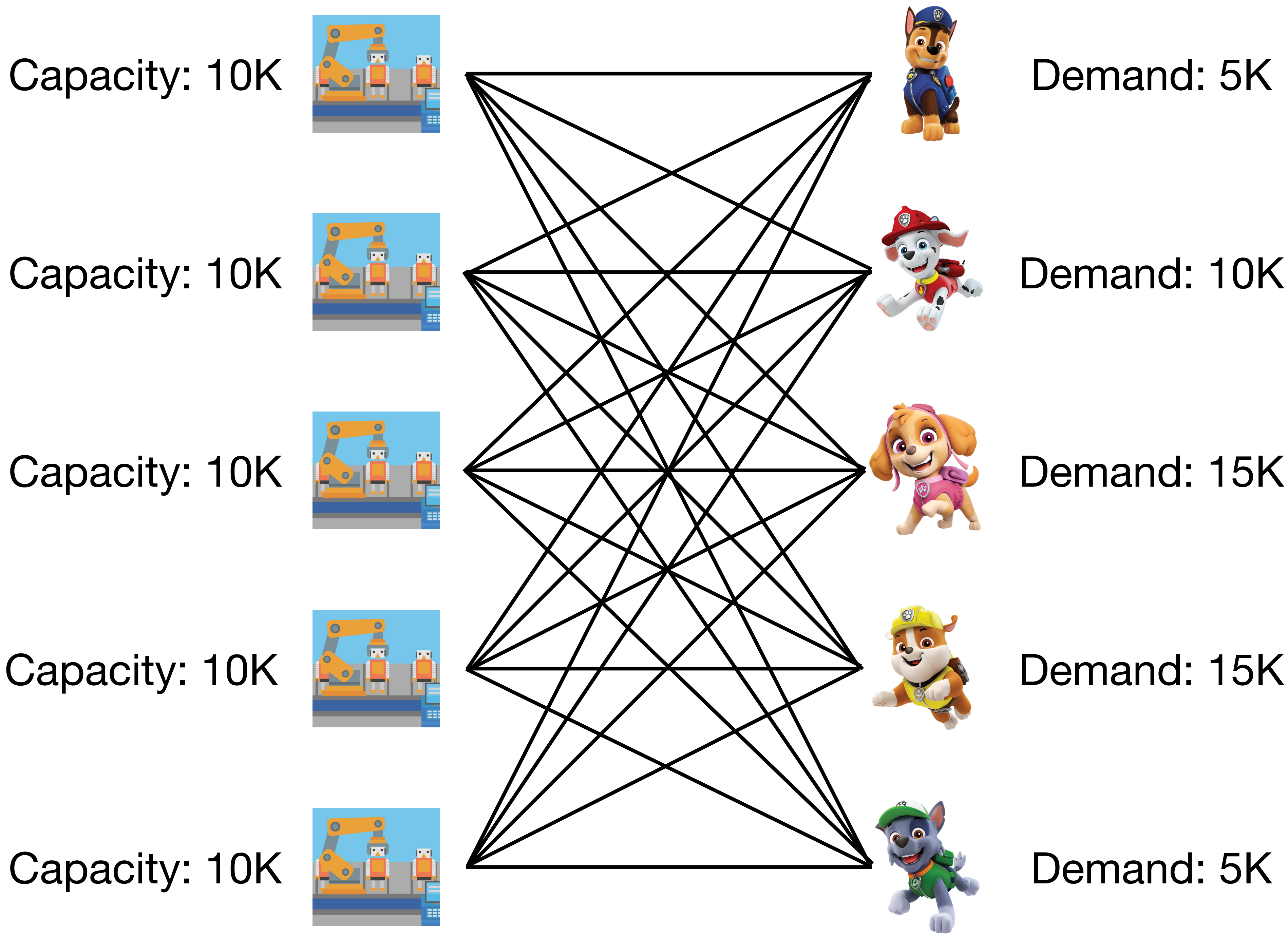
# Flexibility Network Design Under Uncertainty

A Toy Example: Actual demand may change dramatically depending on popularity



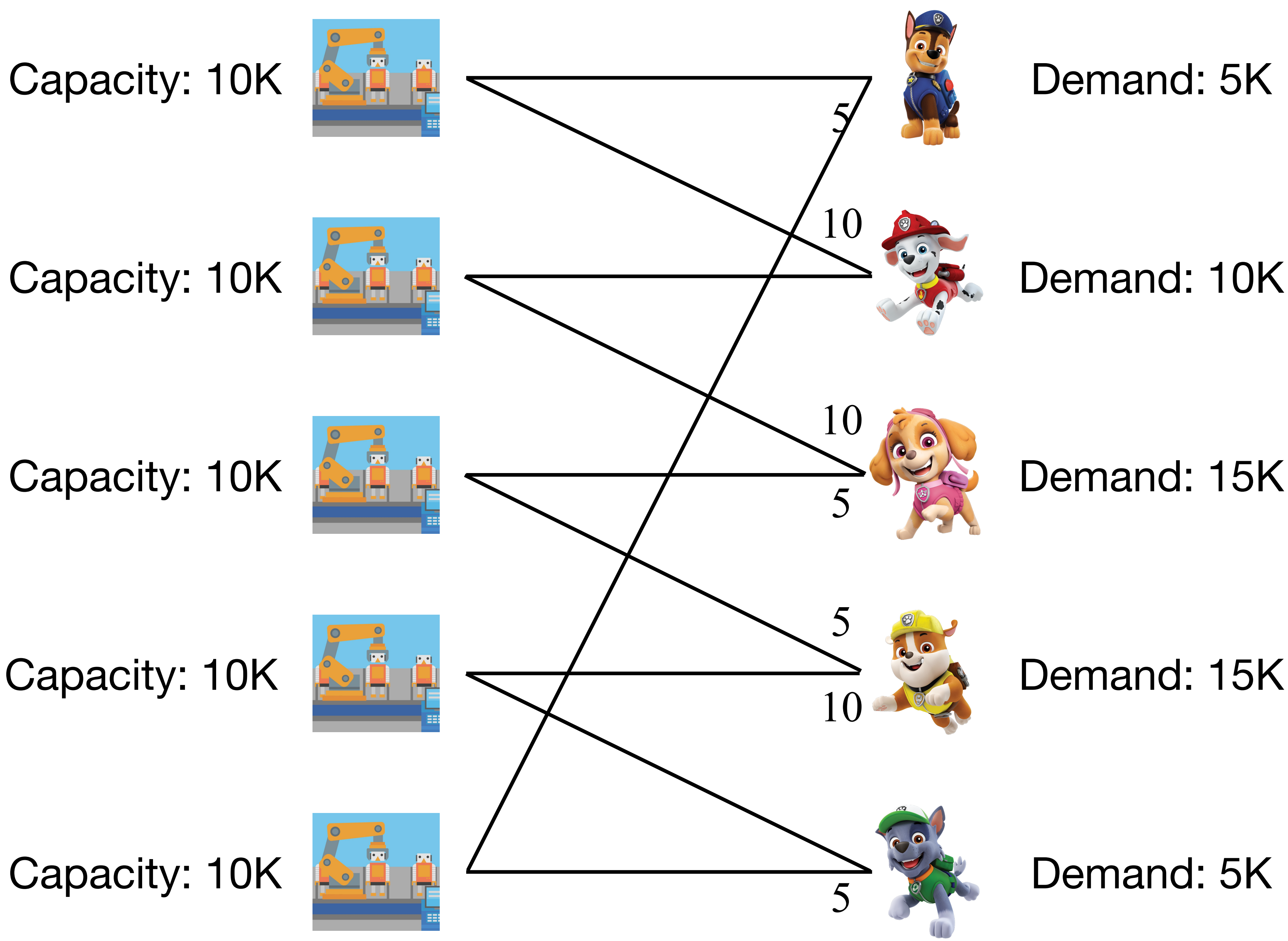
# Flexibility Network Design Under Uncertainty

A Toy Example: Configure all production lines to be fully flexible, but too costly



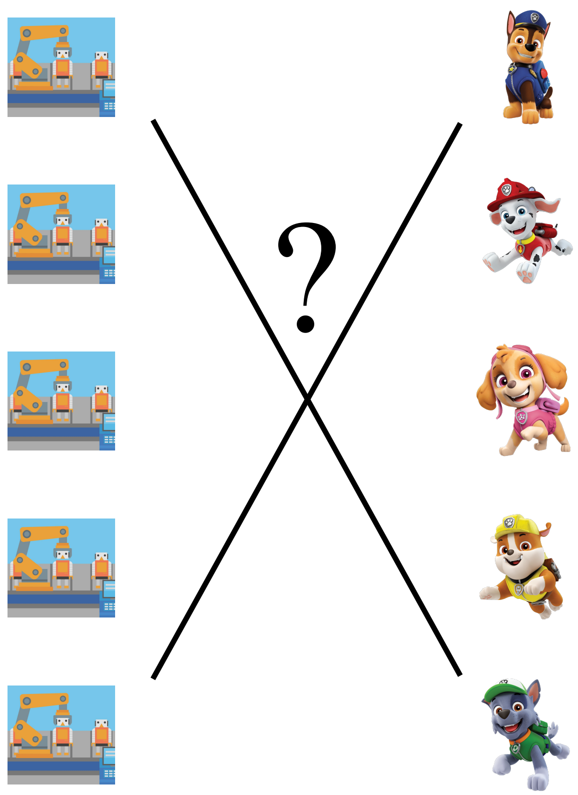
# Flexibility Network Design Under Uncertainty

A Toy Example: Add a few flexible edges (long-chain design)



# Flexibility Network Design Under Uncertainty

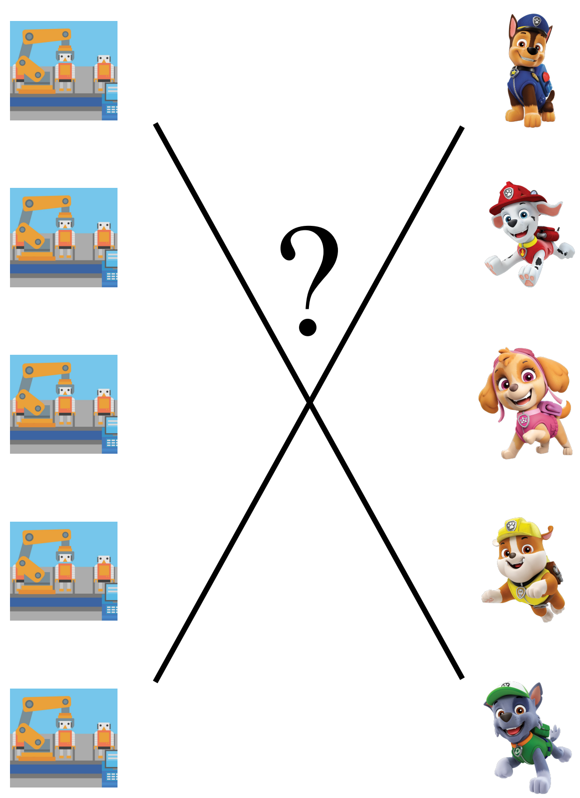
A Toy Example:




**Question**: How to design a sparse graph that achieves efficiency as close as the complete graph?

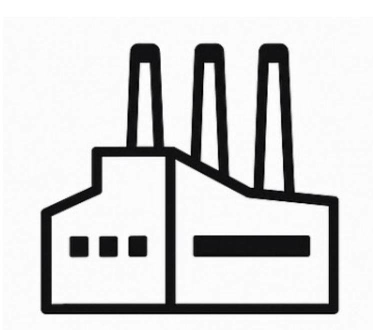
# Flexibility Network Design Under Uncertainty

A Toy Example:

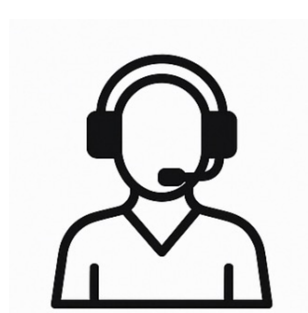



**Question**: How to design a sparse graph that achieves efficiency as close as the complete graph?

Many real-world applications: match capacitated resources with uncertain demand

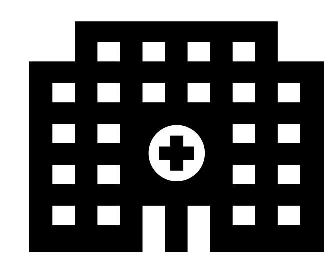


Manufacturing



Service operations

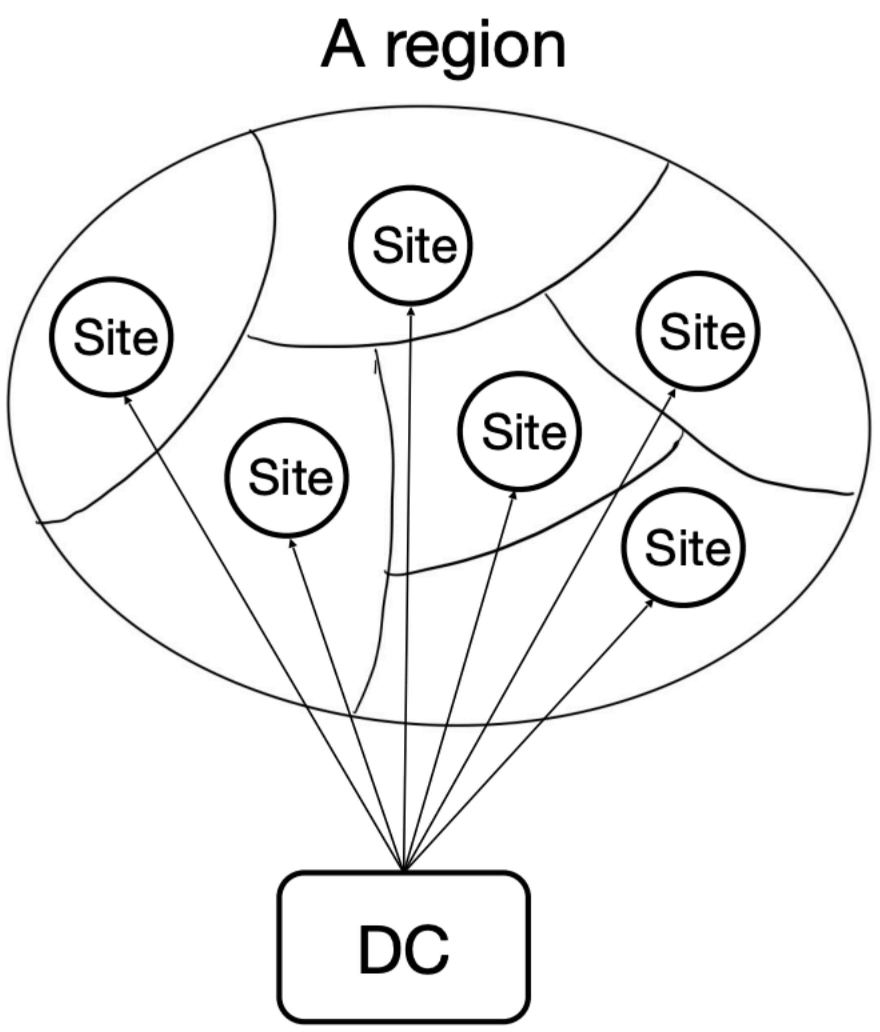
Transportation



Healthcare

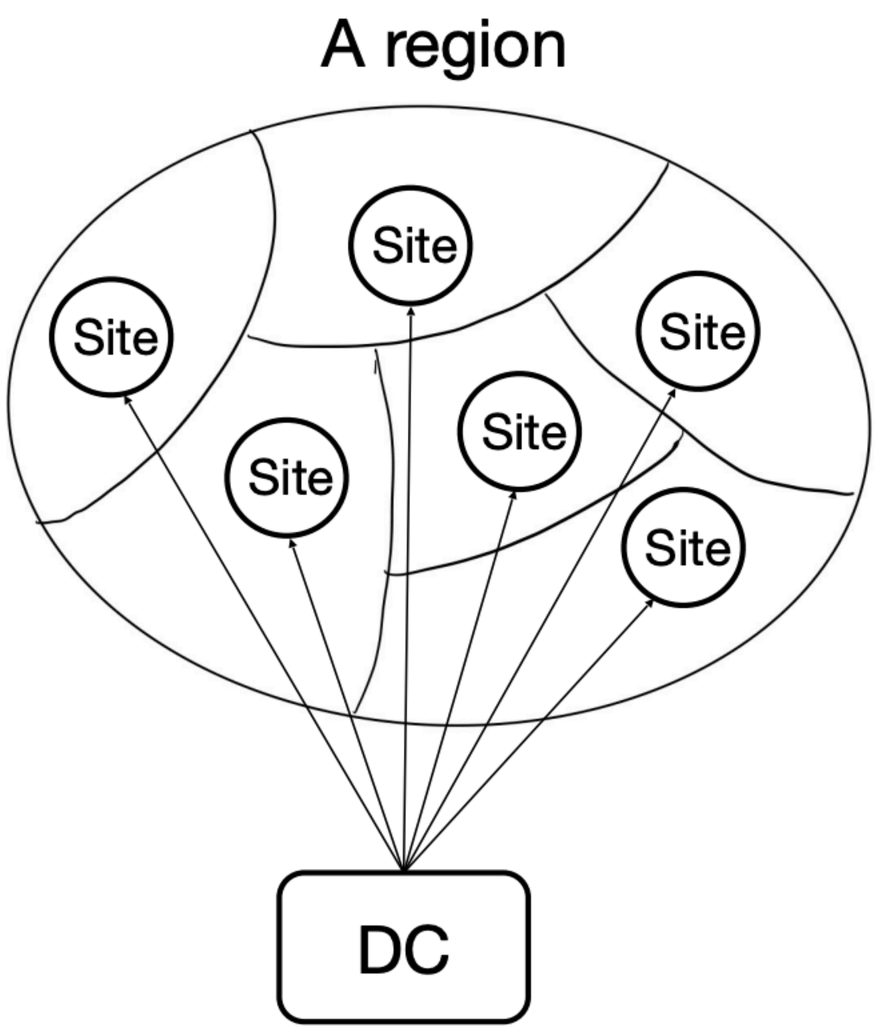
# Motivation: Transportation flexibility

- Proposed in [Feng-Caldentey-Xin-Zhong-Wang-Hu ’24]
- Logistics networks are divided into regions
- With each region:
  - 1 distribution center (DC)
  - Multiple delivery stations (Sites)



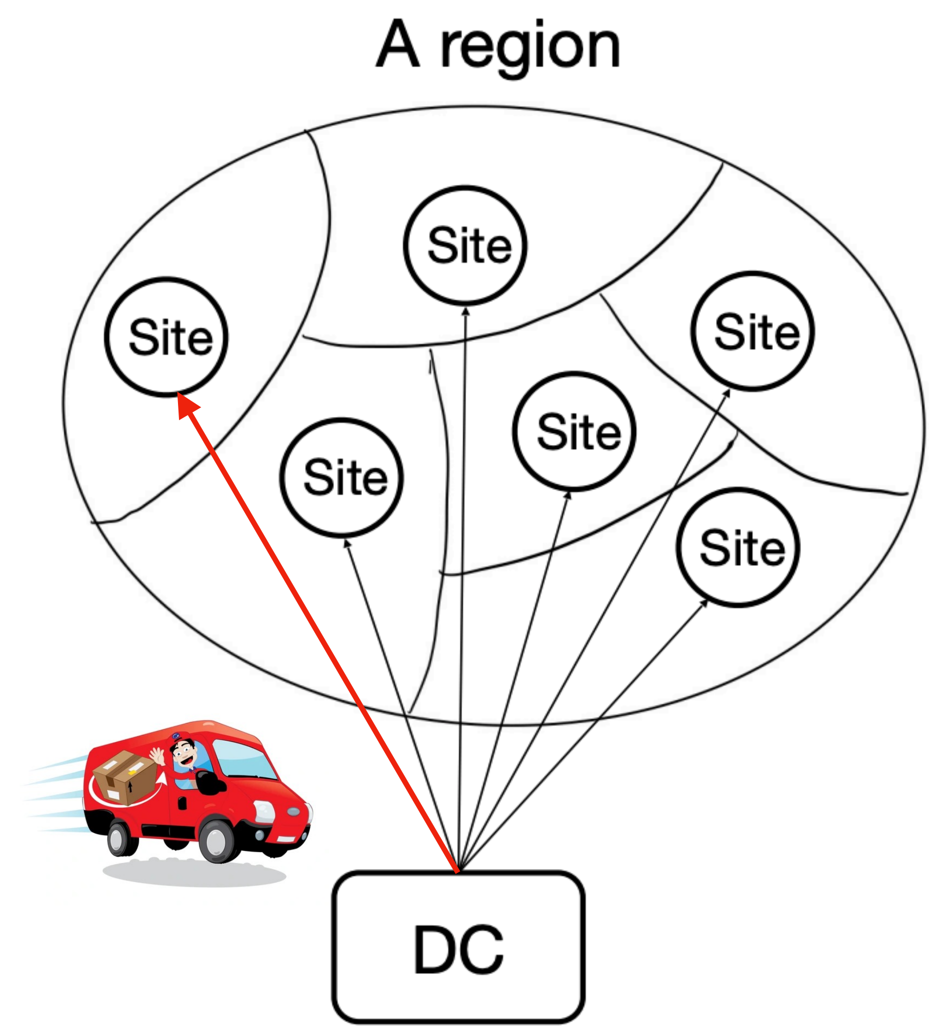
# Motivation: Transportation flexibility

- Proposed in [Feng-Caldentey-Xin-Zhong-Wang-Hu '24]
- Logistics networks are divided into regions
- With each region:
  - 1 distribution center (DC)
  - Multiple delivery stations (Sites)
- On each day:
  - Demands are revealed at each site
  - Packages are transported from DC to Sites



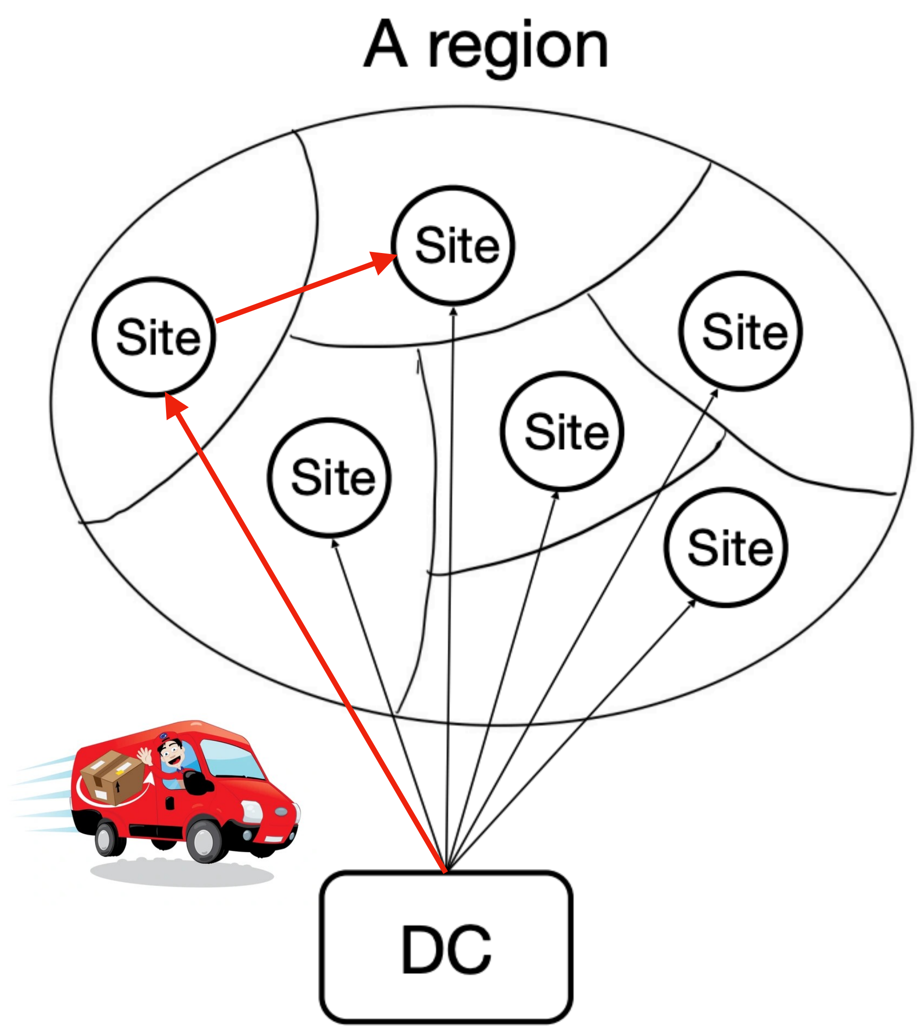
# Motivation: Transportation flexibility

- Proposed in [Feng-Caldentey-Xin-Zhong-Wang-Hu ’24]
- Logistics networks are divided into regions
- With each region:
  - 1 distribution center (DC)
  - Multiple delivery stations (Sites)
- On each day:
  - Demands are revealed at each site
  - Packages are transported from DC to Sites
- Current practice
  - Direct shipping (mostly): one truck one site



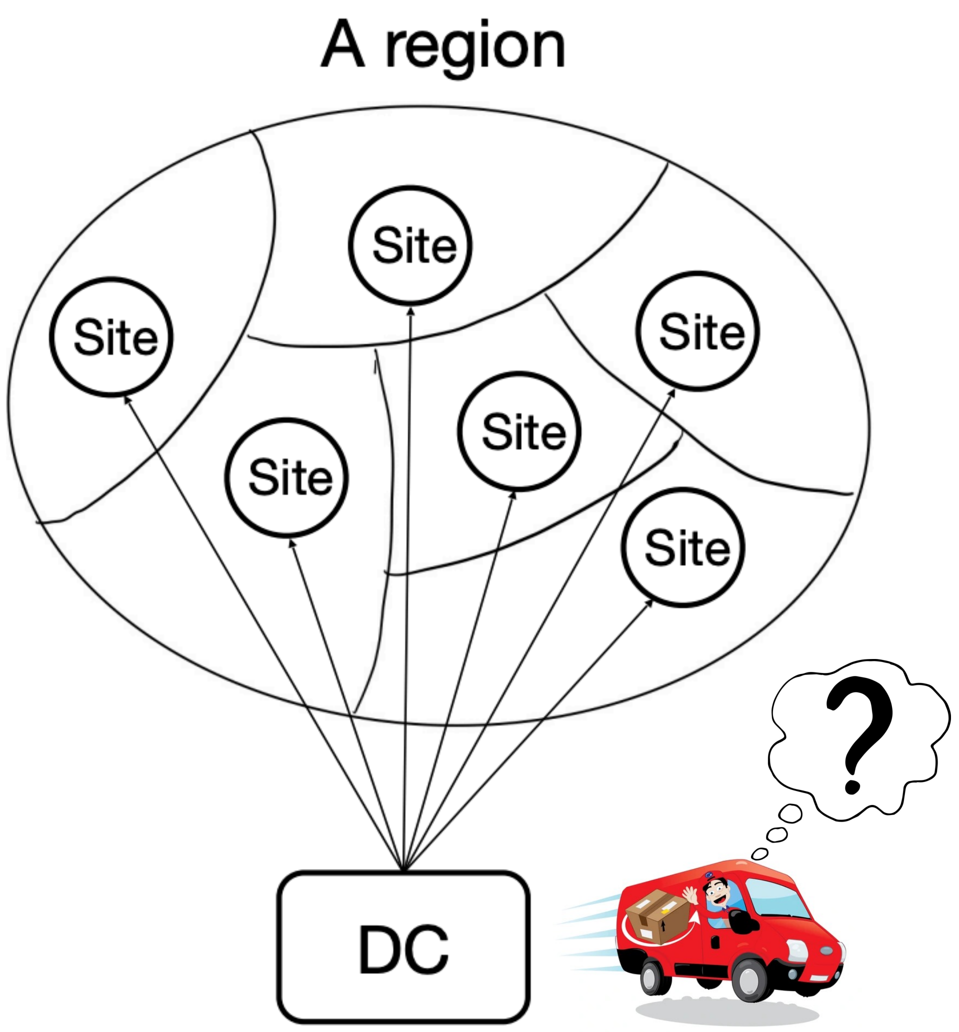
# Motivation: Transportation flexibility

- Proposed in [Feng-Caldentey-Xin-Zhong-Wang-Hu '24]
- Logistics networks are divided into regions
- With each region:
  - 1 distribution center (DC)
  - Multiple delivery stations (Sites)
- On each day:
  - Demands are revealed at each site
  - Packages are transported from DC to Sites
- Current practice
  - Direct shipping (mostly): one truck one site
- Room for Improvement
  - Indirect shipping: one truck serves multiple (two) sites

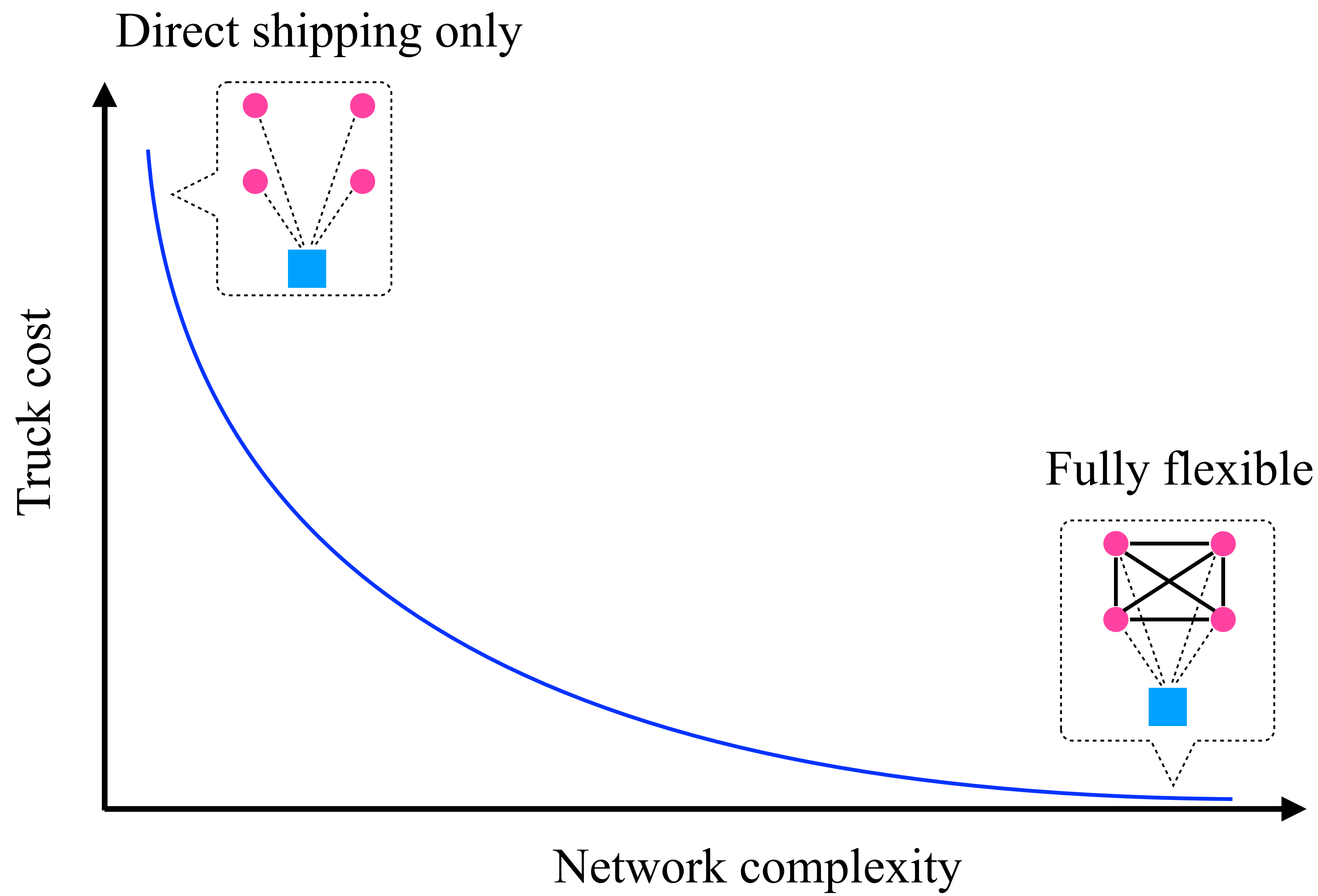


# Motivation: Transportation flexibility

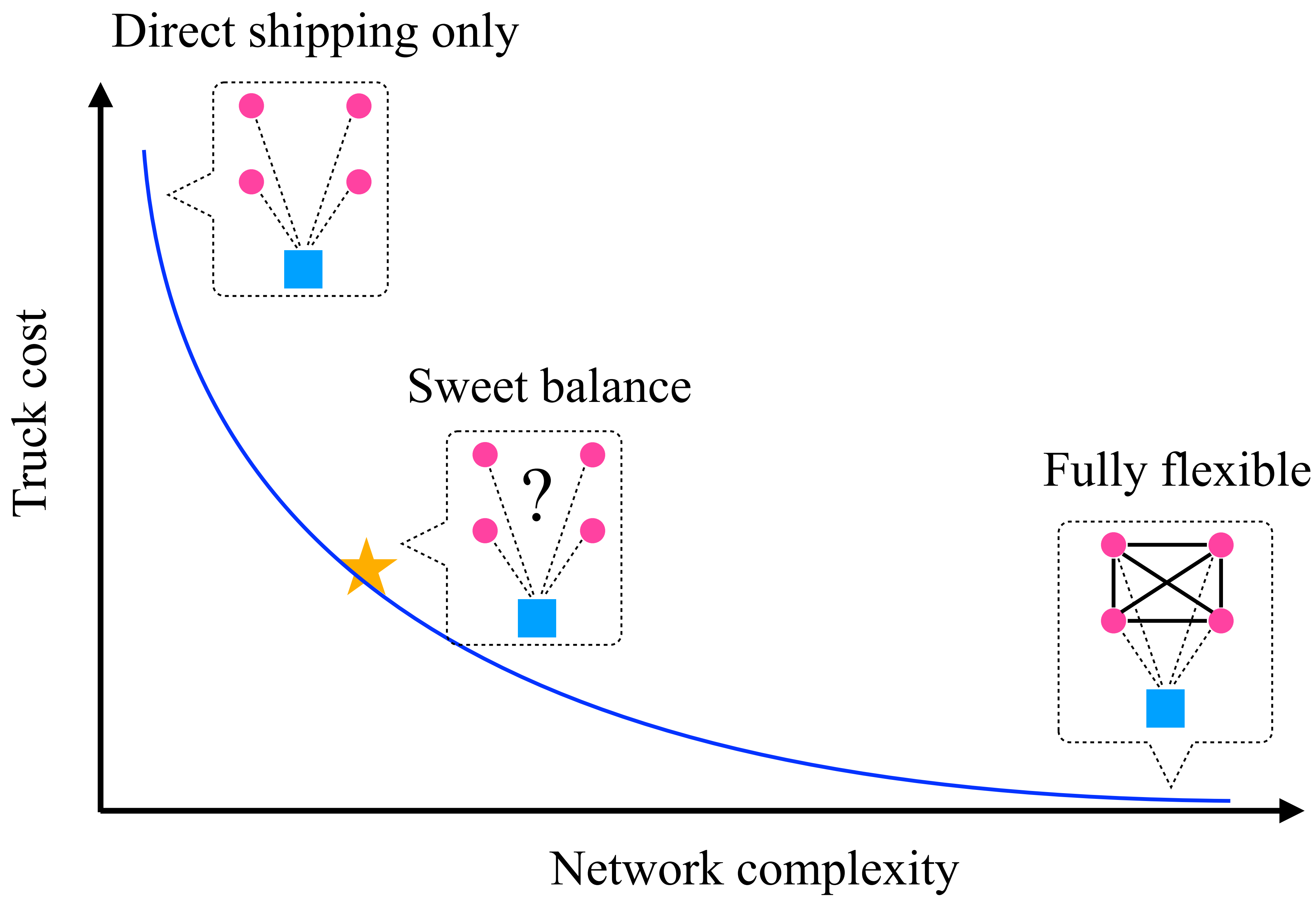
- Proposed in [Feng-Caldentey-Xin-Zhong-Wang-Hu ’24]
- Logistics networks are divided into regions
- With each region:
  - 1 distribution center (DC)
  - Multiple delivery stations (Sites)
- On each day:
  - Demands are revealed at each site
  - Packages are transported from DC to Sites
- Current practice
  - Direct shipping (mostly): one truck one site
- Room for Improvement
  - Indirect shipping: one truck serves multiple (two) sites
- Question
  - How to design delivery routes for indirect shipping under uncertain demands?



# Tradeoff between truck cost and network complexity



# Tradeoff between truck cost and network complexity




How to design a sparse network that achieves low truck cost?

# Literature on Flexibility Network Design

- Flexibility design proposed in [Jordan-Graves'95] and widely studied since then

# Literature on Flexibility Network Design

- Flexibility design proposed in [Jordan-Graves'95] and widely studied since then
- *"A little bit flexibility goes a long way"* (survey [Wang-Wang-Zhang'21])

# Literature on Flexibility Network Design

- Flexibility design proposed in [Jordan-Graves'95] and widely studied since then
- *"A little bit flexibility goes a long way"* (survey [Wang-Wang-Zhang'21])
  - Long-chain design: characterize the performance ratio w.r.t. full flexibility [Chou-Chua-Teo-Zheng'10, Simchi-Levi-Wei'12, Wang-Zhang'15]

# Literature on Flexibility Network Design

- Flexibility design proposed in [Jordan-Graves'95] and widely studied since then
- *"A little bit flexibility goes a long way"* (survey [Wang-Wang-Zhang'21])
  - Long-chain design: characterize the performance ratio w.r.t. full flexibility [Chou-Chua-Teo-Zheng'10, Simchi-Levi-Wei'12, Wang-Zhang'15]
  - Graph expander design [Chou-Chua-Teo-Zheng'11, Chen-Zhang-Zhou'15, Feng-Caldentey-Xin-Zhong-Wang-Hu'24]: $\Theta(\log(1/\epsilon))$ edges per node is both sufficient and necessary for achieving $(1-\epsilon)$ performance ratio

# Literature on Flexibility Network Design

- Flexibility design proposed in [Jordan-Graves'95] and widely studied since then
- *"A little bit flexibility goes a long way"* (survey [Wang-Wang-Zhang'21])
  - Long-chain design: characterize the performance ratio w.r.t. full flexibility [Chou-Chua-Teo-Zheng'10, Simchi-Levi-Wei'12, Wang-Zhang'15]
  - Graph expander design [Chou-Chua-Teo-Zheng'11, Chen-Zhang-Zhou'15, Feng-Caldentey-Xin-Zhong-Wang-Hu'24]: $\Theta(\log(1/\epsilon))$ edges per node is both sufficient and necessary for achieving $(1-\epsilon)$ performance ratio
- Similar principle has been shown in many other related flexible systems: queueing systems [Tsitsiklis-Xu'17, Rutten-Mukherjee'23], dynamic resource allocation [Dong-Hu-Wang'24], online platforms [Freund-Martin-Zhao'24]

# Literature on Flexibility Network Design

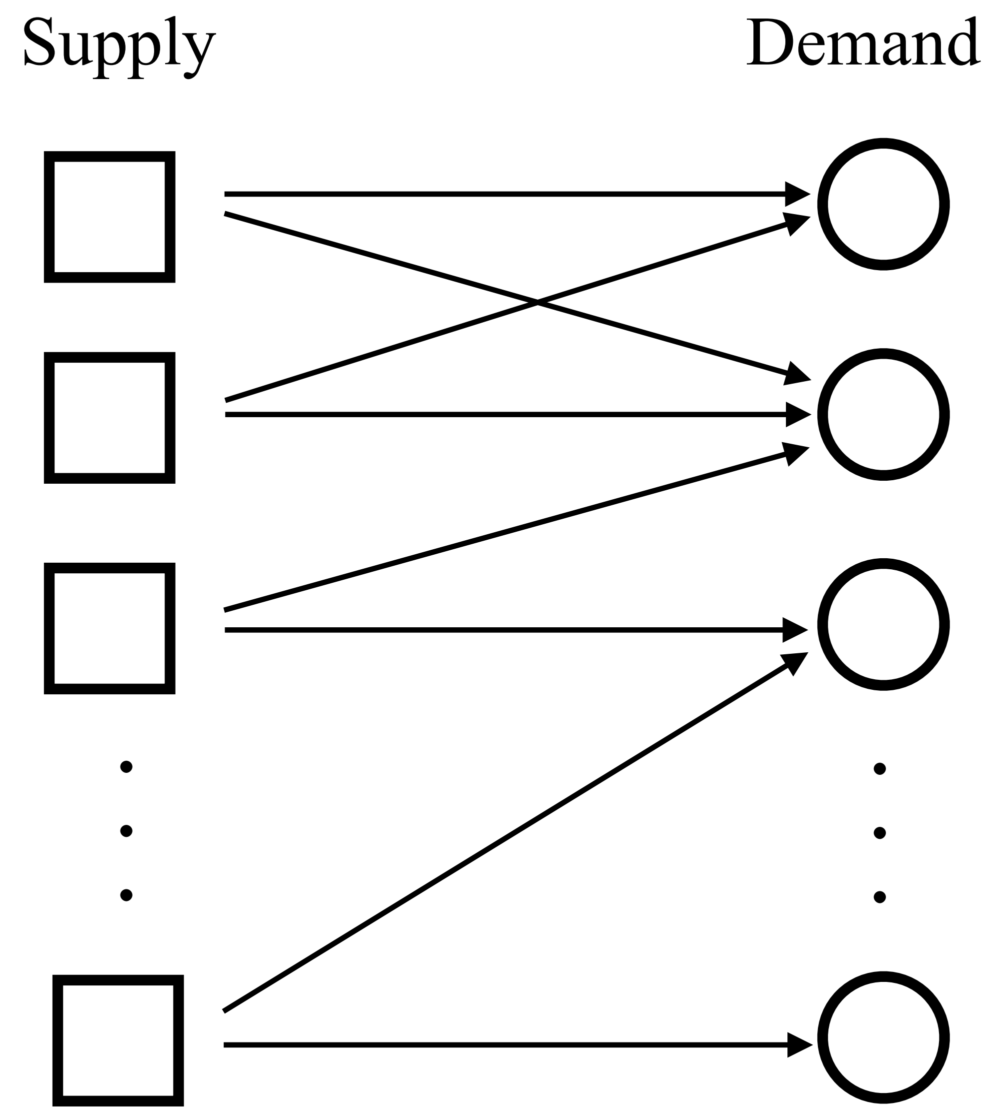
- Flexibility design proposed in [Jordan-Graves'95] and widely studied since then
- *"A little bit flexibility goes a long way"* (survey [Wang-Wang-Zhang'21])
  - Long-chain design: characterize the performance ratio w.r.t. full flexibility [Chou-Chua-Teo-Zheng'10, Simchi-Levi-Wei'12, Wang-Zhang'15]
  - Graph expander design [Chou-Chua-Teo-Zheng'11, Chen-Zhang-Zhou'15, Feng-Caldentey-Xin-Zhong-Wang-Hu'24]: $\Theta(\log(1/\epsilon))$ edges per node is both sufficient and necessary for achieving $(1-\epsilon)$ performance ratio
- Similar principle has been shown in many other related flexible systems: queueing systems [Tsitsiklis-Xu'17, Rutten-Mukherjee'23], dynamic resource allocation [Dong-Hu-Wang'24], online platforms [Freund-Martin-Zhao'24]

**Open Question**: What is the precise limit of flexibility needed for achieving $(1-\epsilon)$ performance ratio?

# Model

# Process Flexibility Model

- Bipartite graph $G = ([n], [n], E)$



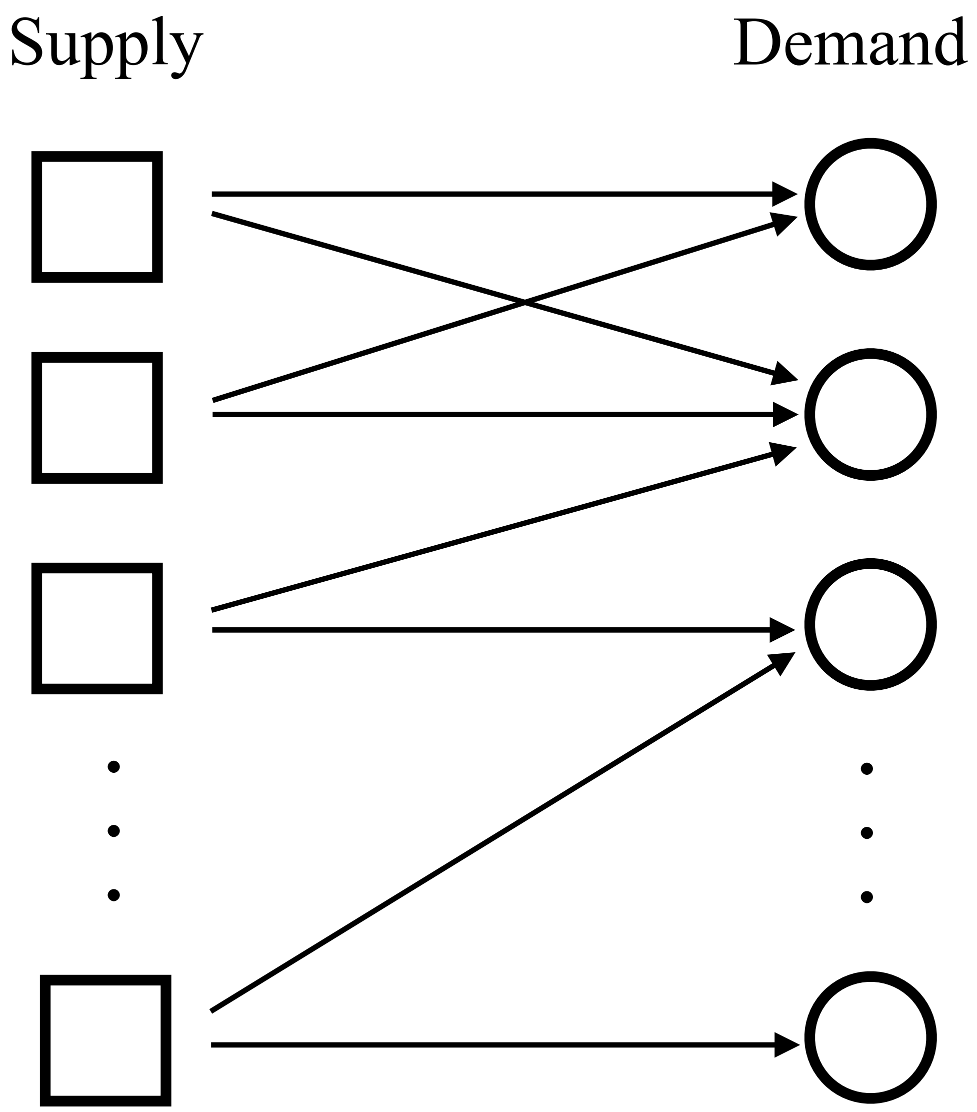
# Process Flexibility Model

- Bipartite graph $G = ([n], [n], E)$
- The total fulfilled demand

$$z(G,D) = \max_{x_{ij} \geq 0} \sum_{(i,j)\in E} x_{ij}$$

$$\text{s.t.} \quad \sum_{j:\,(i,j)\in E} x_{ij} \leq 1 \qquad \forall i$$

$$\sum_{i:\,(i,j)\in E} x_{ij} \leq D_j \qquad \forall j$$



# Process Flexibility Model

- Bipartite graph $G = ([n], [n], E)$
- The total fulfilled demand

$$z(G, D) = \max_{x_{ij} \geq 0} \sum_{(i,j)\in E} x_{ij}$$

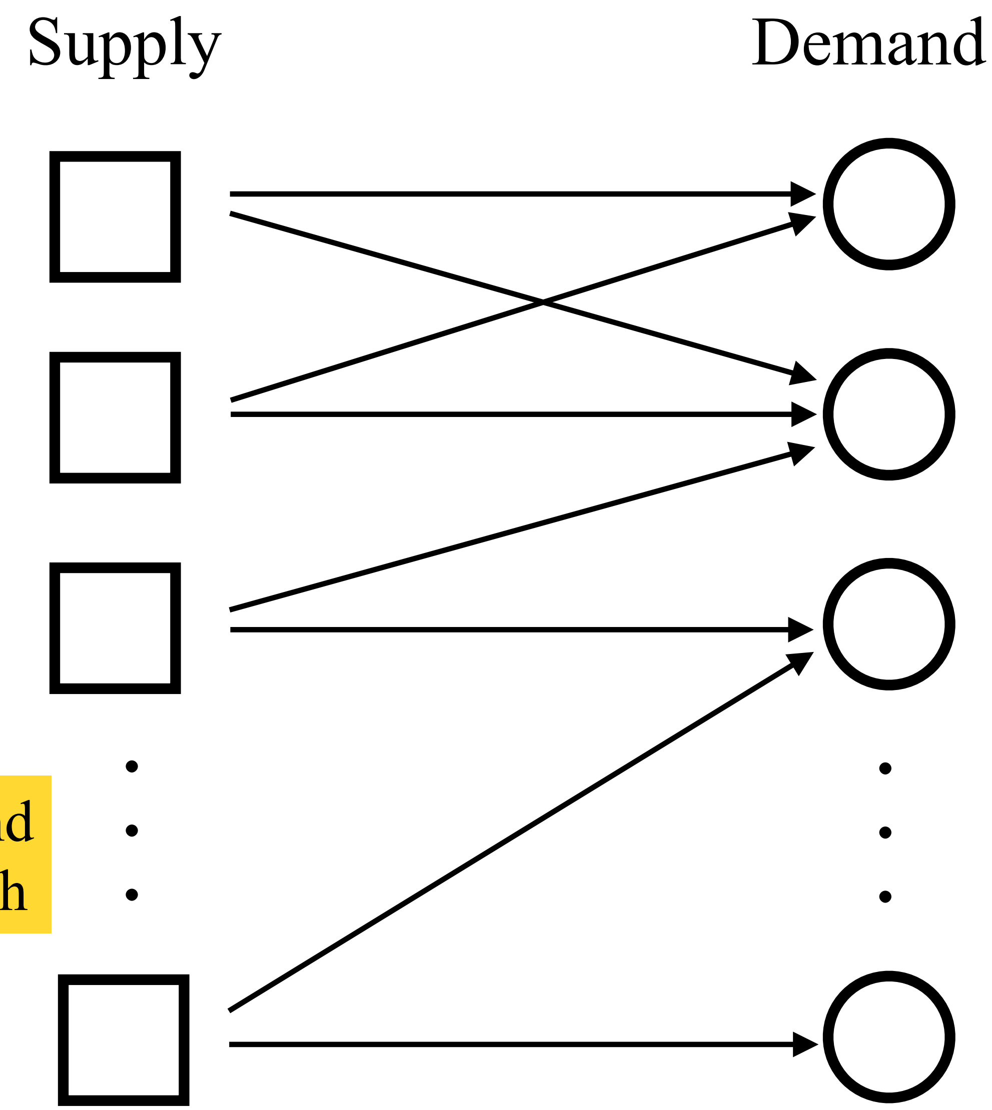
$$\text{s.t.} \quad \sum_{j:\,(i,j)\in E} x_{ij} \leq 1 \qquad \forall i$$

$$\sum_{i:\,(i,j)\in E} x_{ij} \leq D_j \qquad \forall j$$

- The expected loss of $G$:

Total fulfilled demand under complete graph

$$\mathbb{E}_D[L(G, D)] = \mathbb{E}_D[z(K_{n,n}, D) - z(G, D)]$$



# Process Flexibility Model

- Bipartite graph $G = ([n], [n], E)$
- The total fulfilled demand

$$z(G, D) = \max_{x_{ij} \geq 0} \sum_{(i,j) \in E} x_{ij}$$

$$\text{s.t.} \quad \sum_{j:\,(i,j) \in E} x_{ij} \leq 1 \qquad \forall i$$

$$\sum_{i:\,(i,j) \in E} x_{ij} \leq D_j \qquad \forall j$$

- The expected loss of $G$:

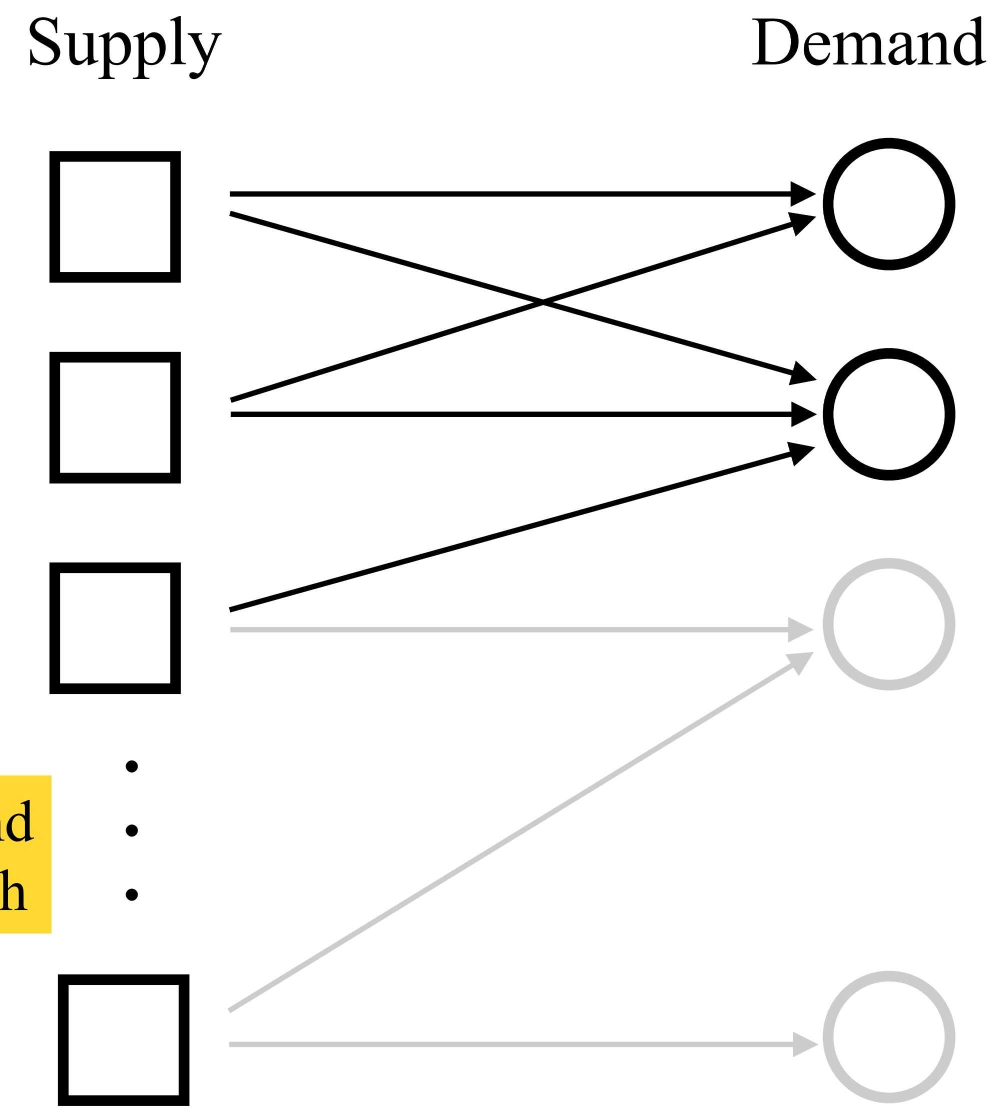
Total fulfilled demand under complete graph

$$\mathbb{E}_D[L(G, D)] = \mathbb{E}_D[z(K_{n,n}, D) - z(G, D)]$$

- IID Bernoulli demand with mean 1:

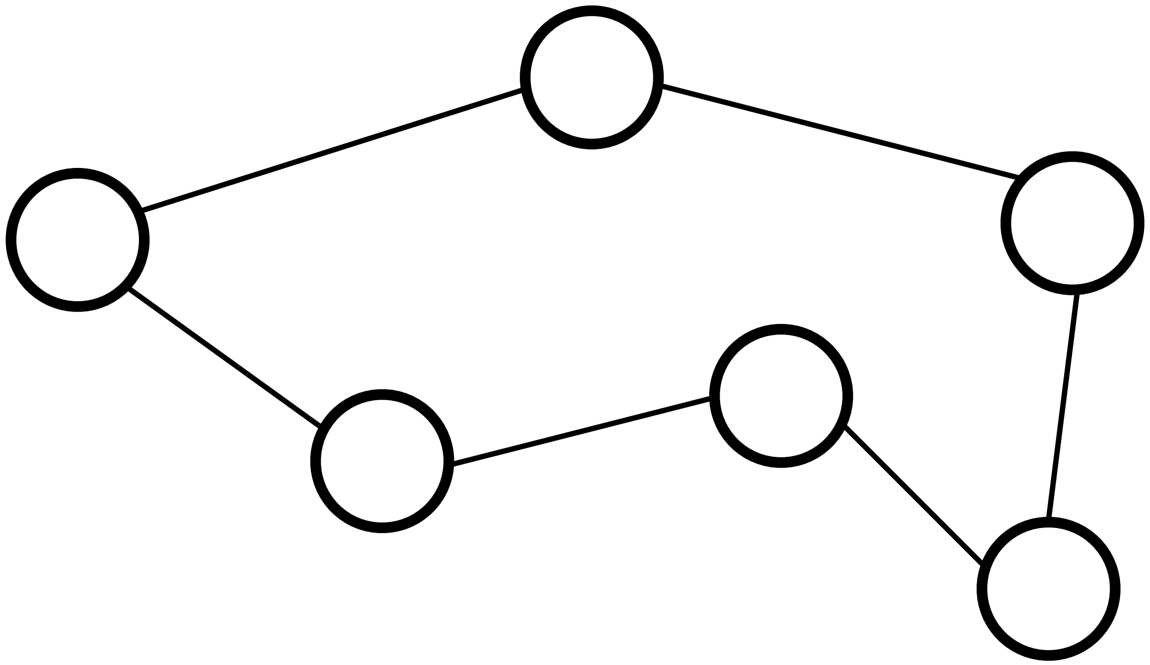
$$D_j = \begin{cases} 0 & \text{w.p. } q \\ 1/p & \text{w.p. } p = 1 - q \end{cases}$$

Random demand node deletion with probability $q$



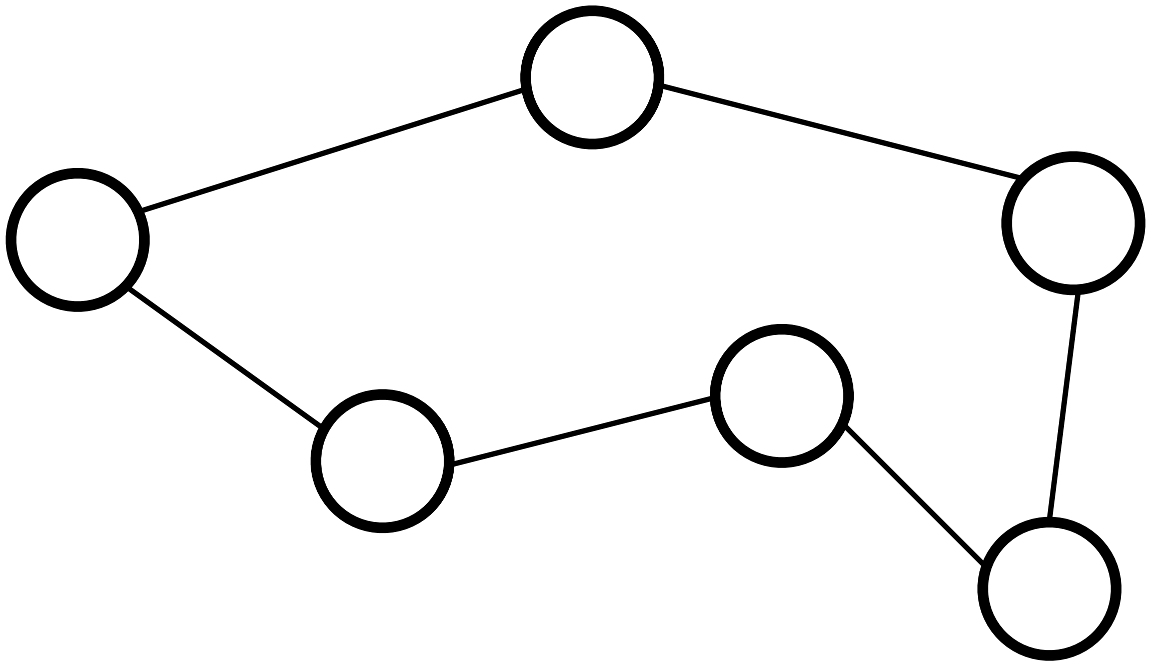
# Transportation Flexibility Model

- Given a graph design $G = ([n], E)$
  - Vertices = sites
  - Edges = delivery routes



# Transportation Flexibility Model

- Given a graph design $G = ([n], E)$
  - Vertices = sites
  - Edges = delivery routes
- The total fulfilled demand



$$\mu(G, D) = \max_{y_{ij} \in \{0,1\}} \sum_{(i,j) \in E} y_{ij}$$

$$\text{s.t.} \quad \sum_{j:\,(i,j) \in E} y_{ij} \leq 1 \qquad \forall i$$

$$y_{ij} \leq \mathbf{1}_{D_i + D_j \leq 1} \qquad \forall (i, j) \in E \longleftarrow$$

Two connected sites can be served by a single truck if their total demand does not exceed the truck capacity

# Transportation Flexibility Model

- Given a graph design $G = ([n], E)$
  - Vertices = sites
  - Edges = delivery routes
- The total fulfilled demand

$$\mu(G, D) = \max_{y_{ij} \in \{0,1\}} \sum_{(i,j)\in E} y_{ij}$$

$$\text{s.t.} \quad \sum_{j:\,(i,j)\in E} y_{ij} \le 1 \qquad \forall i$$

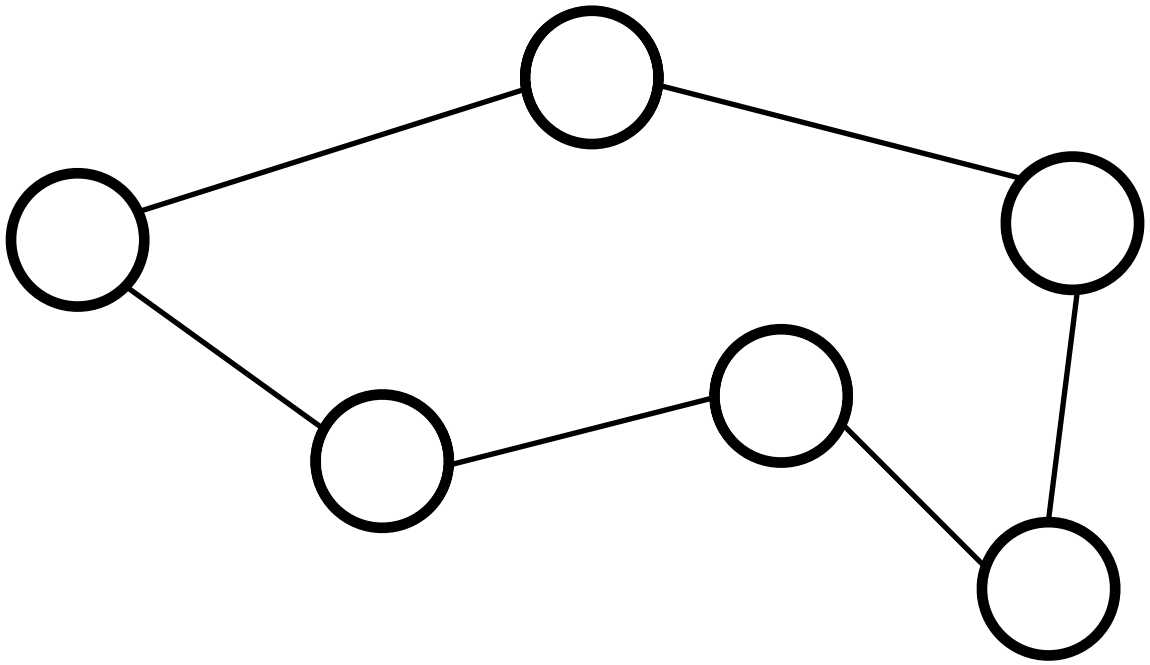
$$y_{ij} \le \mathbf{1}_{D_i + D_j \le 1} \qquad \forall (i,j) \in E$$

Two connected sites can be served by a single truck if their total demand does not exceed the truck capacity

- The expected loss is

Total fulfilled demand under complete graph

$$\mathbb{E}_D[L(G,D)] = 2\mathbb{E}_D\left[\mu(K_n, D) - \mu(G, D)\right]$$



# Transportation Flexibility Model

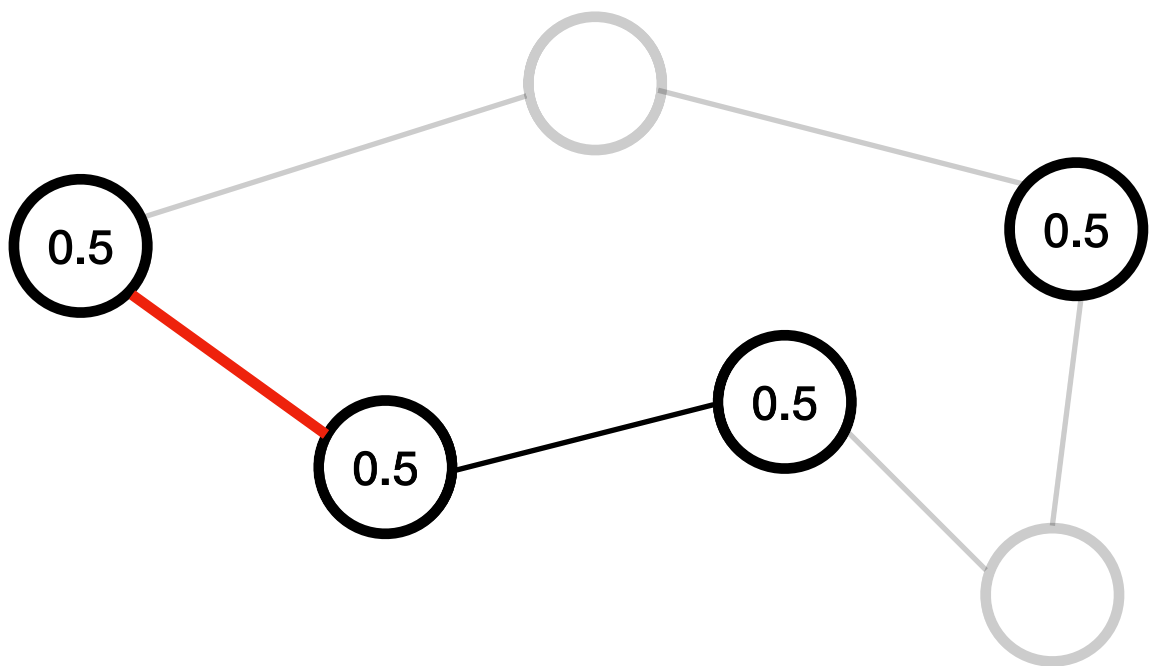
- Given a graph design $G = ([n], E)$
  - Vertices = sites
  - Edges = delivery routes
- The total fulfilled demand

$$\mu(G, D) = \max_{y_{ij} \in \{0,1\}} \sum_{(i,j) \in E} y_{ij}$$

$$\text{s.t.} \quad \sum_{j:\,(i,j) \in E} y_{ij} \leq 1 \qquad \forall i$$

$$y_{ij} \leq \mathbf{1}_{D_i + D_j \leq 1} \qquad \forall (i,j) \in E \longleftarrow$$

Two connected sites can be served by a single truck if their total demand does not exceed the truck capacity




- The expected loss is

$$\mathbb{E}_D[L(G, D)] = 2\mathbb{E}_D\left[\mu(K_n, D) - \mu(G, D)\right] \longleftarrow$$

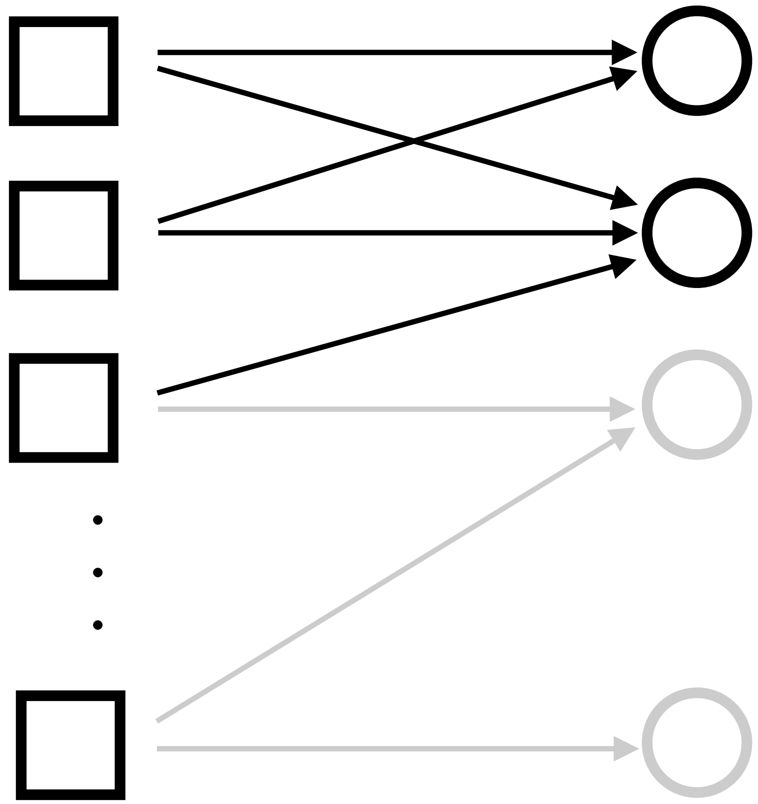
Expected number of unmatched sites

- IID Bernoulli demand with mean 1:

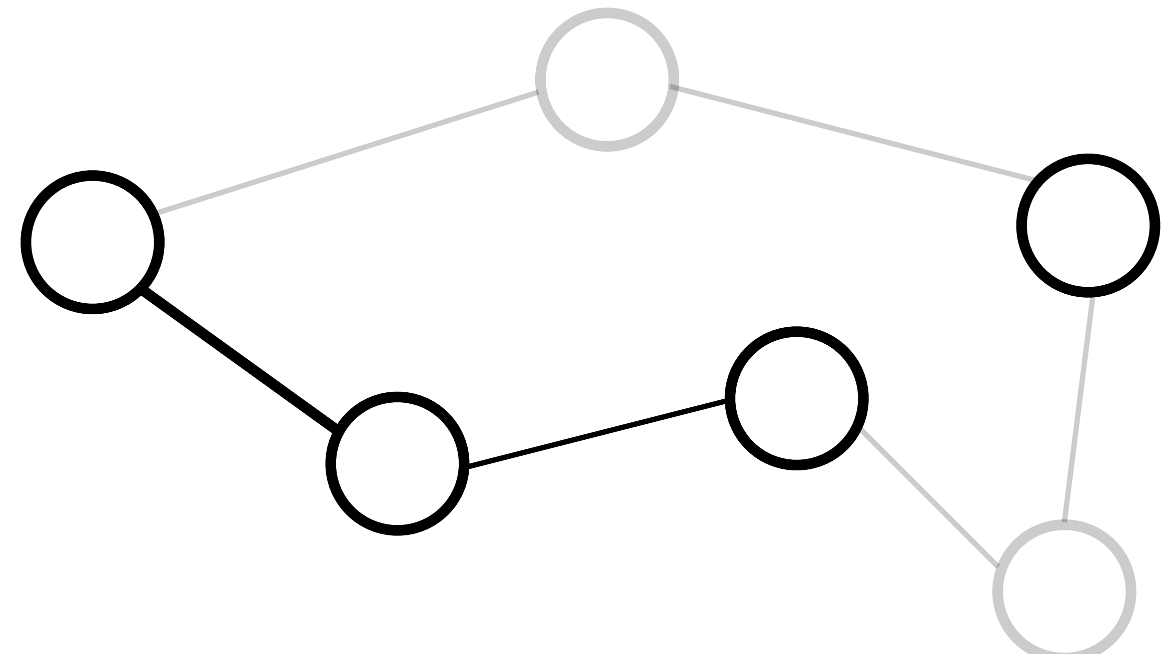
$$D_j = \begin{cases} 1 & \text{w.p. } q \\ 0.5 & \text{w.p. } p = 1 - q \end{cases} \longleftarrow$$

Random node deletion with probability $q$

# Stochastic matching with random node deletion



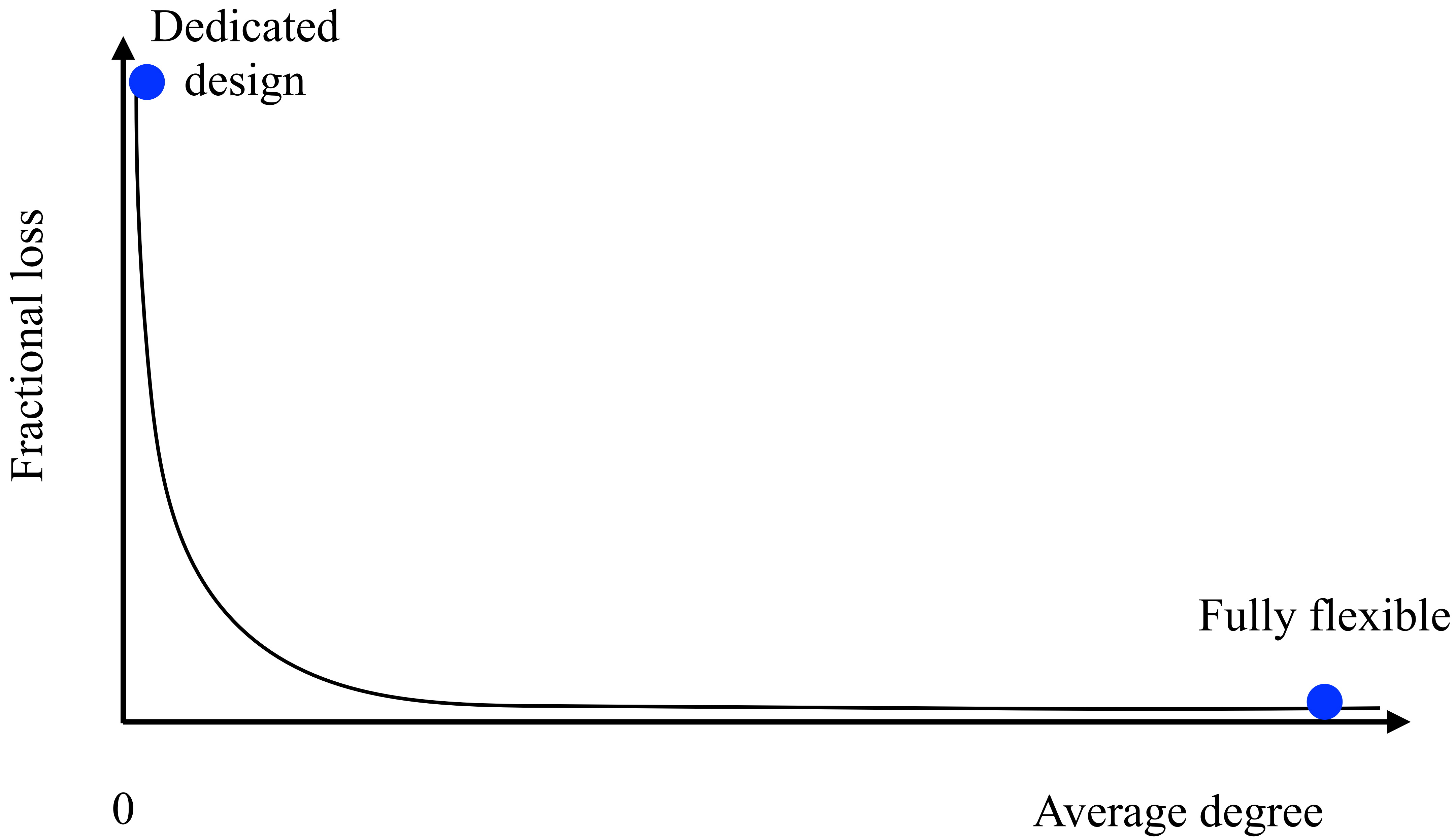
Process flexibility



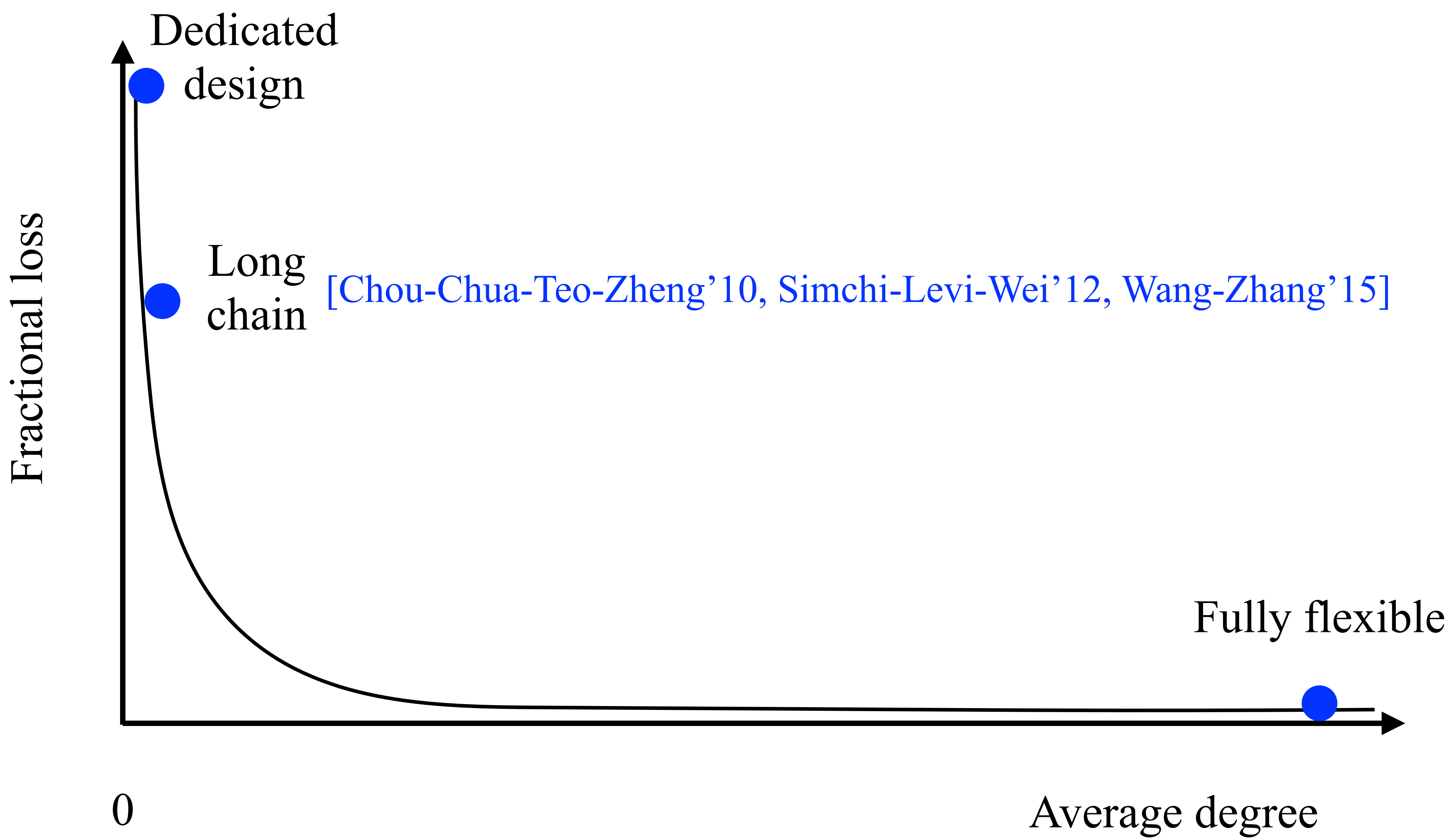
Transportation flexibility

- Both problems reduce to a “matching” problem in the residual graph $G_{\mathrm{res}}$ obtained after each demand node is independently deleted with probability $q$
- Goal: Design a graph that sustains a large “matching” after random node deletions

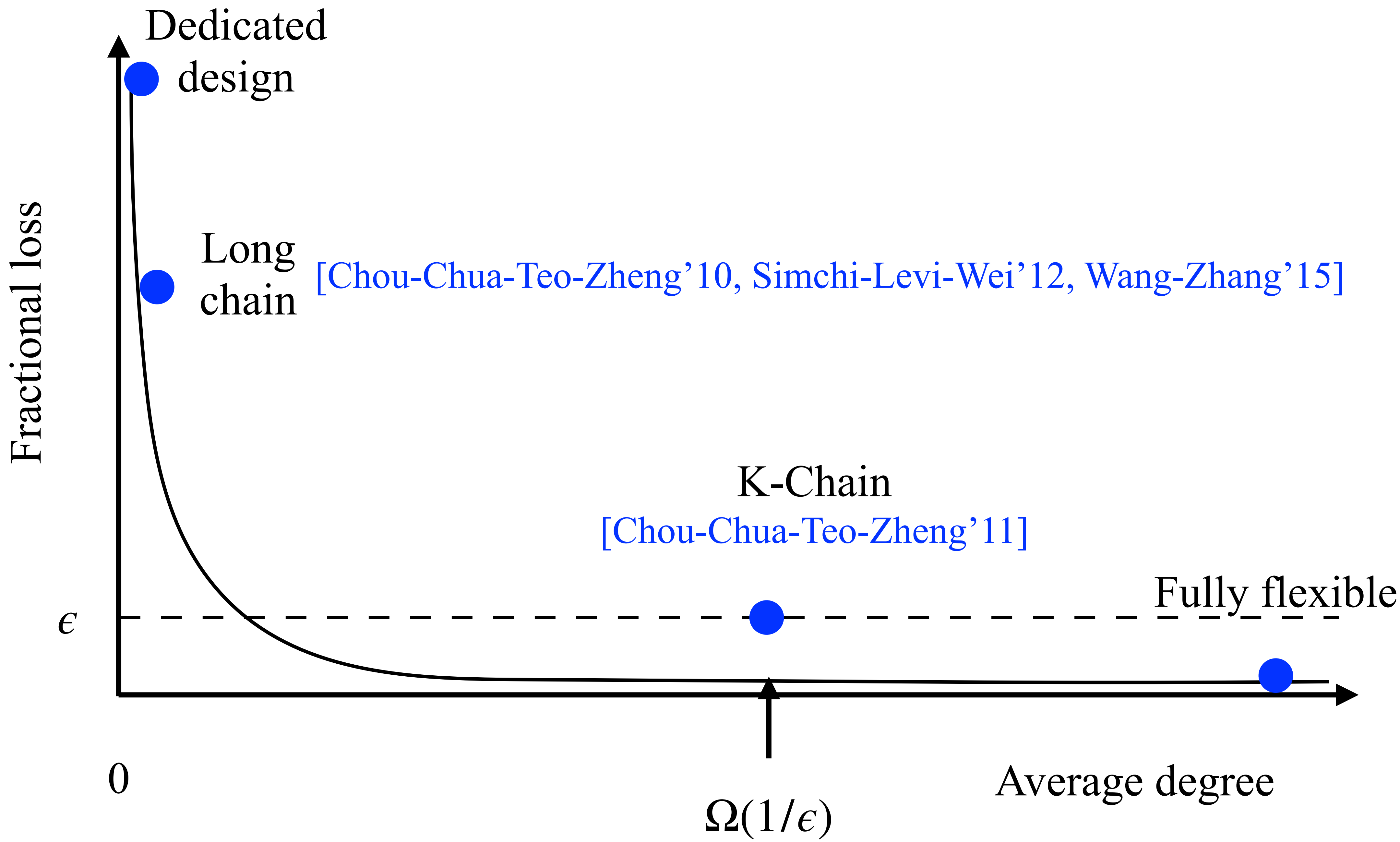
# Tradeoff: Demand Loss vs Network Complexity



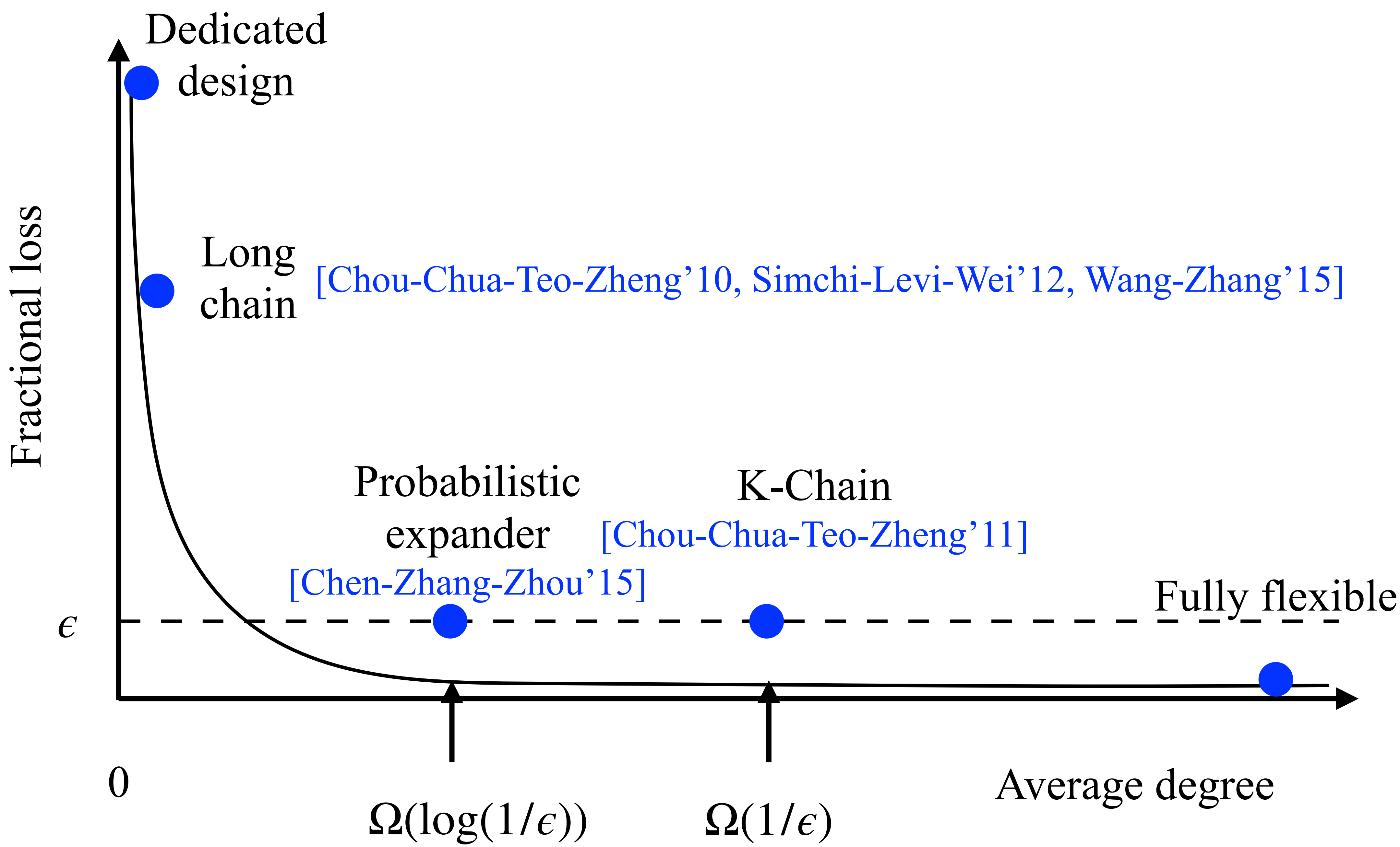
# Tradeoff: Demand Loss vs Network Complexity



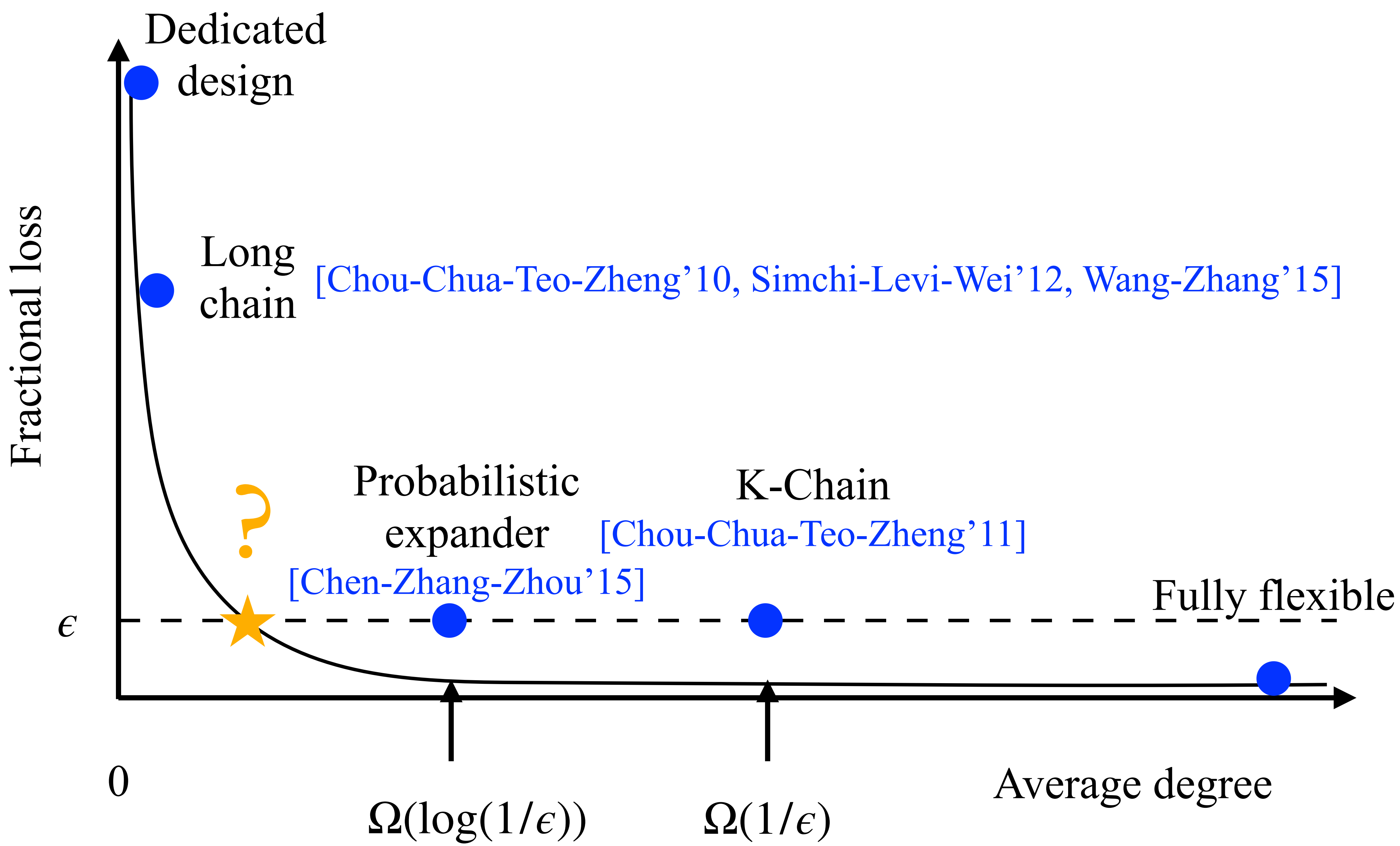
# Tradeoff: Demand Loss vs Network Complexity



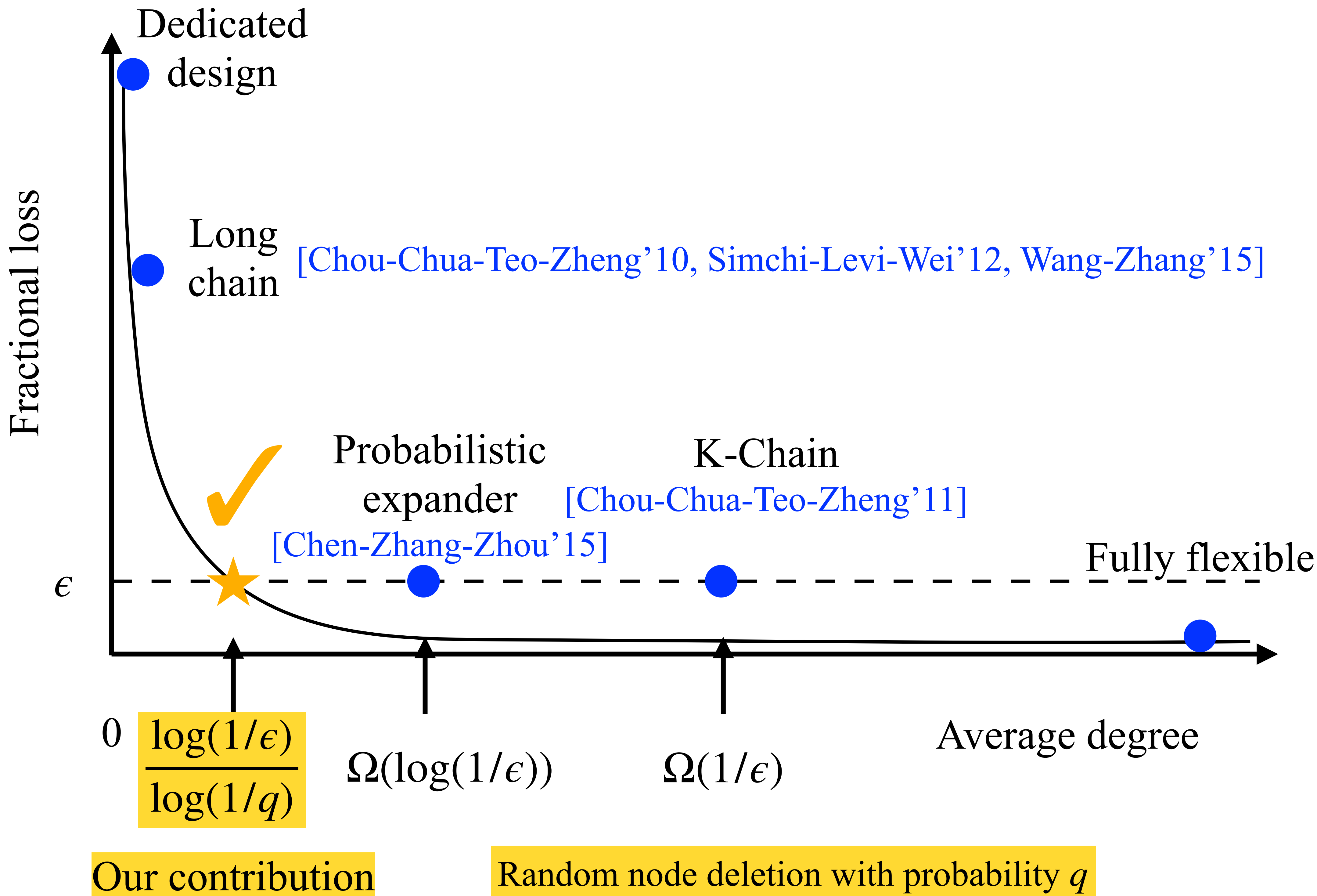
# Tradeoff: Demand Loss vs Network Complexity



# Tradeoff: Demand Loss vs Network Complexity



# Tradeoff: Demand Loss vs Network Complexity



# Lower Bound

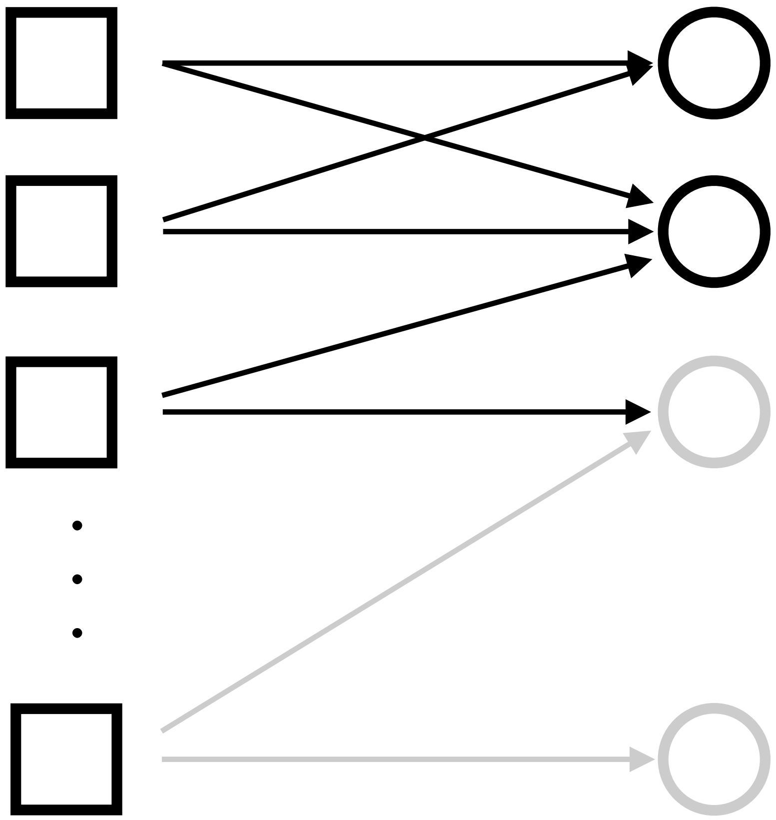
# Process Flexibility: A simple lower bound

**Theorem**

For any bipartite graph $G$ with an average degree $d$,

$$\mathbb{E}_D[L(G,D)] \gtrsim nq^d.$$

Expected demand loss



Demand node deletion w.p. $q$

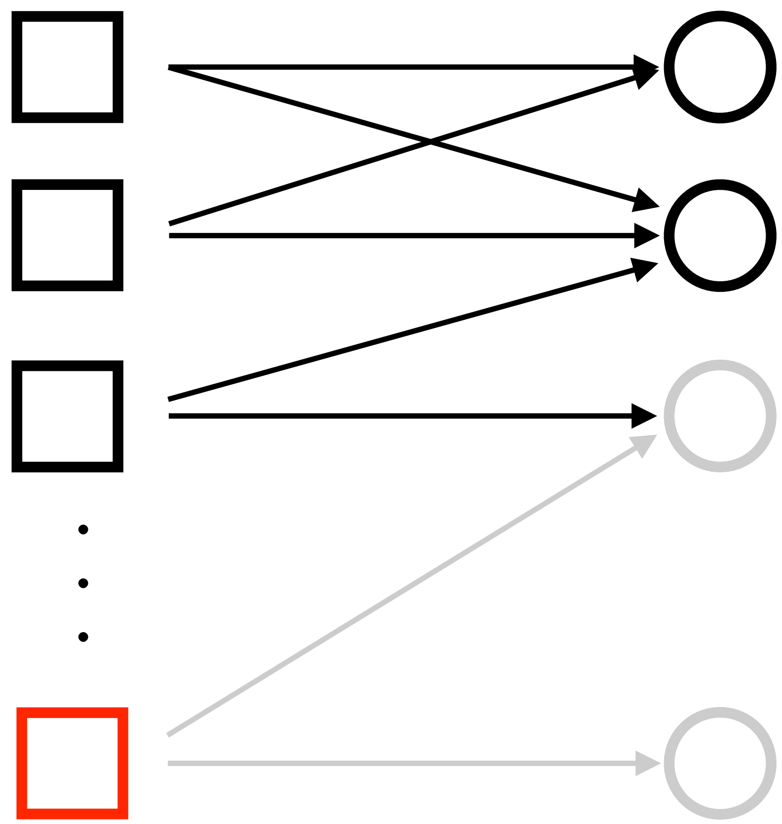
# Process Flexibility: A simple lower bound

**Theorem**

For any bipartite graph $G$ with an average degree $d$,

$$\mathbb{E}_D[L(G, D)] \gtrsim nq^d.$$

- Proof Ideas:
  - $L(G, D) \gtrsim$ # of isolated supply nodes in $G_{\text{res}}$



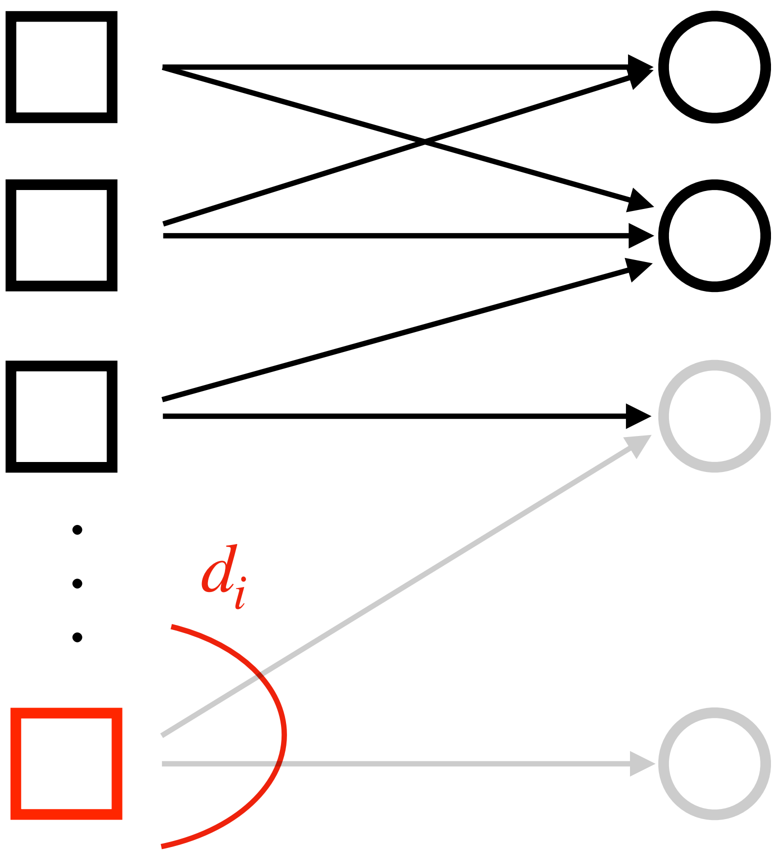
Demand node deletion w.p. $q$

# Process Flexibility: A simple lower bound

**Theorem**

For any bipartite graph $G$ with an average degree $d$,

$$\mathbb{E}_D[L(G,D)] \geq nq^d.$$




Demand node deletion w.p. $q$

- Proof Ideas:
  - $L(G,D) \gtrsim$ # of isolated supply nodes in $G_{\text{res}}$
  - $\mathbb{E}[\text{ \# of isolated supply nodes in } G_{\text{res}}] = \sum_{i=1}^{n} q^{d_i}$

# Process Flexibility: A simple lower bound

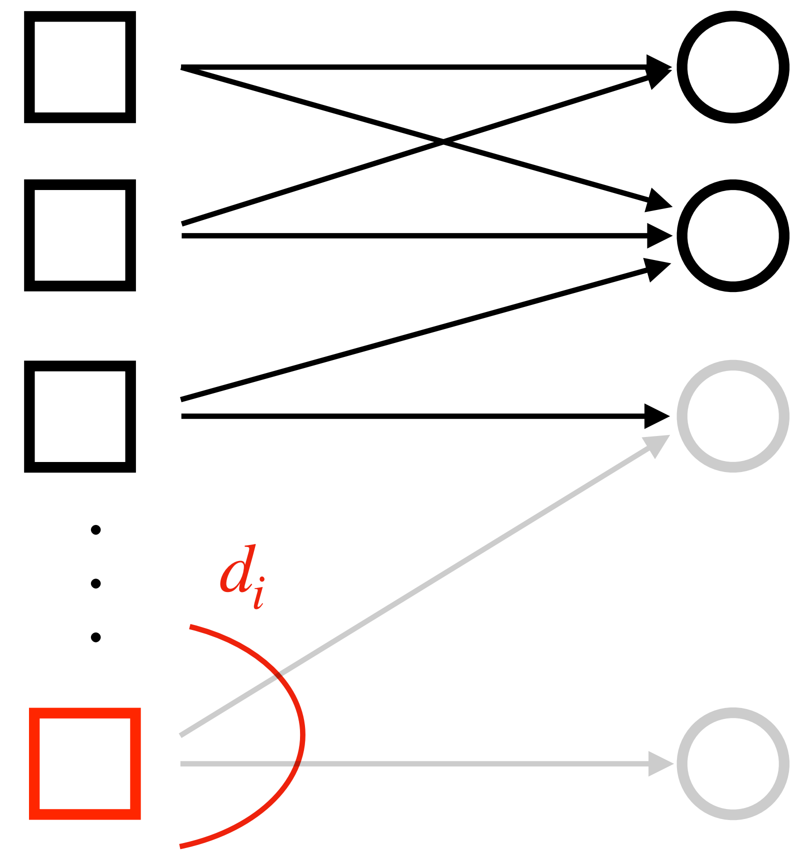
**Theorem**

For any bipartite graph $G$ with an average degree $d$,

$$\mathbb{E}_D[L(G,D)] \gtrsim nq^d.$$

- Proof Ideas:
  - $L(G,D) \gtrsim$ # of isolated supply nodes in $G_{\text{res}}$
  - $\mathbb{E}[\text{ \# of isolated supply nodes in } G_{\text{res}}] = \sum_{i=1}^{n} q^{d_i} \geq nq^d$

Jensen's; Equality holds for regular graphs



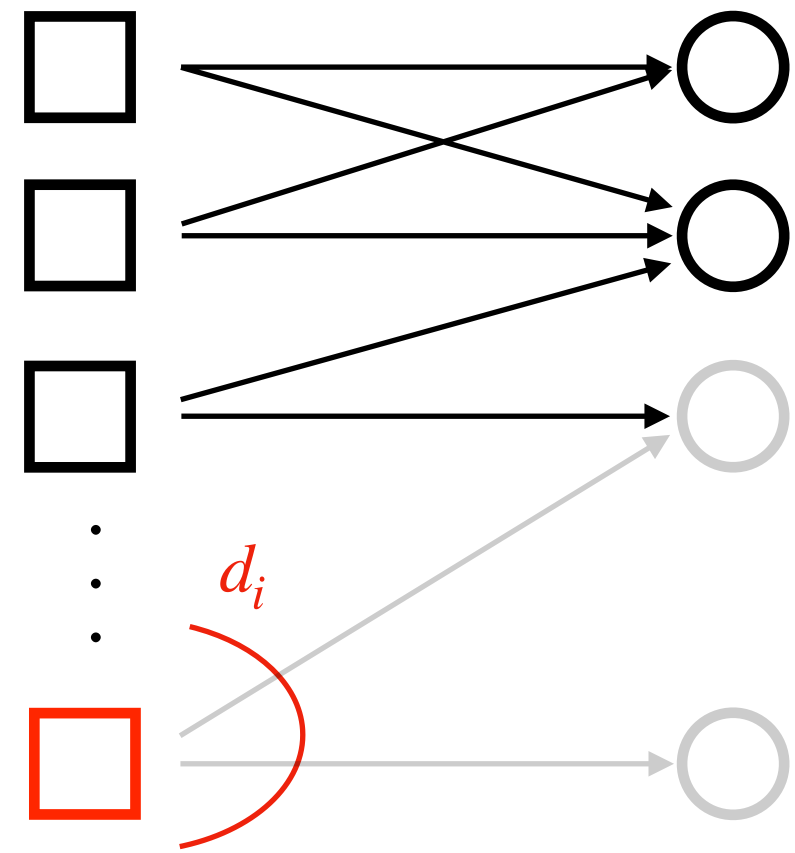

Demand node deletion w.p. $q$

# Process Flexibility: A simple lower bound

**Theorem**

For any bipartite graph $G$ with an average degree $d$,

$$\mathbb{E}_D[L(G,D)] \gtrsim nq^d.$$




- Proof Ideas:
  - $L(G,D) \gtrsim$ # of isolated supply nodes in $G_{\text{res}}$
  - $\mathbb{E}[\text{ \# of isolated supply nodes in } G_{\text{res}}] = \sum_{i=1}^{n} q^{d_i} \geq nq^d$

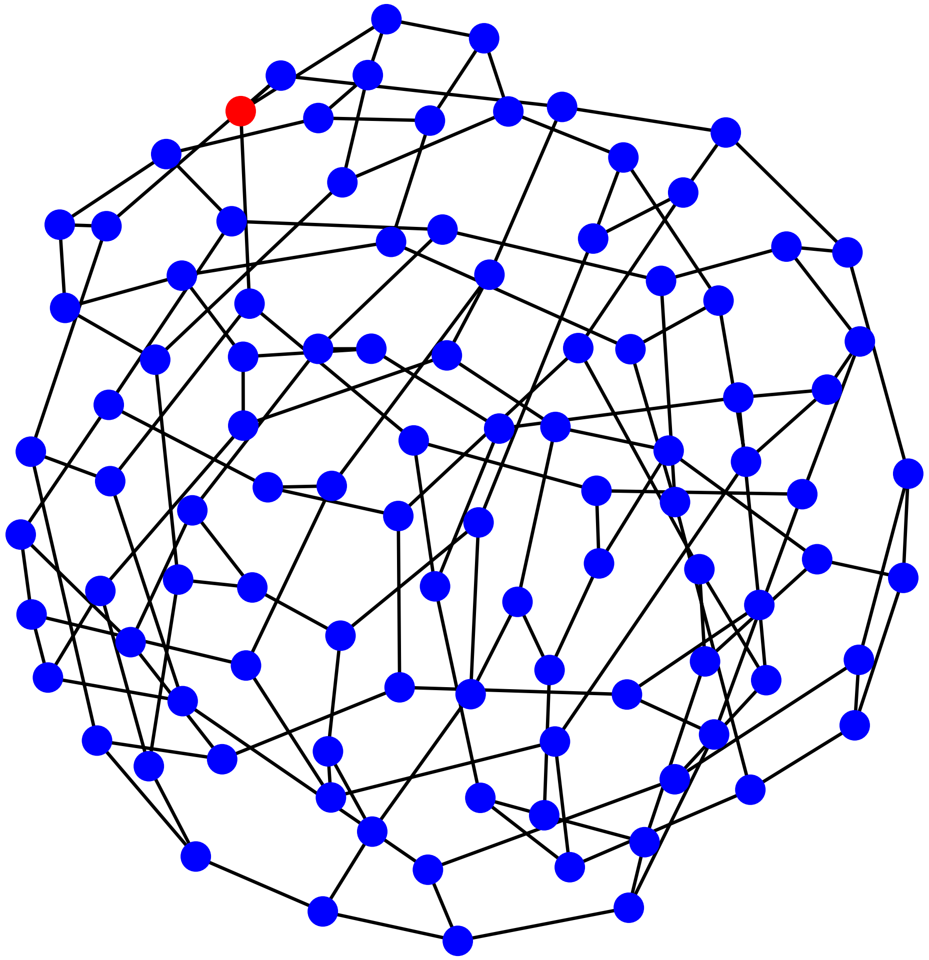
Jensen's; Equality holds for regular graphs

- Implications:
  - $\epsilon n$-loss requires $d \geq \log(1/\epsilon)/\log(1/q)$ for any graph;
  - An optimal graph design needs to be regular
  - K-Chain is also regular, but it still requires $d \geq \Omega(1/\epsilon)$ [Chou-Chua-Teo-Zheng'11]

# Our solution: Random regular graphs

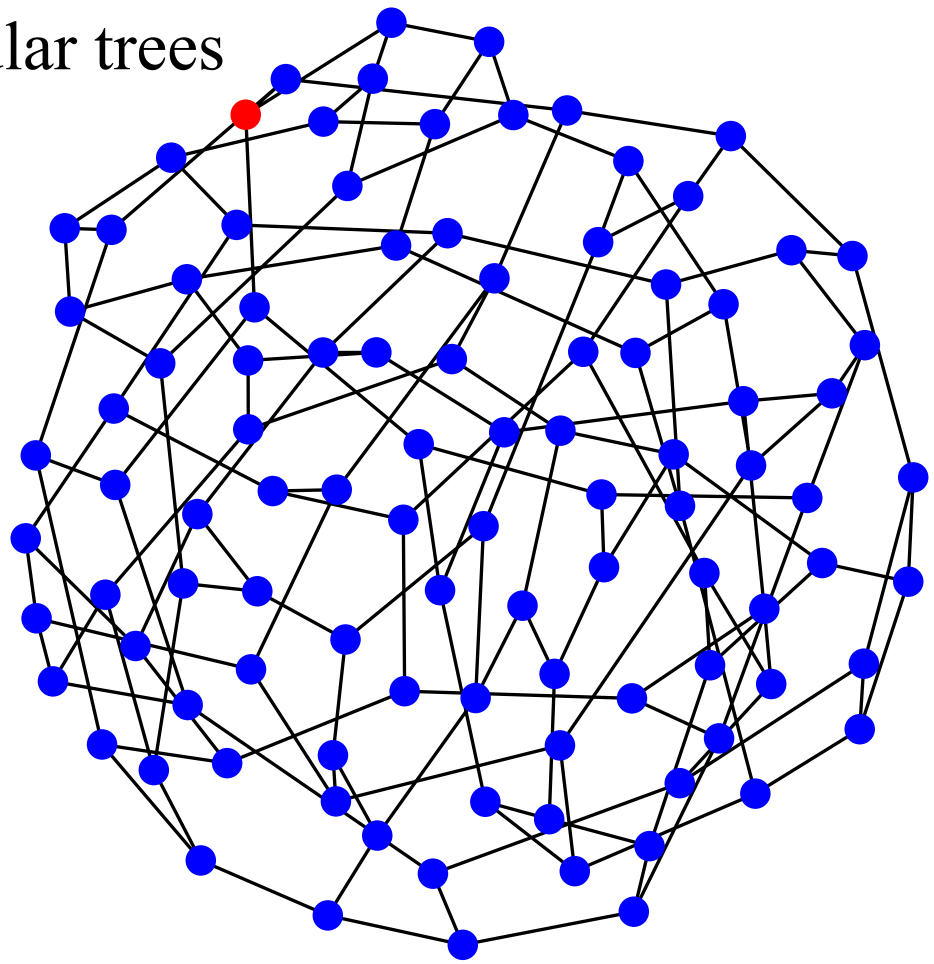
# Random Regular Graphs

- An object widely study in probability and TCS



# Random Regular Graphs

- An object widely study in probability and TCS
- Locally tree-like: local neighborhoods look like regular trees



# Random Regular Graphs

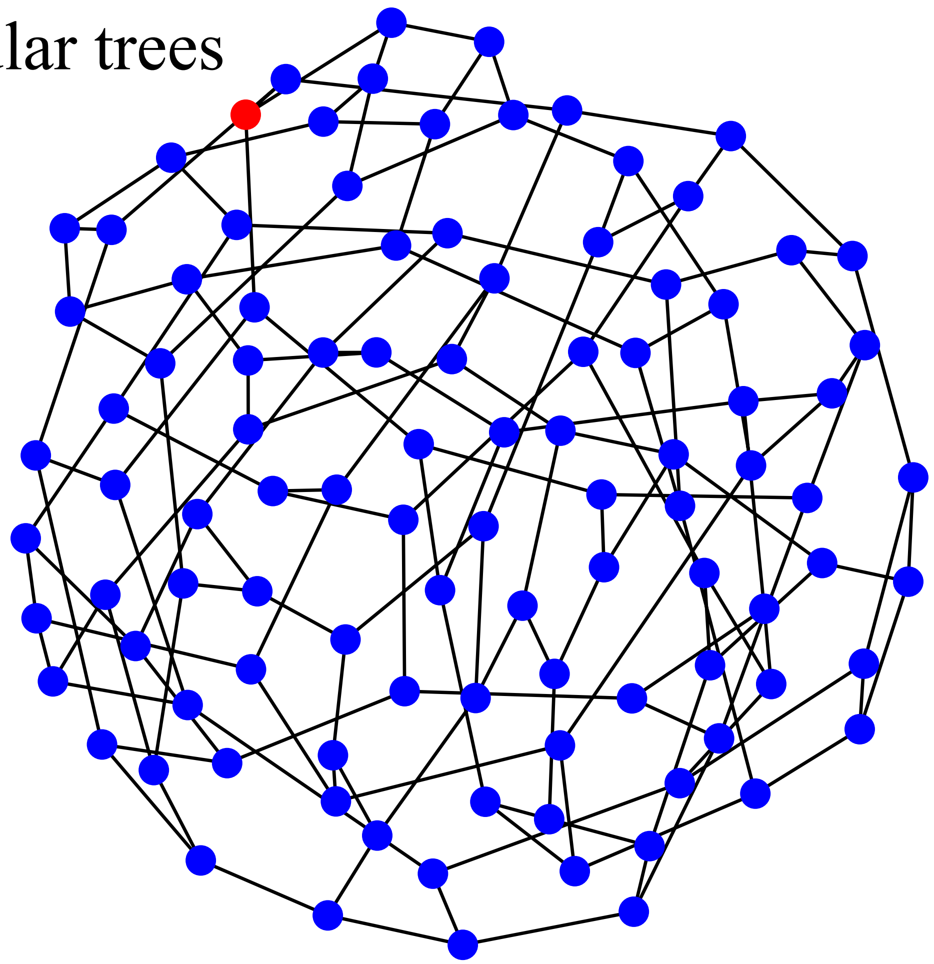
- An object widely study in probability and TCS
- Locally tree-like: local neighborhoods look like regular trees
- Spectral expansion: Almost Ramanujan graphs

**Freedman's Theorem (2003)**

For random $d$-regular graph with eigenvalues $\lambda_n \leq \ldots \leq \lambda_2 \leq \lambda_1 = d$ ,

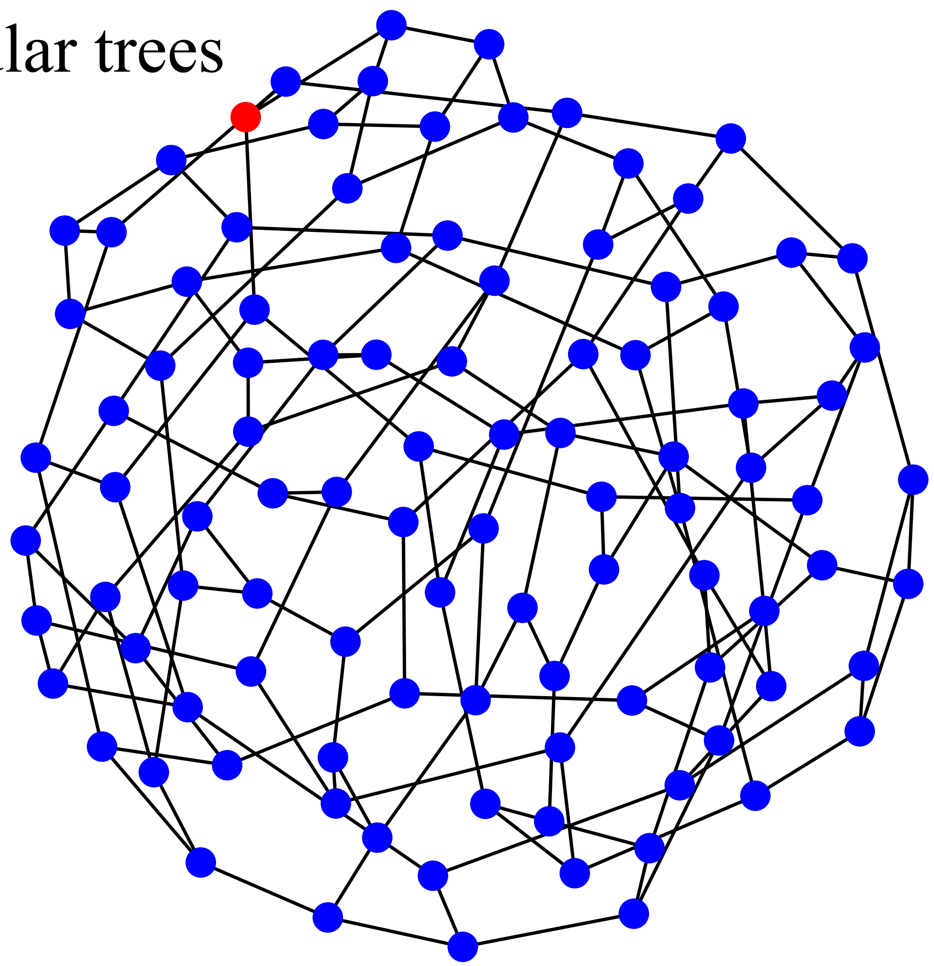
$$\max\{\lambda_2, |\lambda_n|\} \leq 2\sqrt{d-1} + o(1)$$

Optimal among all $d$-regular graphs



# Random Regular Graphs

- An object widely study in probability and TCS
- Locally tree-like: local neighborhoods look like regular trees
- Spectral expansion: Almost Ramanujan graphs



**Freedman's Theorem (2003)**

For random $d$-regular graph with eigenvalues $\lambda_n \leq \ldots \leq \lambda_2 \leq \lambda_1 = d$ ,

$$\max\{\lambda_2, |\lambda_n|\} \leq 2\sqrt{d-1} + o(1)$$

Optimal among all $d$-regular graphs

**Question**: Do random regular graphs enjoy optimality in flexibility design?

# Process Flexibility: Random Regular Graphs Are Optimal

**Theorem** [Li-Niu-Wei-Xu '25]

For any constant $\delta > 0$, a random $d$-regular bipartite graph $G$ achieves whp:

- $\epsilon$-fractional loss ($\mathbb{E}_D[L(G, D)] \leq \epsilon n$), if

$$d \geq \frac{(1+\delta)\log(1/\epsilon)}{\log(1/q)}$$

- Constant loss ($\mathbb{E}_D[L(G, D)] = O(1)$), if

$$d \geq \frac{(1+\delta)\log n}{\log(1/q)}$$

Remarks

- Match the lower bounds with the sharp constant;
- Previous best-known upper bound $d \geq \Omega(\log(1/\epsilon))$ [Chen-Zhang-Zhou'15];
- $K$-chain needs $d \geq \Omega(1/\epsilon)$ [Chou-Chua-Teo-Zheng'11].

# Transportation Flexibility

**Theorem** [Li-Niu-Wei-Xu '25]

For any constant $\delta > 0$, a random $d$-regular graph $G$ achieves whp:

- $\epsilon$-fractional loss ($\mathbb{E}_D[L(G, D)] \leq \epsilon n$), if

$$d \geq \frac{(1+\delta)\log(1/\epsilon)}{\log(1/q)}$$

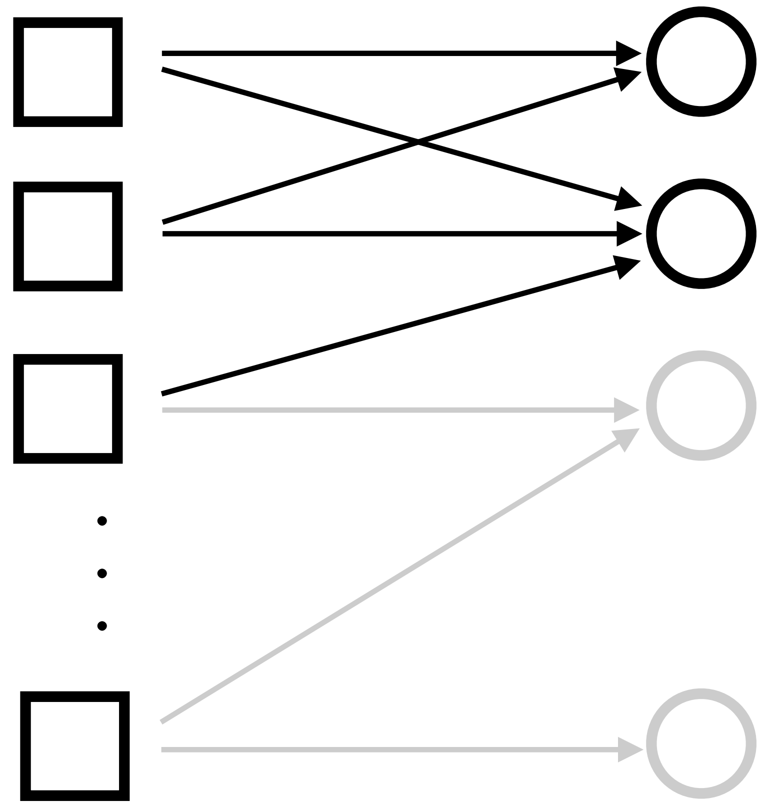
- Constant loss ($\mathbb{E}_D[L(G, D)] = O(1)$), if

$$d \geq \frac{(1+\delta)\log n}{\log(1/q)}$$

Remarks

- Match the lower bounds with the sharp constant;
- Erdős–Rényi random graph needs $d > \log(1/\epsilon)/(1-q)$ [Li-Niu-Wei-Xu'25];
- $K$-chain needs $d \geq 2\log(1/\epsilon)/\log(1/q)$ [Li-Niu-Wei-Xu'25].

# Proof ideas for process flexibility

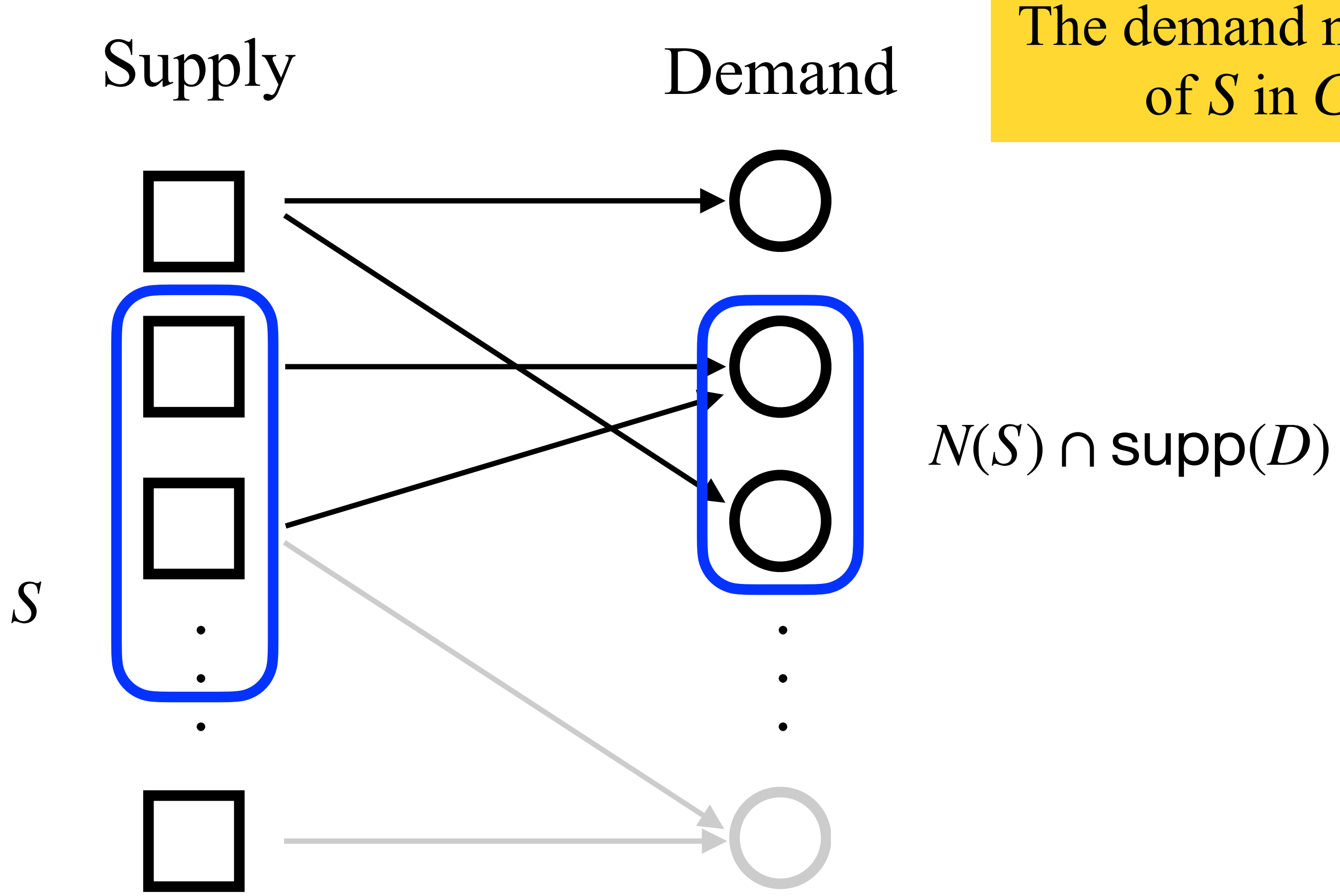


# Strong Connectivity Property

By max-flow min-cut theorem:

$$Z(G, D) = \min_{S \subset [n]} \{n - |S| + |N(S) \cap \mathrm{supp}(D)|/p\}$$

The demand neighbors of $S$ in $G_{\mathrm{res}}$

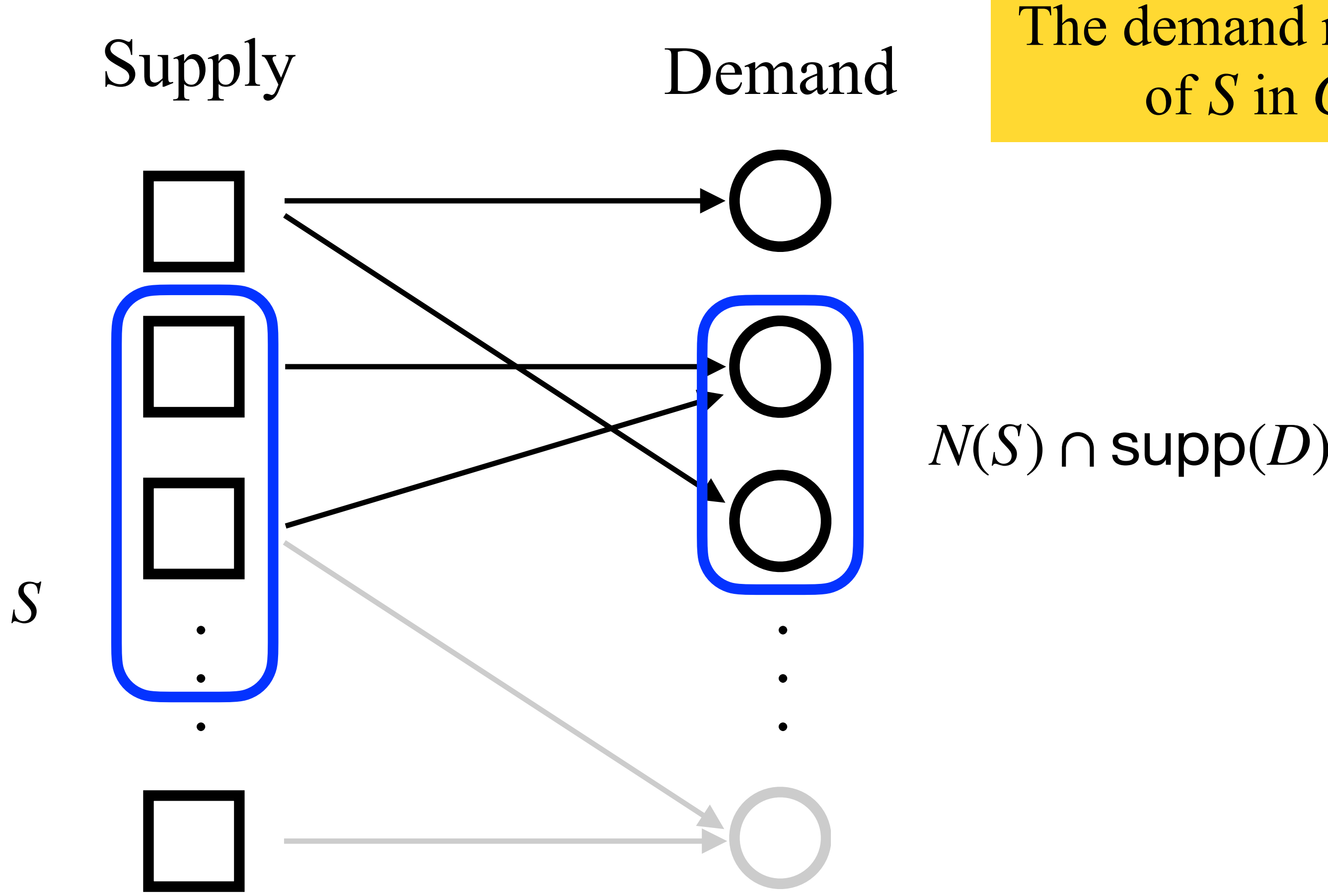


# Strong Connectivity Property

By max-flow min-cut theorem:

$$Z(G, D) = \min_{S \subset [n]} \{n - |S| + |N(S) \cap \mathrm{supp}(D)|/p\}$$

The demand neighbors of $S$ in $G_{\mathrm{res}}$




To ensure $\epsilon n$-loss:

Strong "$\epsilon n$-connectivity"

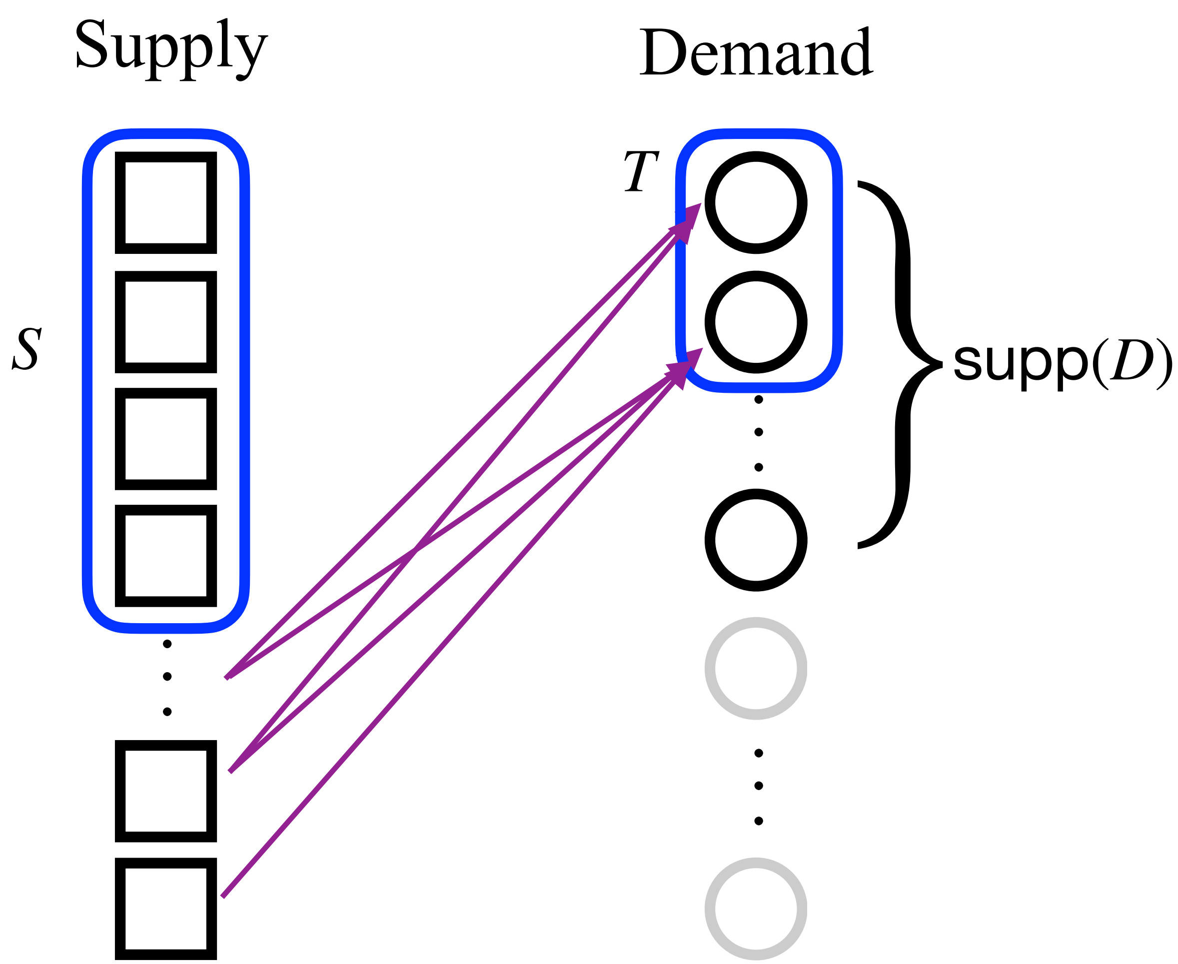
$$L(G, D) \geq (1 - \epsilon)n \iff \forall S \subset [n] : \ |N(S) \cap \mathrm{supp}(D)| \geq p(|S| - \epsilon n)$$

# Ruling out Large Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \text{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "large", if $|S| \geq \epsilon^c n$ for some choice of $c \in (0,1)$

# Ruling out Large Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \mathrm{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "large", if $|S| \geq \epsilon^c n$ for some choice of $c \in (0,1)$

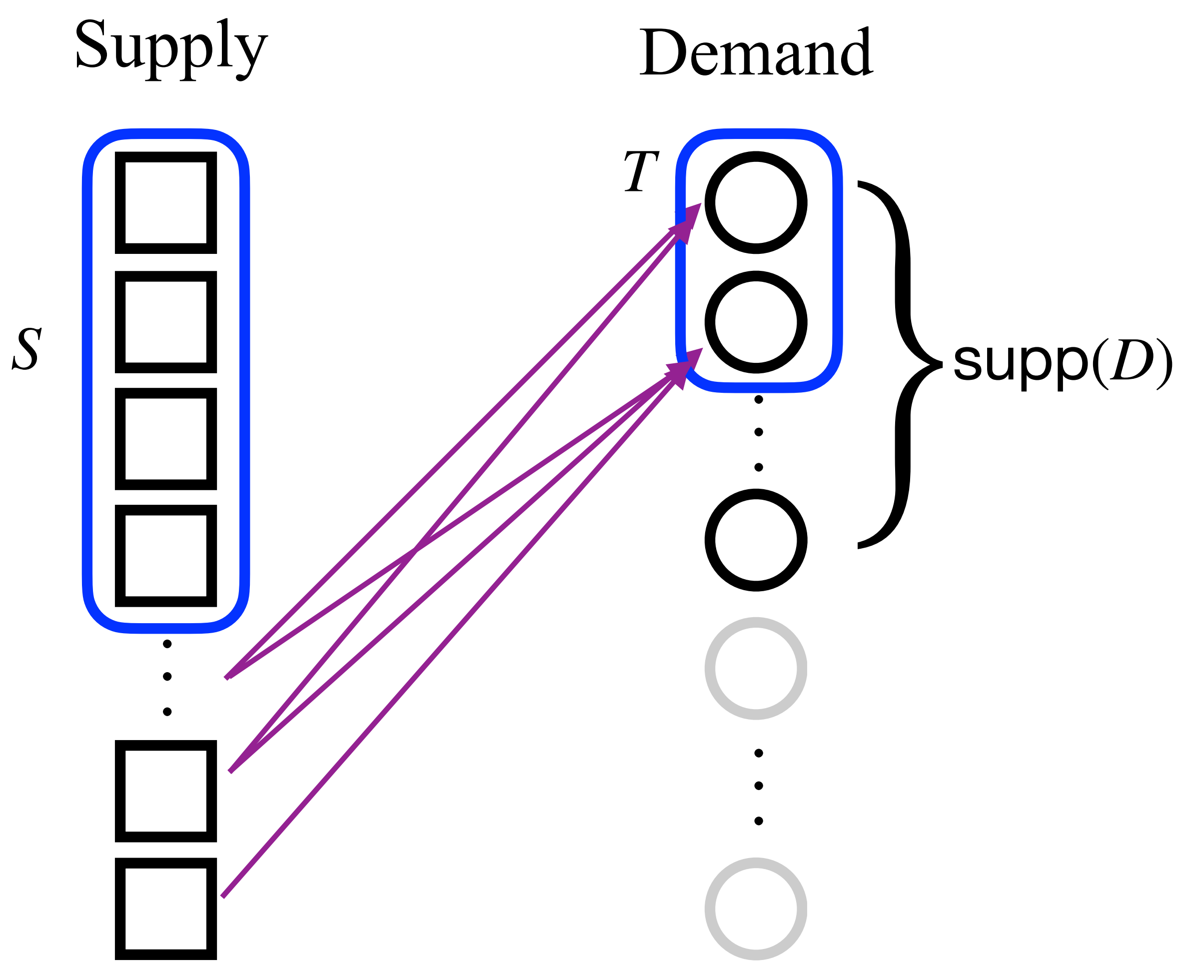



Large Bottleneck $S$

- $S$ is a "large" bottleneck $\Longrightarrow$
- $\exists$ demand subset $T \subset \mathrm{supp}(D)$ with $|T| = |\mathrm{supp}(D)| - p(|S| - \epsilon n)$ s.t. there is no cross-edge between $S$ and $T$, i.e., $e(S, T) = 0$

# Ruling out Large Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \text{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "large", if $|S| \geq \epsilon^c n$ for some choice of $c \in (0,1)$

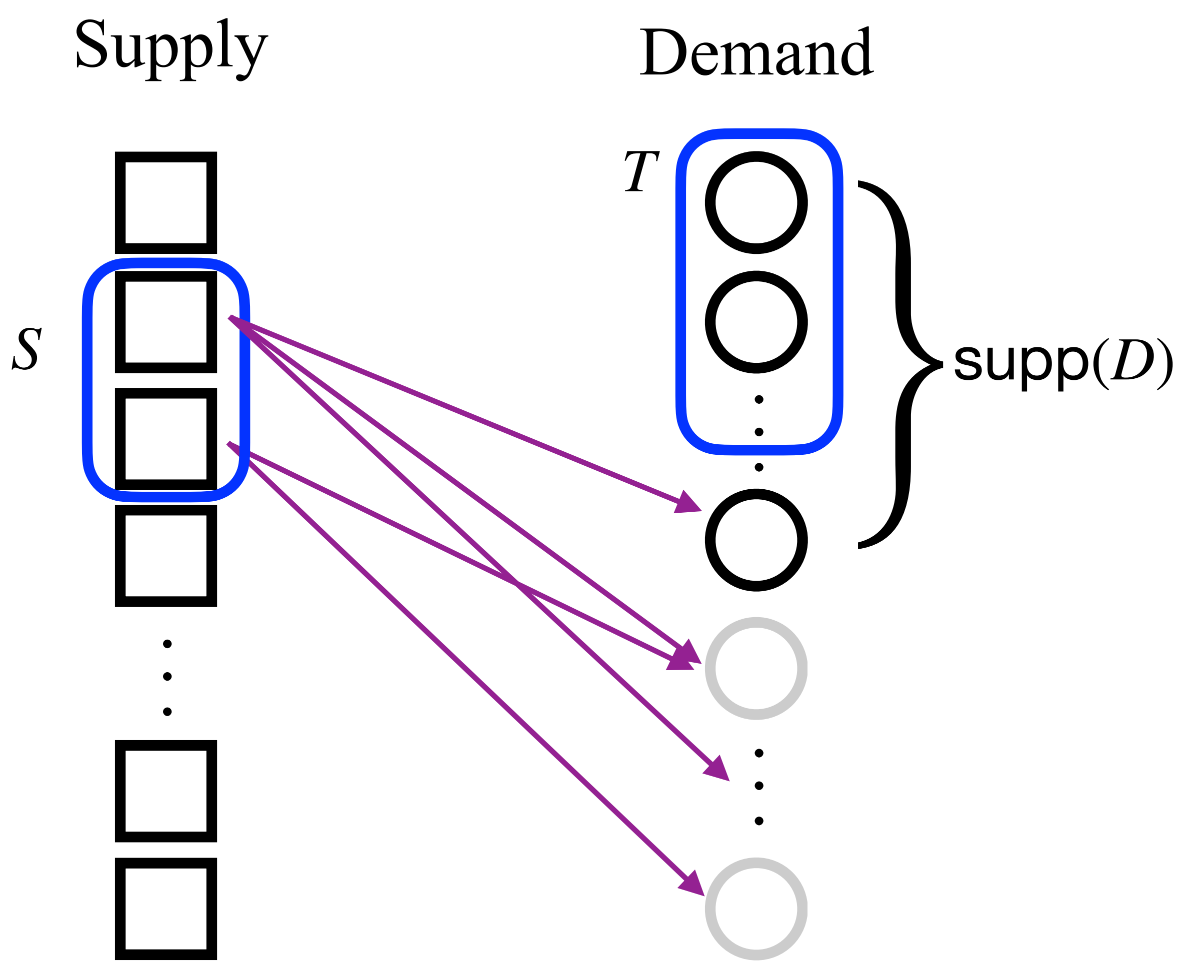



Large Bottleneck $S$

- $S$ is a "large" bottleneck $\Longrightarrow$
- $\exists$ demand subset $T \subset \text{supp}(D)$ with $|T| = |\text{supp}(D)| - p(|S| - \epsilon n)$ s.t. there is no cross-edge between $S$ and $T$, i.e., $e(S, T) = 0$
- A simple union bound over $S$ and $T$ shows that there is no such big cut whp
- Regularity on the demand side and random connections are important

# Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \mathrm{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "small", if $\epsilon n \leq |S| \leq \epsilon^c n$



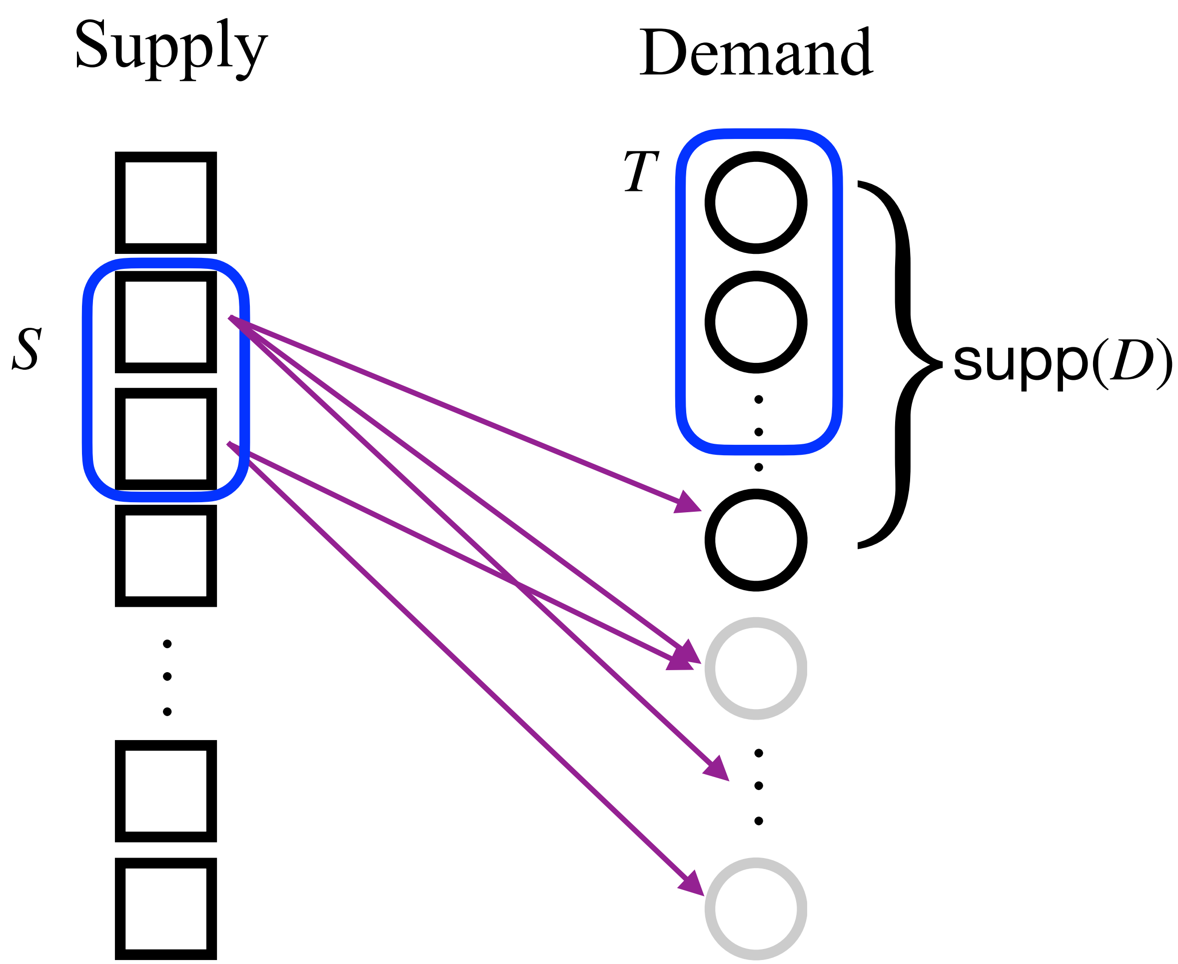

Small Bottleneck $S$

Key challenge:

$\{e(S, T)\}$ are highly correlated across different $T \subset \mathrm{supp}(D)$, so a simple union bound over $S$ and $T$ does not work

# Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \mathrm{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "small", if $\epsilon n \leq |S| \leq \epsilon^c n$

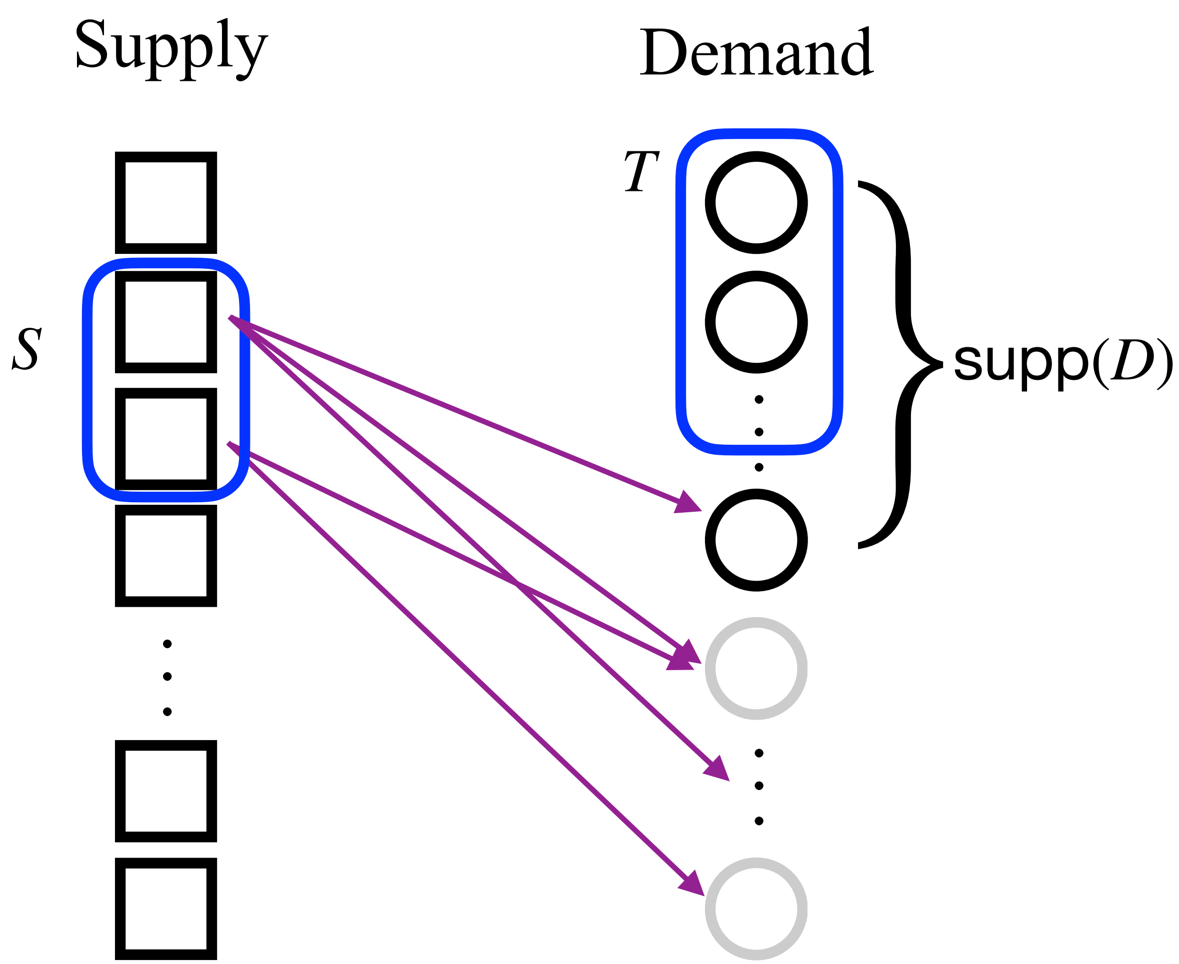



Small Bottleneck $S$
Rule out by (1)

(1) *Little waste of supply*:
$e(S, \mathrm{supp}(D)) > \gamma |S|$ for all $S$ with $|S| \geq \epsilon n$, for some $\gamma > 0$

# Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \mathrm{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "small", if $\epsilon n \leq |S| \leq \epsilon^c n$




Small Bottleneck $S$

Rule out by (1)

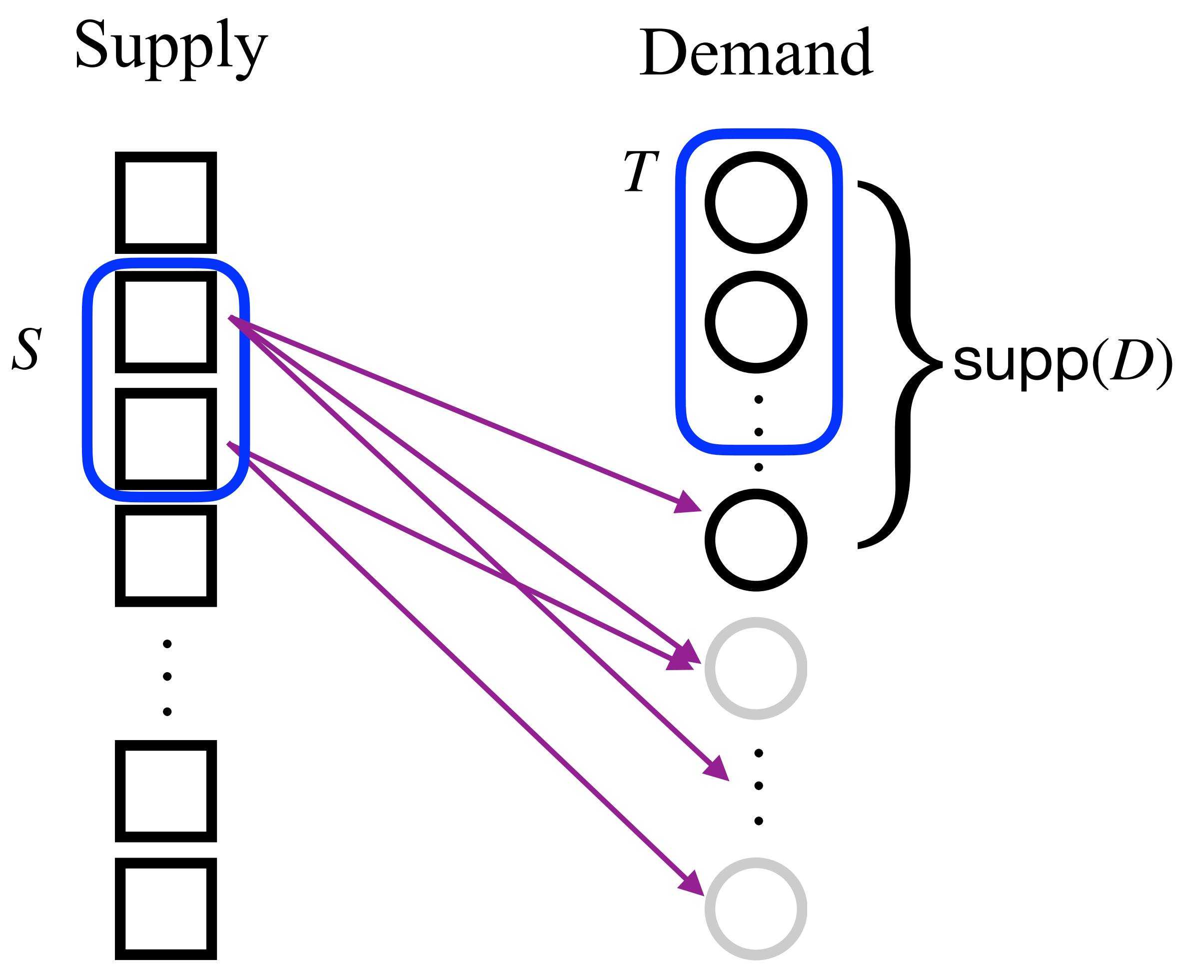
(1) *Little waste of supply*:

$e(S, \mathrm{supp}(D)) > \gamma |S|$ for all $S$ with $|S| \geq \epsilon n$, for some $\gamma > 0$

A stronger requirement than the bound on the # isolated supplier nodes

# Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \text{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "small", if $\epsilon n \le |S| \le \epsilon^c n$




Small Bottleneck $S$

Rule out by (1)

(1) *Little waste of supply*:

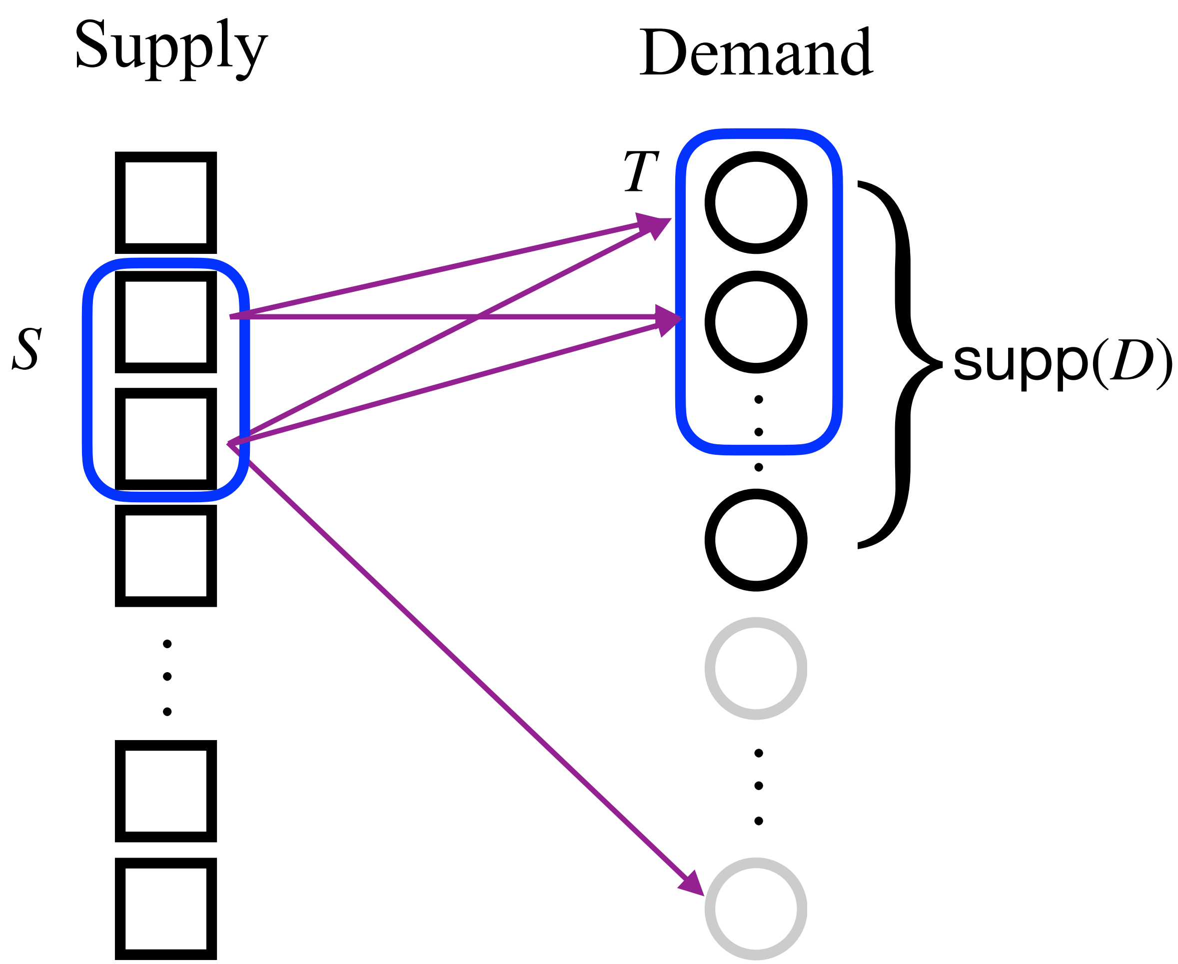
$e(S, \text{supp(D)}) > \gamma |S|$ for all $S$ with $|S| \ge \epsilon n$, for some $\gamma > 0$

A stronger requirement than the bound on the # isolated supplier nodes

Satisfied when $d \ge \dfrac{(1+\delta)\log(1/\epsilon)}{\log(1/q)}$

# Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \mathrm{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "small", if $\epsilon n \leq |S| \leq \epsilon^c n$




Small Bottleneck $S$

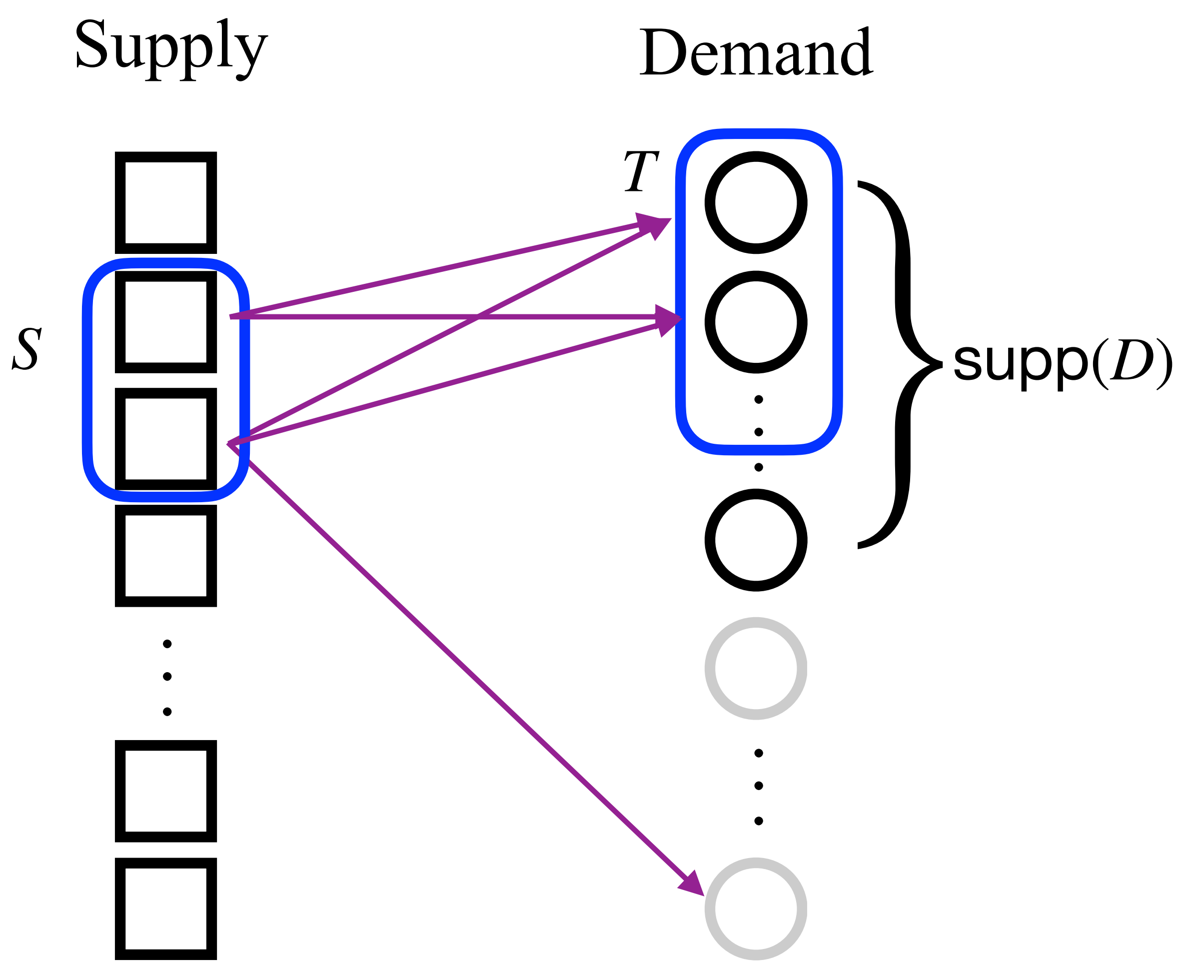
(1) *Little waste of supply*:
$e(S, \mathrm{supp(D)}) > \gamma |S|$ for all $S$ with $|S| \geq \epsilon n$, for some $\gamma > 0$

A stronger requirement than the bound on the # isolated supplier nodes

Satisfied when $d \geq \dfrac{(1+\delta)\log(1/\epsilon)}{\log(1/q)}$

# Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \mathrm{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "small", if $\epsilon n \leq |S| \leq \epsilon^c n$




Small Bottleneck $S$

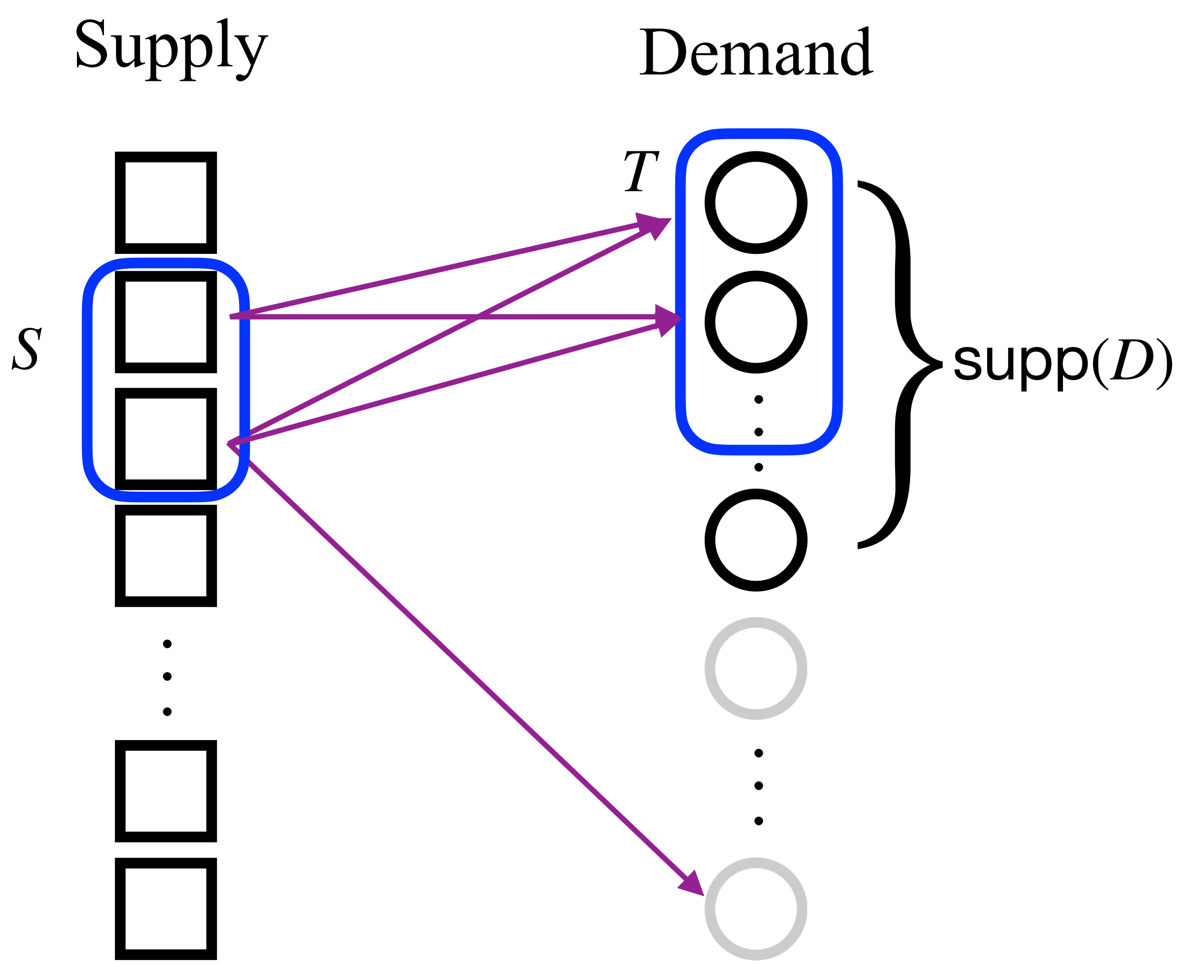
Rule out by (2)

(1) *Little waste of supply*:
$e(S, \mathrm{supp(D)}) > \gamma |S|$ for all $S$ with $|S| \geq \epsilon n$, for some $\gamma > 0$

(2) *Little congestion*: $e(S, T) < \gamma |S|$
for all $S$ with $\epsilon n \leq |S| \leq \epsilon^c n$ and all $T \subset \mathrm{supp}(D)$ with $|T| \leq p(|S| - \epsilon n)$

# Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $|N(S) \cap \mathrm{supp}(D)| < p(|S| - \epsilon n)$
- A bottleneck $S$ is called "small", if $\epsilon n \leq |S| \leq \epsilon^c n$




Small Bottleneck $S$

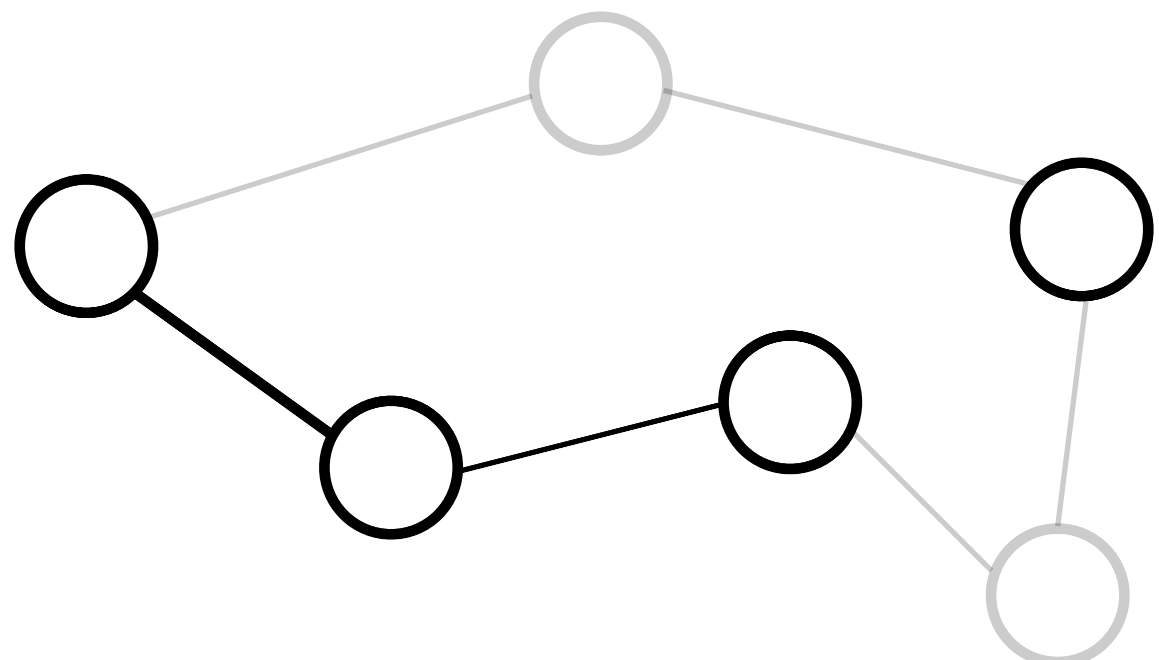
Rule out by (2)

(1) *Little waste of supply*: $e(S, \mathrm{supp(D)}) > \gamma |S|$ for all $S$ with $|S| \geq \epsilon n$, for some $\gamma > 0$

(2) *Little congestion*: $e(S, T) < \gamma |S|$ for all $S$ with $\epsilon n \leq |S| \leq \epsilon^c n$ and all $T \subset \mathrm{supp}(D)$ with $|T| \leq p(|S| - \epsilon n)$

- (1) + (2) $\Rightarrow$ small bottleneck $S$ cannot occur
- Regularity on the supply side and random connections are important

# Proof ideas for transportation flexibility



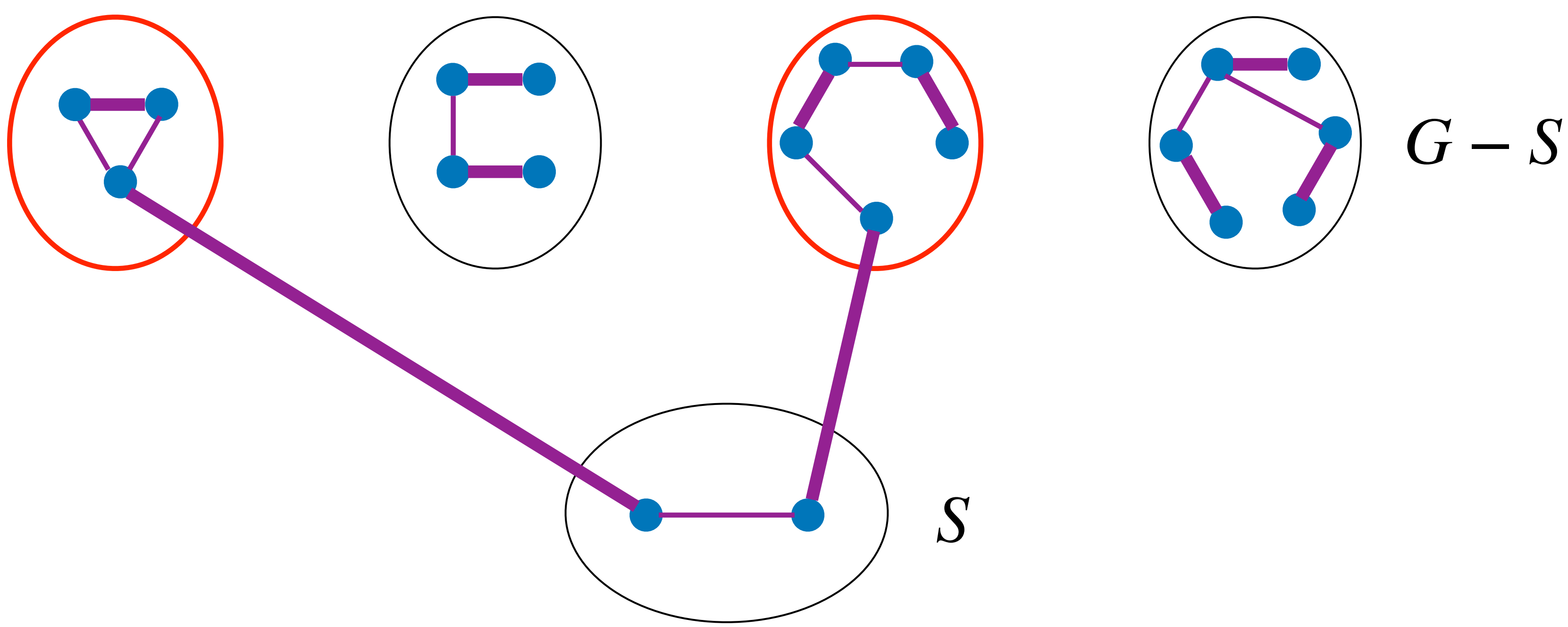
# Tutte-Berge theorem

**Theorem** [Tutte '47, Berge'58]

The number of unmatched vertices in a maximum matching of graph $G$ is given by

$$\max_{S \subseteq V(G)} \left( \mathrm{odd}(G - S) - |S| \right)$$

# of odd components when $S$ removed



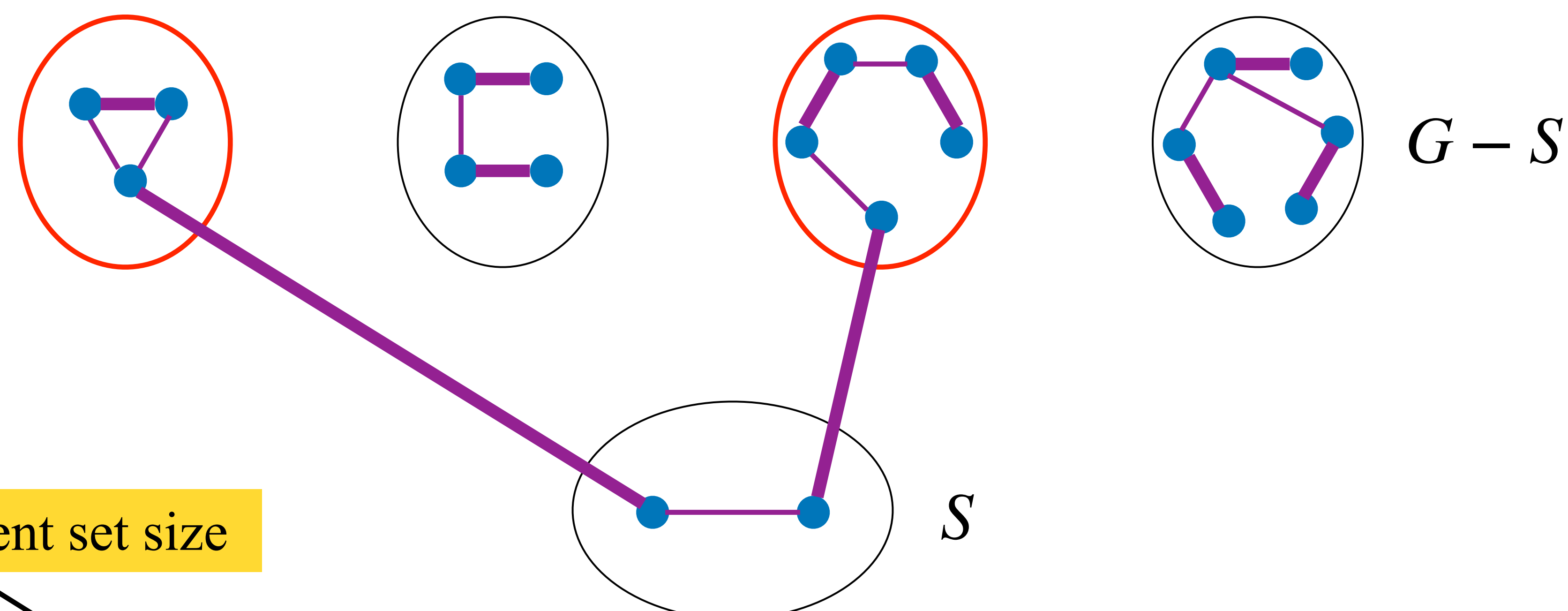
# Tutte-Berge theorem

**Theorem** [Tutte '47, Berge'58]

The number of unmatched vertices in a maximum matching of graph $G$ is given by

$$\max_{S \subseteq V(G)} \Big( \mathrm{odd}(G - S) - |S| \Big)$$

\# of odd components when $S$ removed




Maximum independent set size

Remark: $\mathrm{odd}(G - S) \leq \mathrm{Ind}(G) \Rightarrow$ It suffices to check $S$ with $|S| \leq \mathrm{Ind}(G)$

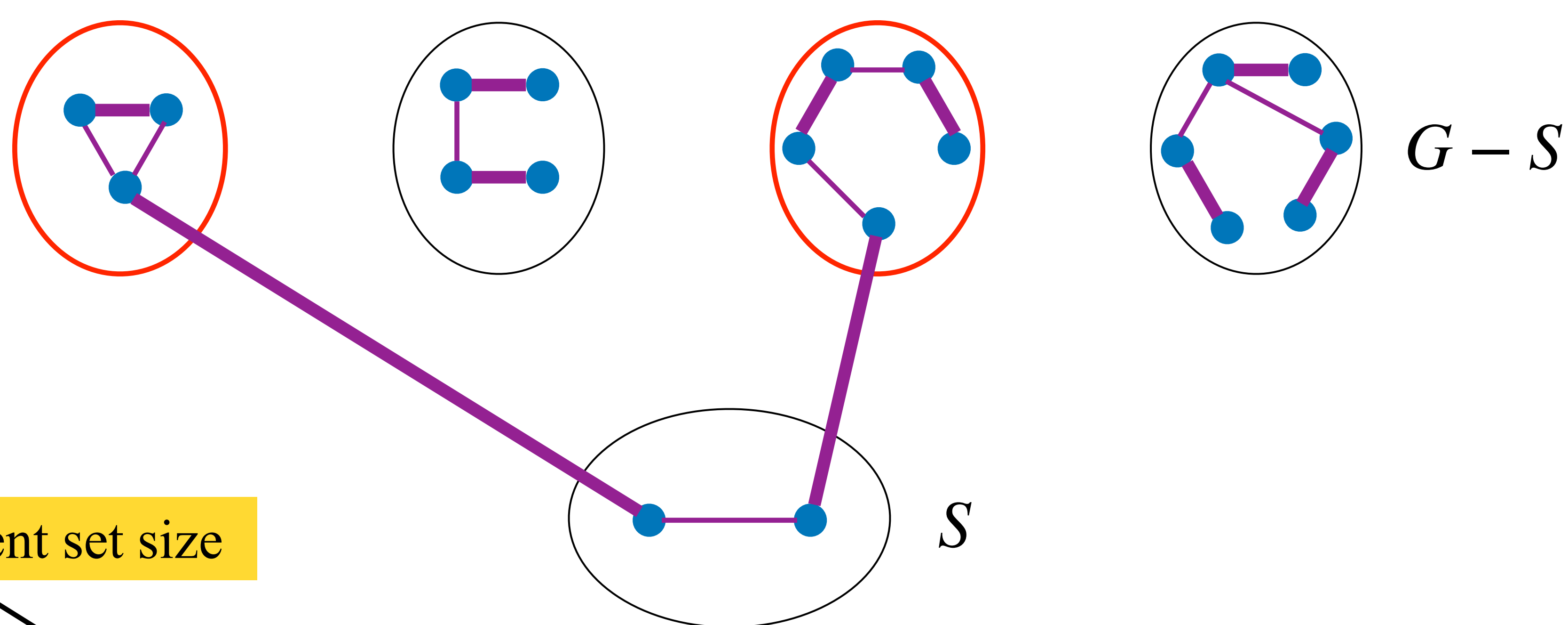
# Tutte-Berge theorem

**Theorem** [Tutte '47, Berge'58]

The number of unmatched vertices in a maximum matching of graph $G$ is given by

$$\max_{S \subseteq V(G)} \left( \mathrm{odd}(G-S) - |S| \right)$$

# of odd components when $S$ removed



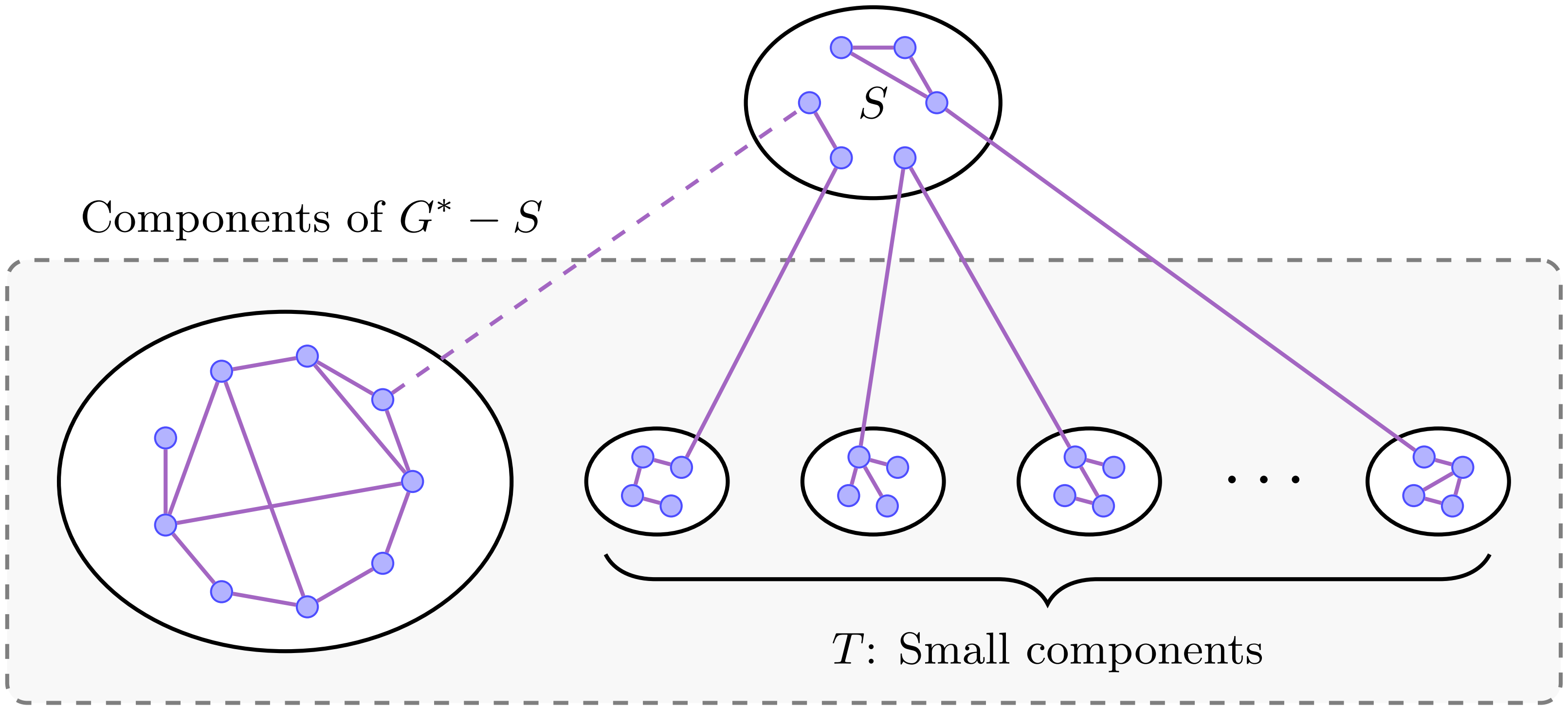

Maximum independent set size

Remark: $\mathrm{odd}(G-S) \leq \mathrm{Ind}(G) \Rightarrow$ It suffices to check $S$ with $|S| \leq \mathrm{Ind}(G)$

For a random $d$-regular graph $G$, $\mathrm{Ind}(G) \leq 2n \log(d)/d$

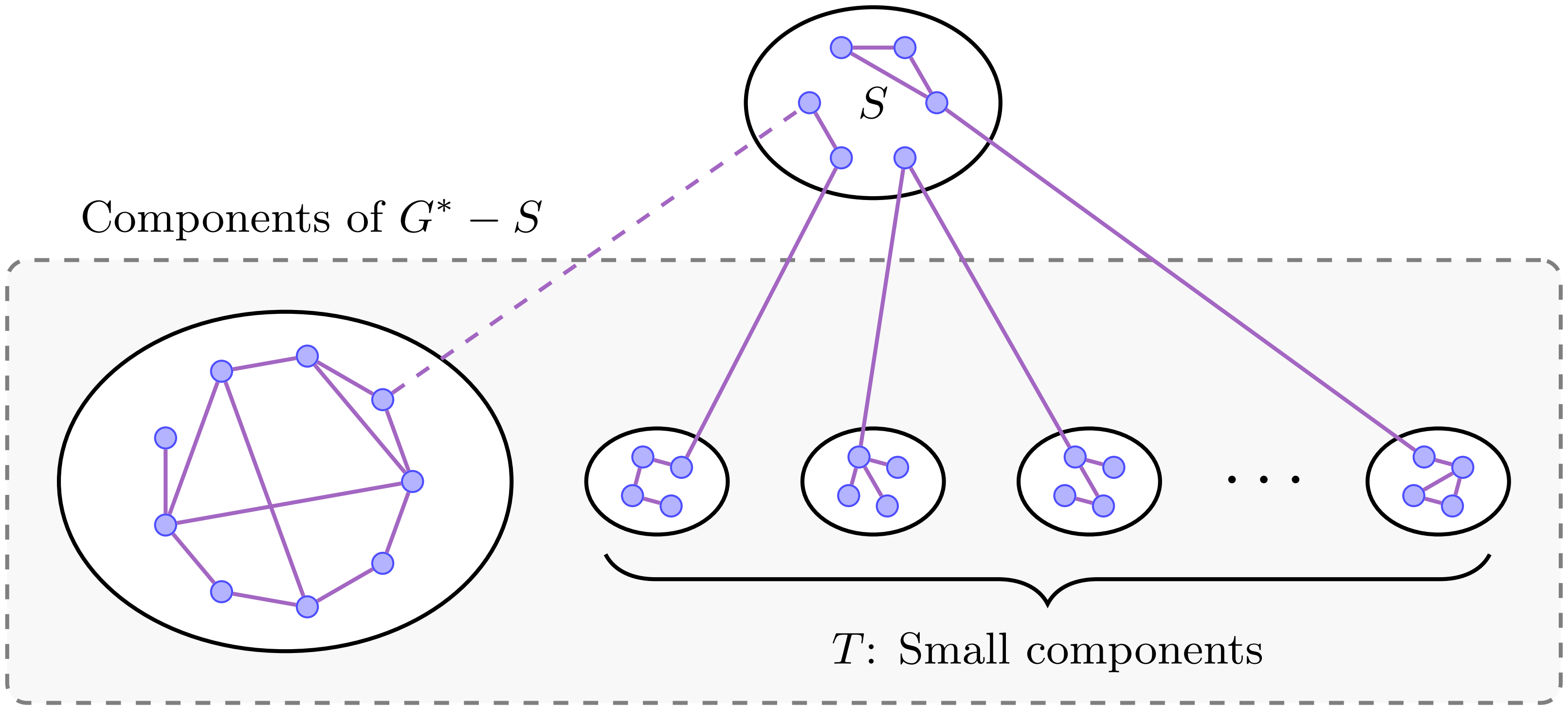
# Constant Loss Regime: Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $\text{odd}(G^* - S) \geq |S| + 2$ , where $G^* := G[\text{supp}(D)]$
- A bottleneck $S$ is called "small", if $|S| \leq 2n \log(d)/d$



# Constant Loss Regime: Ruling out Small Bottleneck $S$

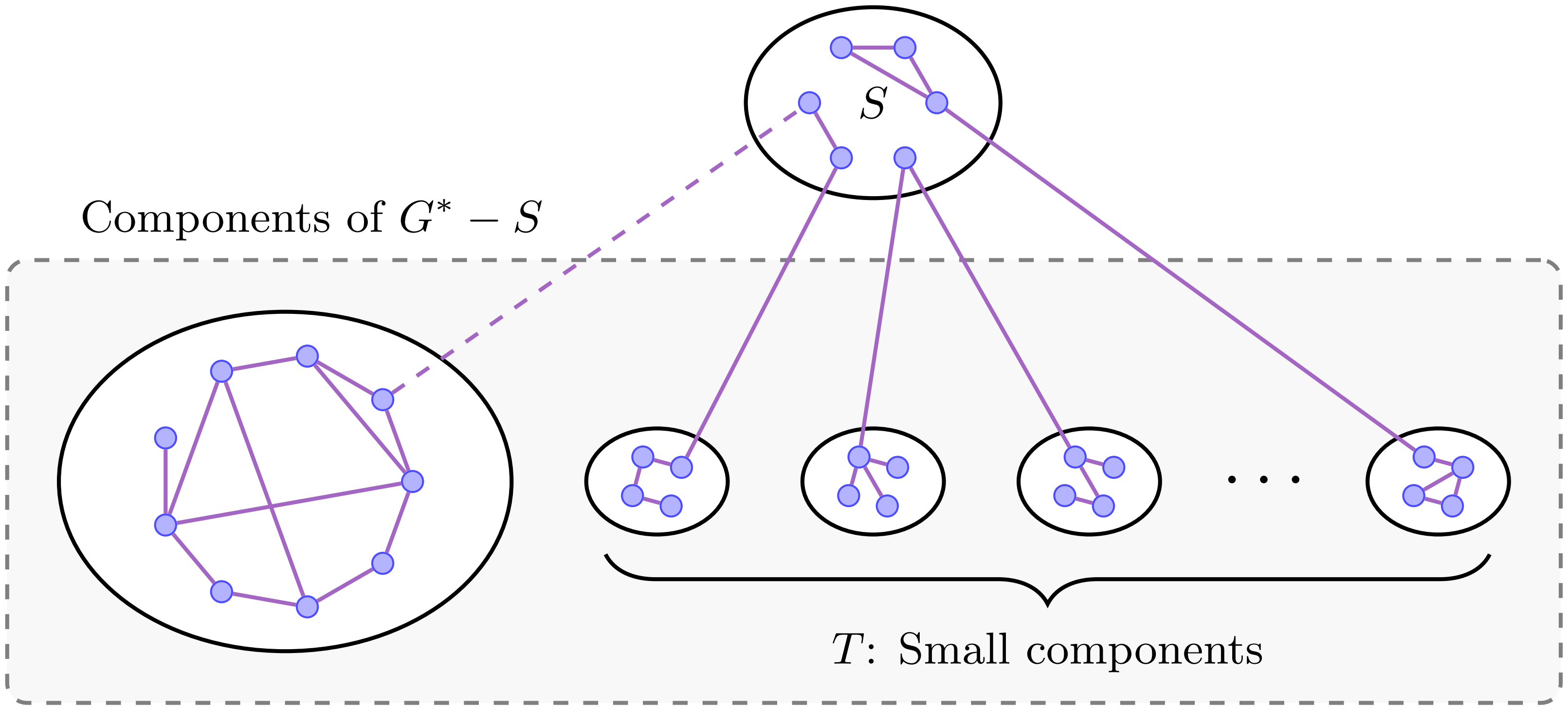
- $S$ is called a bottleneck, if $\text{odd}(G^* - S) \geq |S| + 2$ , where $G^* := G[\text{supp}(D)]$
- A bottleneck $S$ is called "small", if $|S| \leq 2n\log(d)/d$




- (No big cut) $\Rightarrow$ there exist at least $|S| + 1$ small components

# Constant Loss Regime: Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $\mathrm{odd}(G^* - S) \geq |S| + 2$ , where $G^* := G[\mathrm{supp}(D)]$
- A bottleneck $S$ is called "small", if $|S| \leq 2n\log(d)/d$



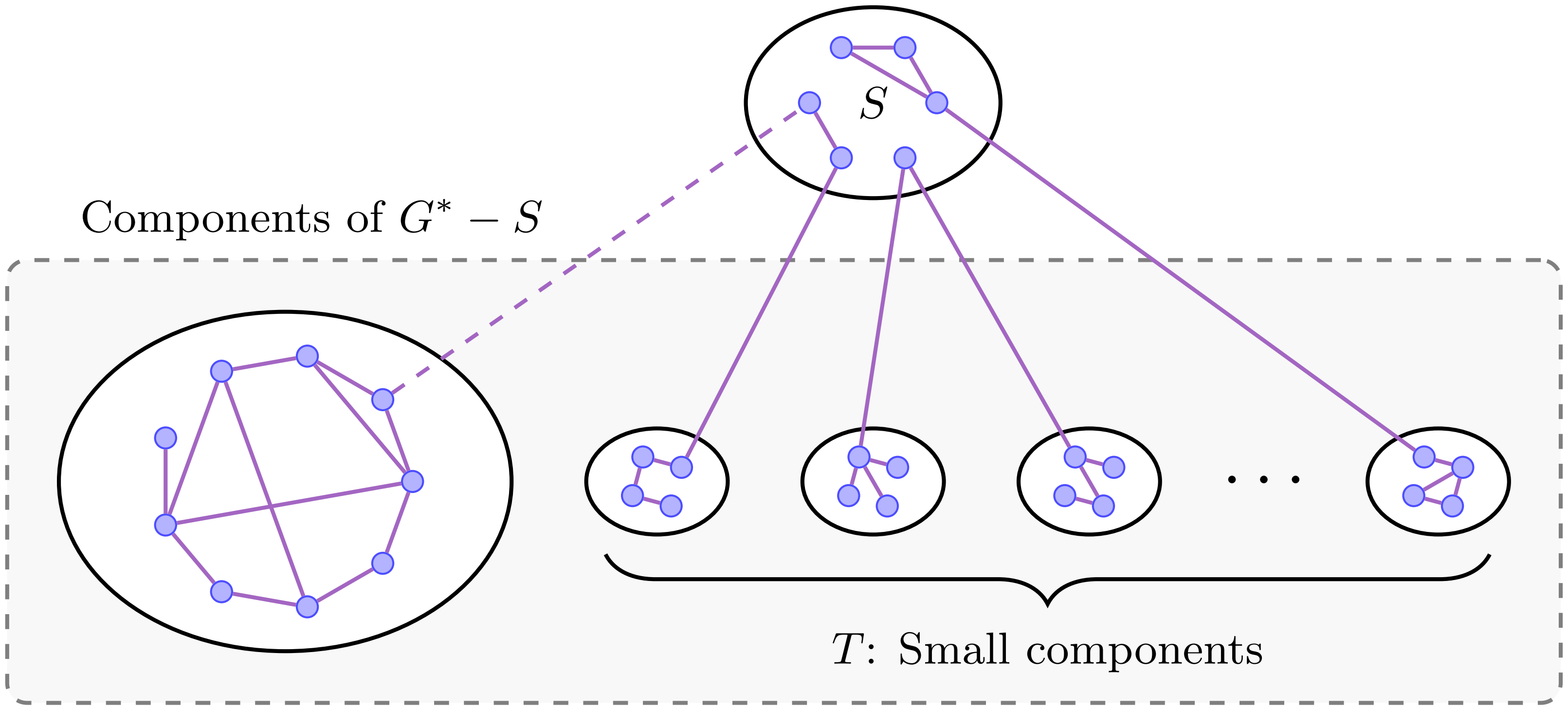

- (No big cut) $\Rightarrow$ there exist at least $|S| + 1$ small components
- (No small component in $G^*$) $\Rightarrow$ each of these small components must have at least one edge to $S$ $\Rightarrow e(S, T) \geq |S| + 1$

A stronger requirement than no isolated node

Satisfied when $d \geq \dfrac{(1+\delta)\log n}{\log(1/q)}$

# Constant Loss Regime: Ruling out Small Bottleneck $S$

- $S$ is called a bottleneck, if $\text{odd}(G^* - S) \geq |S| + 2$ , where $G^* := G[\text{supp}(D)]$
- A bottleneck $S$ is called "small", if $|S| \leq 2n \log(d)/d$




- (No big cut) $\Rightarrow$ there exist at least $|S| + 1$ small components
- (No small component in $G^*$) $\Rightarrow$ each of these small components must have at least one edge to $S$ $\Rightarrow e(S, T) \geq |S| + 1$
- (Little congestion) $\Rightarrow$ since both $S$ and $T$ are small, $e(S, T) \leq |S|$ (contradiction)

# Fractional Loss Regime: Local Weak Convergence

Asymptotic size of maximum matching can be determined locally

**Theorem** [Bordenave-Lelarge-Salez '13]

Suppose $G_n$ admits a local weak convergence to a Galton-Watson tree with degree distribution $\pi$. The fraction of unmatched vertices in a maximum matching

$$\alpha(G_n) \xrightarrow{\mathbb{P}} \max_{t \in [0,1]} F(t),$$

where $F(t) = t\phi'(1-t) + \phi(1-t) + \phi\left(1 - \frac{\phi'(1-t)}{\phi'(1)}\right) - 1$ and $\phi(t) = \sum_k \pi_k t^k$.

# Fractional Loss Regime: Local Weak Convergence

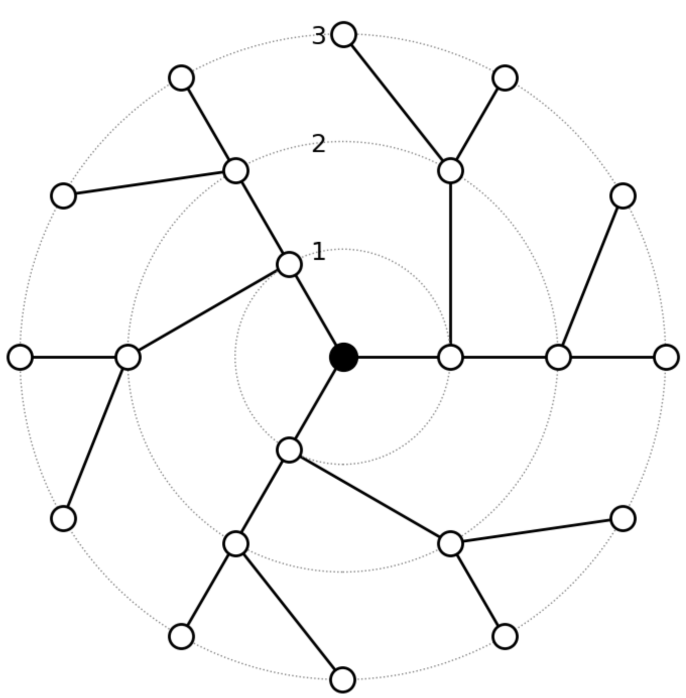
Asymptotic size of maximum matching can be determined locally

**Theorem** [Bordenave-Lelarge-Salez '13]

Suppose $G_n$ admits a local weak convergence to a Galton-Watson tree with degree distribution $\pi$. The fraction of unmatched vertices in a maximum matching

$$\alpha(G_n) \xrightarrow{\mathbb{P}} \max_{t \in [0,1]} F(t)\,,$$

where $F(t) = t\phi'(1-t) + \phi(1-t) + \phi\left(1 - \frac{\phi'(1-t)}{\phi'(1)}\right) - 1$ and $\phi(t) = \sum_k \pi_k t^k$.




Local neighborhood of 3-regular graph

# Fractional Loss Regime: Local Weak Convergence

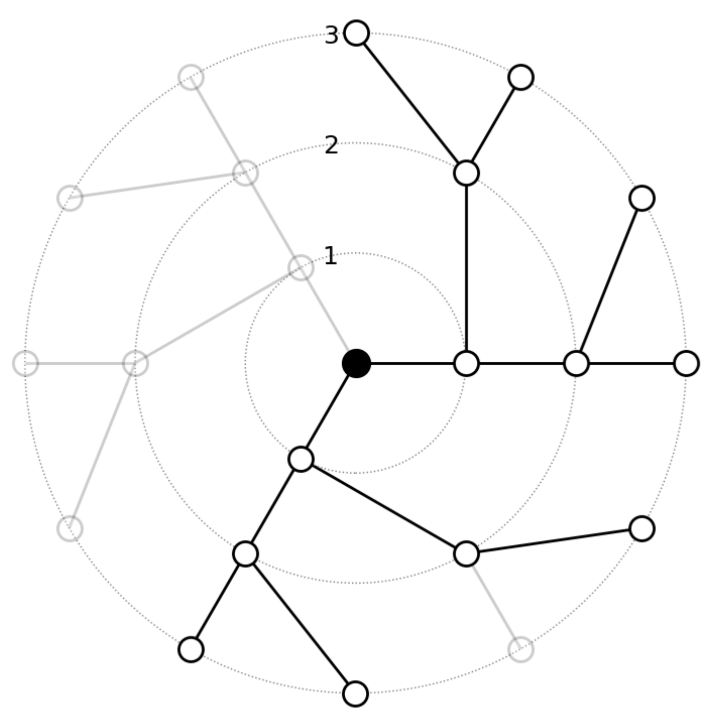
Asymptotic size of maximum matching can be determined locally

**Theorem** [Bordenave-Lelarge-Salez '13]

Suppose $G_n$ admits a local weak convergence to a Galton-Watson tree with degree distribution $\pi$. The fraction of unmatched vertices in a maximum matching

$$\alpha(G_n) \xrightarrow{\mathbb{P}} \max_{t \in [0,1]} F(t)\,,$$

where $F(t) = t\phi'(1-t) + \phi(1-t) + \phi\left(1 - \frac{\phi'(1-t)}{\phi'(1)}\right) - 1$ and $\phi(t) = \sum_k \pi_k t^k$.




Local neighborhood of 3-regular graph after vertex deletion

# Fractional Loss Regime: Local Weak Convergence

Asymptotic size of maximum matching can be determined locally

**Theorem** [Bordenave-Lelarge-Salez '13]

Suppose $G_n$ admits a local weak convergence to a Galton-Watson tree with degree distribution $\pi$. The fraction of unmatched vertices in a maximum matching

$$\alpha(G_n) \xrightarrow{\mathbb{P}} \max_{t\in[0,1]} F(t)\,,$$

where $F(t) = t\phi'(1-t) + \phi(1-t) + \phi\left(1 - \dfrac{\phi'(1-t)}{\phi'(1)}\right) - 1$ and $\phi(t) = \sum_k \pi_k t^k$.

- In our setting, $G_n$ is the random $d$-regular graph with random vertex removal (with probability $q$). We show it converges locally to GW tree with degree distribution $\pi = \mathrm{Binom}(d, p)$ with $p = 1 - q$

# Fractional Loss Regime: Local Weak Convergence

Asymptotic size of maximum matching can be determined locally

**Theorem** [Bordenave-Lelarge-Salez '13]

Suppose $G_n$ admits a local weak convergence to a Galton-Watson tree with degree distribution $\pi$. The fraction of unmatched vertices in a maximum matching

$$\alpha(G_n) \xrightarrow{\mathbb{P}} \max_{t \in [0,1]} F(t)\,,$$

where $F(t) = t\phi'(1-t) + \phi(1-t) + \phi\left(1 - \dfrac{\phi'(1-t)}{\phi'(1)}\right) - 1$ and $\phi(t) = \sum_k \pi_k t^k$.

- In our setting, $G_n$ is the random $d$-regular graph with random vertex removal (with probability $q$). We show it converges locally to GW tree with degree distribution $\pi = \mathrm{Binom}(d,p)$ with $p = 1 - q$
- Thus, $\phi(t) = (pt + 1 - p)^d$ and we show $F(t)$ has three extremum points with the first one being the maximum and $\max_{t \in [0,1]} F(t) = (1 + o_d(1))q^d$

# Fractional Loss Regime: Local Weak Convergence

Asymptotic size of maximum matching can be determined locally

**Theorem** [Bordenave-Lelarge-Salez '13]

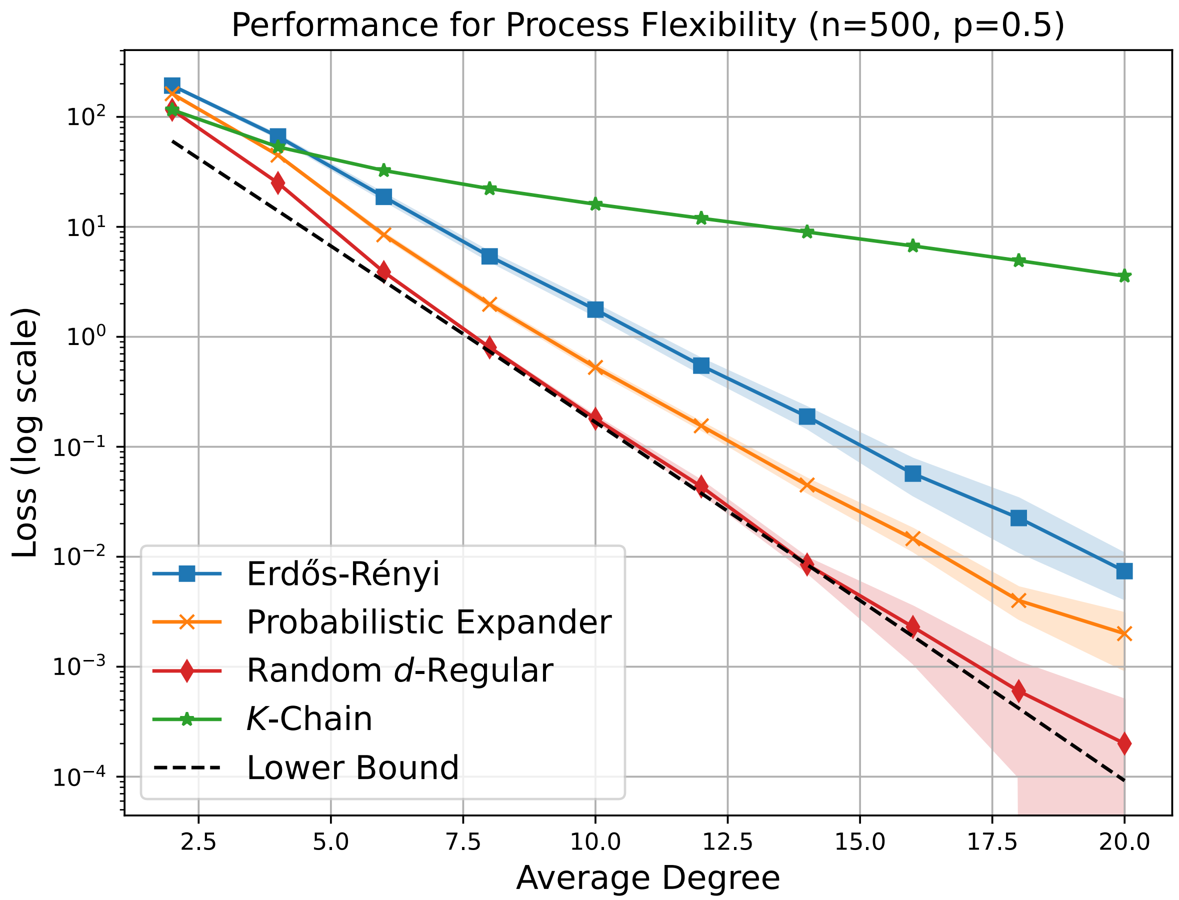
Suppose $G_n$ admits a local weak convergence to a Galton-Watson tree with degree distribution $\pi$. The fraction of unmatched vertices in a maximum matching

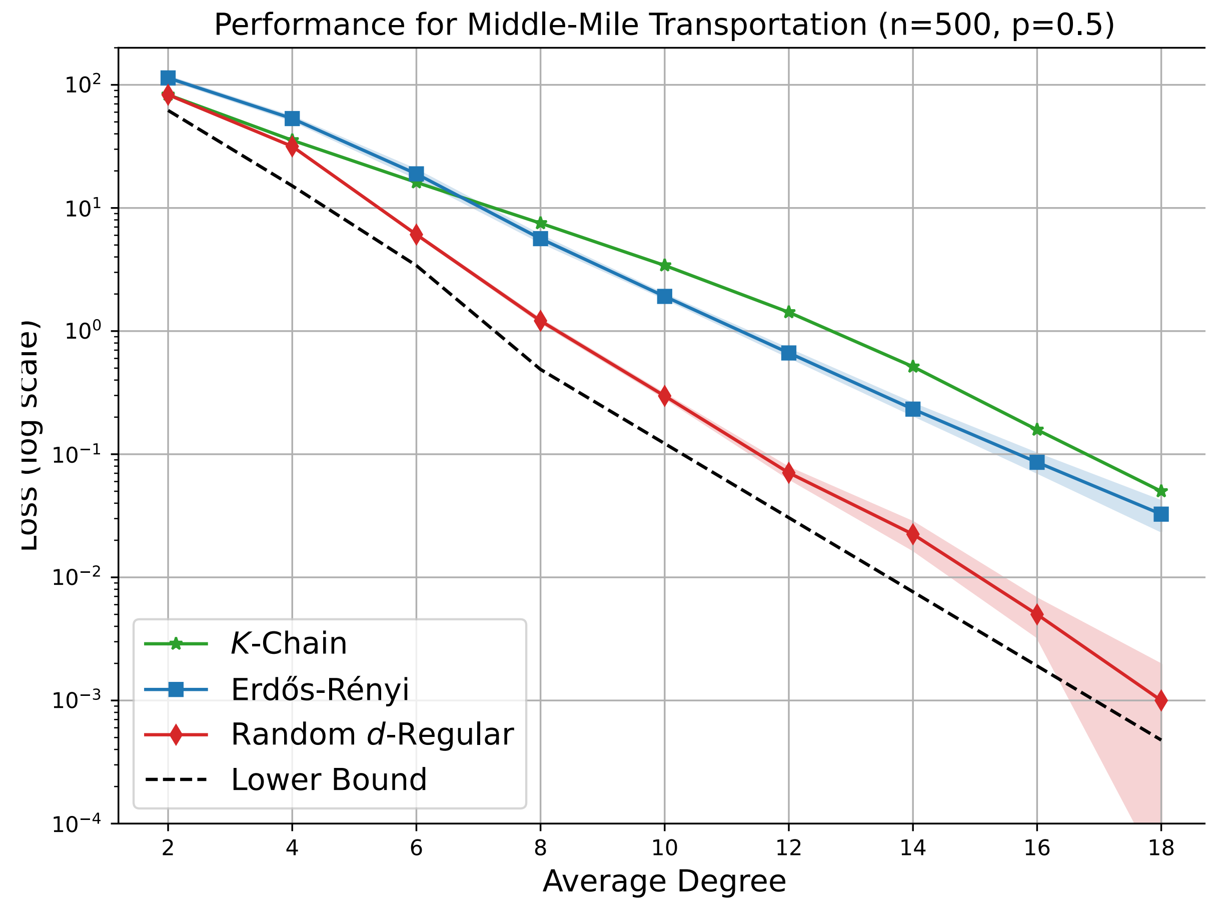
$$\alpha(G_n) \xrightarrow{\mathbb{P}} \max_{t \in [0,1]} F(t),$$

where $F(t) = t\phi'(1-t) + \phi(1-t) + \phi\left(1 - \frac{\phi'(1-t)}{\phi'(1)}\right) - 1$ and $\phi(t) = \sum_k \pi_k t^k$.

- In our setting, $G_n$ is the random $d$-regular graph with random vertex removal (with probability $q$). We show it converges locally to GW tree with degree distribution $\pi = \text{Binom}(d, p)$ with $p = 1 - q$
- Thus, $\phi(t) = (pt + 1 - p)^d$ and we show $F(t)$ has three extremum points with the first one being the maximum and $\max_{t \in [0,1]} F(t) = (1 + o_d(1))q^d$
- So, when $d \geq (1 + \delta)\log(1/\epsilon)/\log(1/q)$, $\alpha(G_n) \leq \epsilon$ and thus $\mathbb{E}_D[L(G, D)] \leq \epsilon n$

# Numerical Experiments




**Random regular graphs** can achieve performance close to the theoretical lower bounds even for moderate graph sizes and relatively small degrees

# Summary

| | Setting | Graph design | $\epsilon n$-loss | Constant loss |
|---|---|---|---|---|
| Lower bound | Both | Any | $\dfrac{\log(1/\epsilon)}{\log(1/q)}$ | $\dfrac{\log(n)}{\log(1/q)}$ |
| Chou et al. (2011) | Bipartite | Chain | $\Omega(1/\epsilon)$ | - |
| Chen et al. (2015) | Bipartite | Probabilistic Expander | $\Omega(\log(1/\epsilon))$ | - |
| Feng et al. (2024) | Unipartite | Ring; Chain; Cluster; Erdős–Rényi | $\Omega(\log(1/\epsilon))$ | $\Omega(\log(n))$ |
| Our work | Both | Random regular | $\dfrac{\log(1/\epsilon)}{\log(1/q)}^{*}$ | $\dfrac{\log(n)}{\log(1/q)}^{*}$ |

*Within a factor of $(1+\delta)$ for arbitrarily small $\delta > 0$.

# Summary

| | Setting | Graph design | $\epsilon n$-loss | Constant loss |
|---|---|---|---|---|
| Lower bound | Both | Any | $\frac{\log(1/\epsilon)}{\log(1/q)}$ | $\frac{\log(n)}{\log(1/q)}$ |
| Chou et al. (2011) | Bipartite | Chain | $\Omega(1/\epsilon)$ | - |
| Chen et al. (2015) | Bipartite | Probabilistic Expander | $\Omega(\log(1/\epsilon))$ | - |
| Feng et al. (2024) | Unipartite | Ring; Chain; Cluster; Erdős–Rényi | $\Omega(\log(1/\epsilon))$ | $\Omega(\log(n))$ |
| Our work | Both | Random regular | $\frac{\log(1/\epsilon)}{\log(1/q)}$ * | $\frac{\log(n)}{\log(1/q)}$ * |

*Within a factor of $(1+\delta)$ for arbitrarily small $\delta > 0$.

**Future directions:**

General demand distribution? Other flexibility network design problems?

# Summary

| | Setting | Graph design | $\epsilon n$-loss | Constant loss |
|---|---|---|---|---|
| Lower bound | Both | Any | $\frac{\log(1/\epsilon)}{\log(1/q)}$ | $\frac{\log(n)}{\log(1/q)}$ |
| Chou et al. (2011) | Bipartite | Chain | $\Omega(1/\epsilon)$ | - |
| Chen et al. (2015) | Bipartite | Probabilistic Expander | $\Omega(\log(1/\epsilon))$ | - |
| Feng et al. (2024) | Unipartite | Ring; Chain; Cluster; Erdős–Rényi | $\Omega(\log(1/\epsilon))$ | $\Omega(\log(n))$ |
| Our work | Both | Random regular | $\frac{\log(1/\epsilon)}{\log(1/q)}$ * | $\frac{\log(n)}{\log(1/q)}$ * |

*Within a factor of $(1+\delta)$ for arbitrarily small $\delta > 0$.

**Future directions:**

General demand distribution? Other flexibility network design problems?

**Reference:**

W. Li, X. Niu, Y. Wei, & J. Xu, *Optimality of Random Regular Graphs in Sparse Network Designs*, SSRN Preprint, March 2026