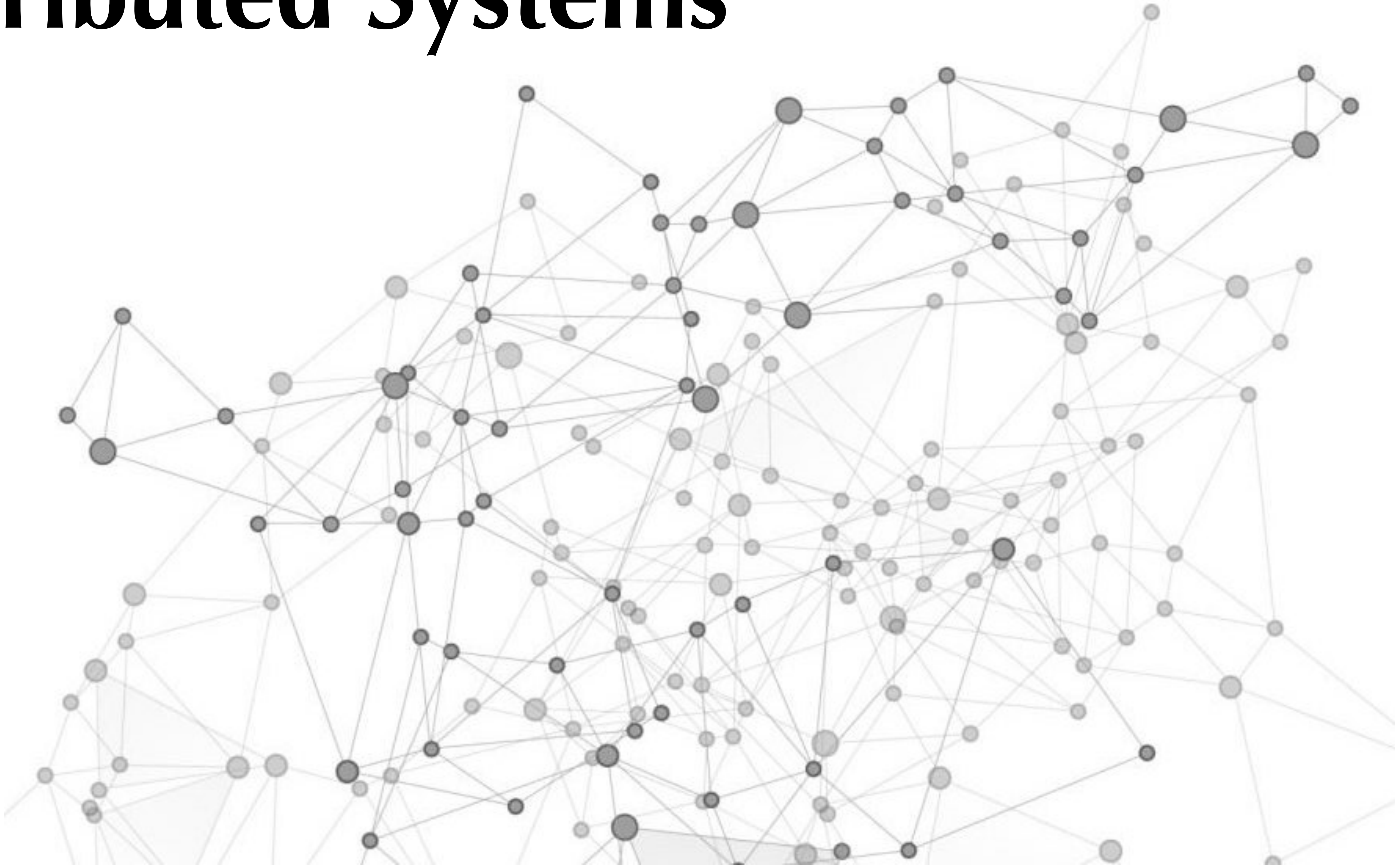
Network Design and Performance Analysis for Reliable Inference in Distributed Systems

12/11/2023

SNAPP Seminar



Mansi Sood

PhD Candidate

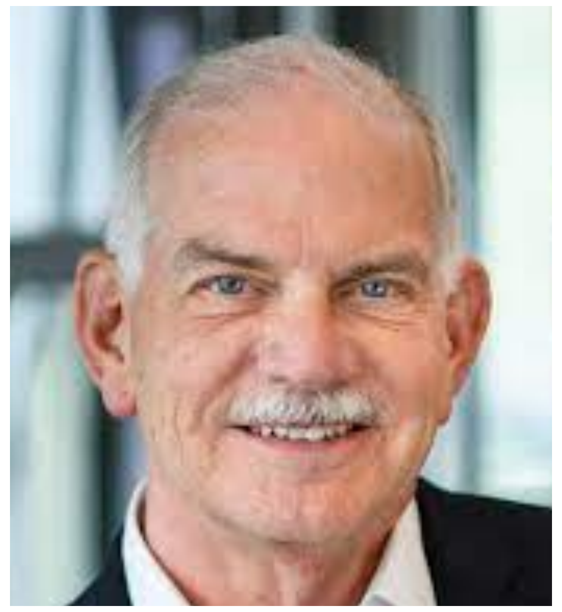
Carnegie Mellon University

Supported by Philip and Marsha Dowd Fellowship, Cylab Presidential Fellowship, David H. Barakat and LaVerne Owen-Barakat Fellowship,, Lee-Stanziale Ohana Endowed Fellowship, Knight Fellowship, NSF CCF #1617934, Office of Naval Research (ONR) #N00014-23-1-2275, the Air Force Office of Scientific Research (AFOSR) # FA9550-22-1-0233, and IoT@CyLab

Acknowledgements

Osman Yagan
CMU

Chai Wah Wu
IBM Research



H. Vincent Poor
Princeton

Simon A. Levin
Princeton

Giulia Fanti
CMU

Ani Sridhar
MIT

Eray Can Elumar
CMU

Rashad Eletreby
Rocket Travel Inc

Rachid El-Azouzi
Avignon University

Gireeja Ranade
UC Berkeley

Supported by Philip and Marsha Dowd Fellowship, Cylab Presidential Fellowship, David H. Barakat and LaVerne Owen-Barakat Fellowship,, Lee-Stanziale Ohana Endowed Fellowship, Knight Fellowship, NSF CCF #1617934, Office of Naval Research (ONR) #N00014-23-1-2275, the Air Force Office of Scientific Research (AFOSR) # FA9550-22-1-0233, and IoT@CyLab

Research Overview

Theme *How can we leverage network structure to better understand and design socio-technical systems?*

Thrusts

Network design and performance analysis for reliable inference in distributed systems

Modeling, analyzing, and controlling spreading processes in social networks

Research Overview

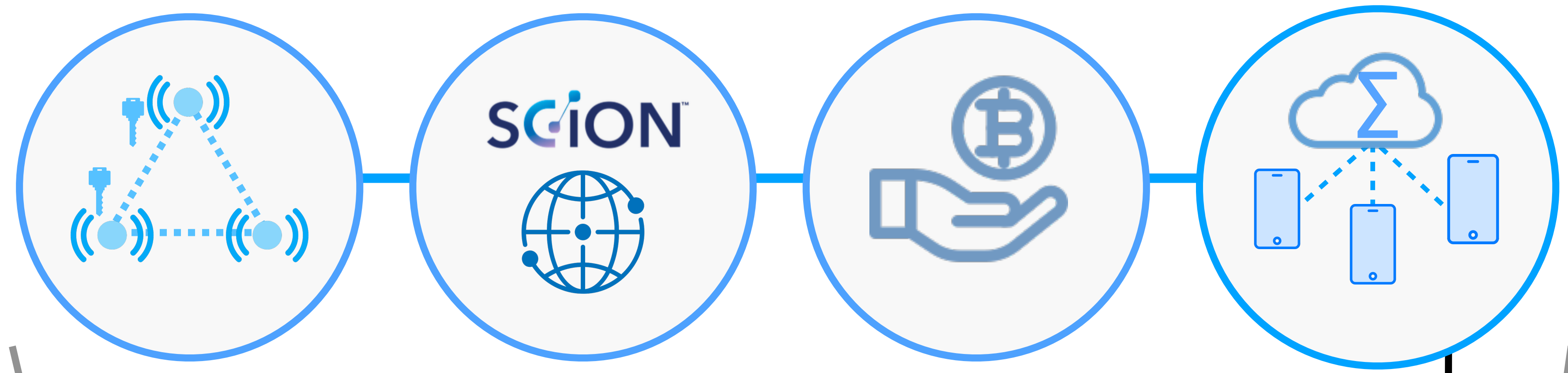
Theme

How can we leverage network structure to better understand and design socio-technical systems?

Thrusts

Network design and performance analysis for reliable inference in distributed systems

Modeling, analyzing, and controlling spreading processes in social networks



Contributions

Formal characterization of strength of connectivity of 'random K-out graphs'

IEEE Transactions on Information Theory '21, '23
IEEE ICC '21 (Best Paper Award)
ISIT '21, '20, CDC '20, Globecom '19

Ongoing, future work

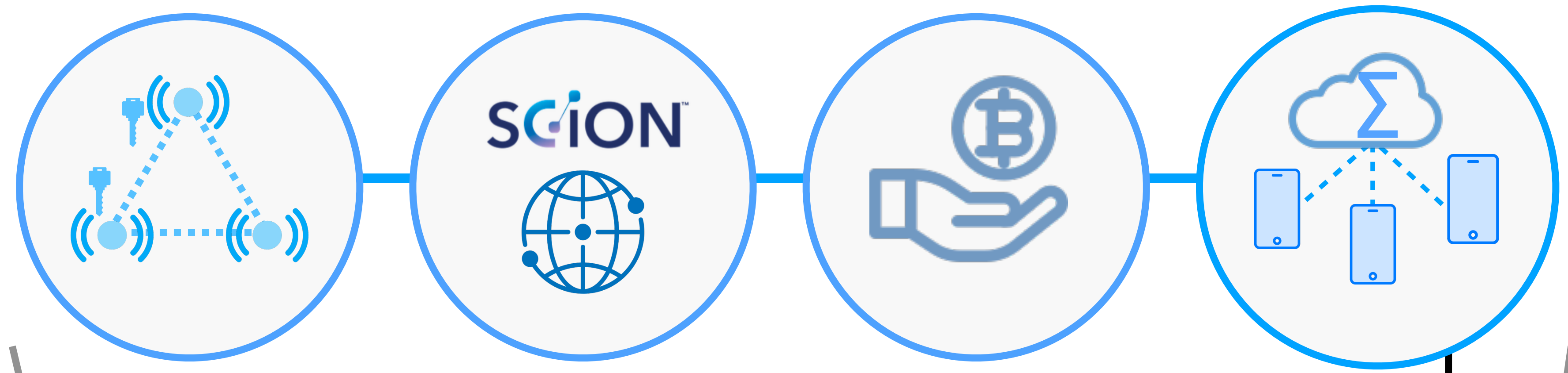
Privacy-scalability frontiers in distributed & decentralized learning

Research Overview

Theme *How can we leverage network structure to better understand and design socio-technical systems?*

Thrusts

Network design and performance analysis for reliable inference in distributed systems



Modeling, analyzing, and controlling spreading processes in social networks

Contributions

Formal characterization of strength of connectivity of 'random K-out graphs'

IEEE Transactions on Information Theory '21, '23
IEEE ICC '21 (Best Paper Award)
ISIT '21, '20, CDC '20, Globecom '19

Analyzing spreading processes trigerred by evolving contagions

Proceedings of the National Academy of Sciences, '23
IEEE ICC '23

Ongoing, future work

Privacy-scalability frontiers in distributed & decentralized learning

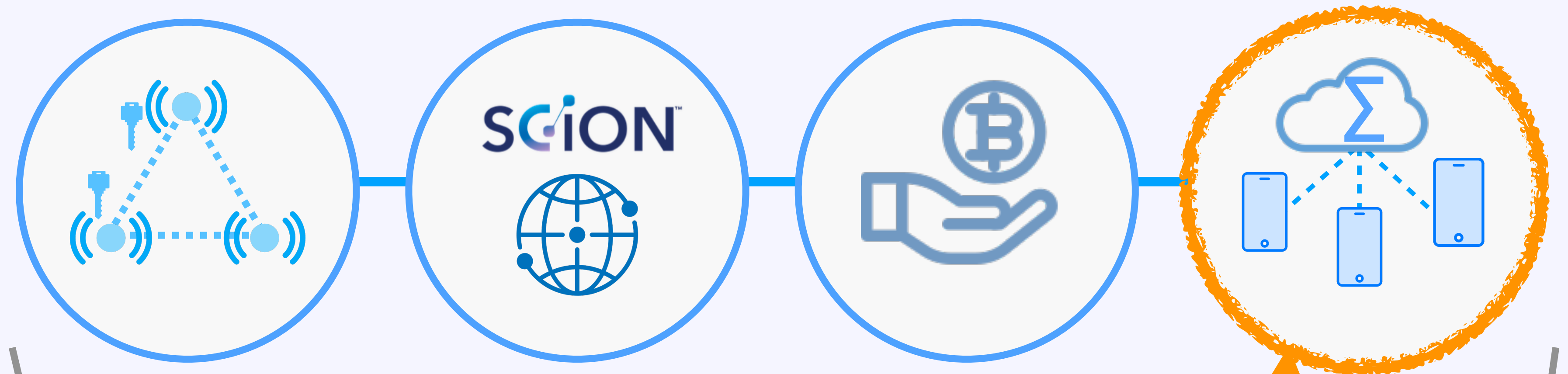
Decentralized contentent moderation
Cross-platform interactions & information spread

Research Overview

Theme *How can we leverage network structure to better understand and design socio-technical systems?*

Thrusts

Network design and performance analysis for reliable inference in distributed systems



Modeling, analyzing, and controlling spreading processes in social networks

Contributions

(Today's focus)

Formal characterization of strength of connectivity of 'random K-out graphs'

IEEE Transactions on Information Theory '21, '23
IEEE ICC '21 (Best Paper Award)
ISIT '21, '20, CDC '20, Globecom '19

(Joint work with O. Yagan and E. C. Elumar)

Analyzing spreading processes trigerred by evolving contagions

Proceedings of the National Academy of Sciences, '23
IEEE ICC '23

Random K-out Graphs

$\mathbb{H}(n, K)$

Each node selects K neighboring nodes

chosen uniformly at random from all n-1 nodes

Edge (i, j) exists if

node i selects node j

or

node j selects node i

Random K-out Graphs

$\mathbb{H}(n, K)$

Each node selects K neighboring nodes

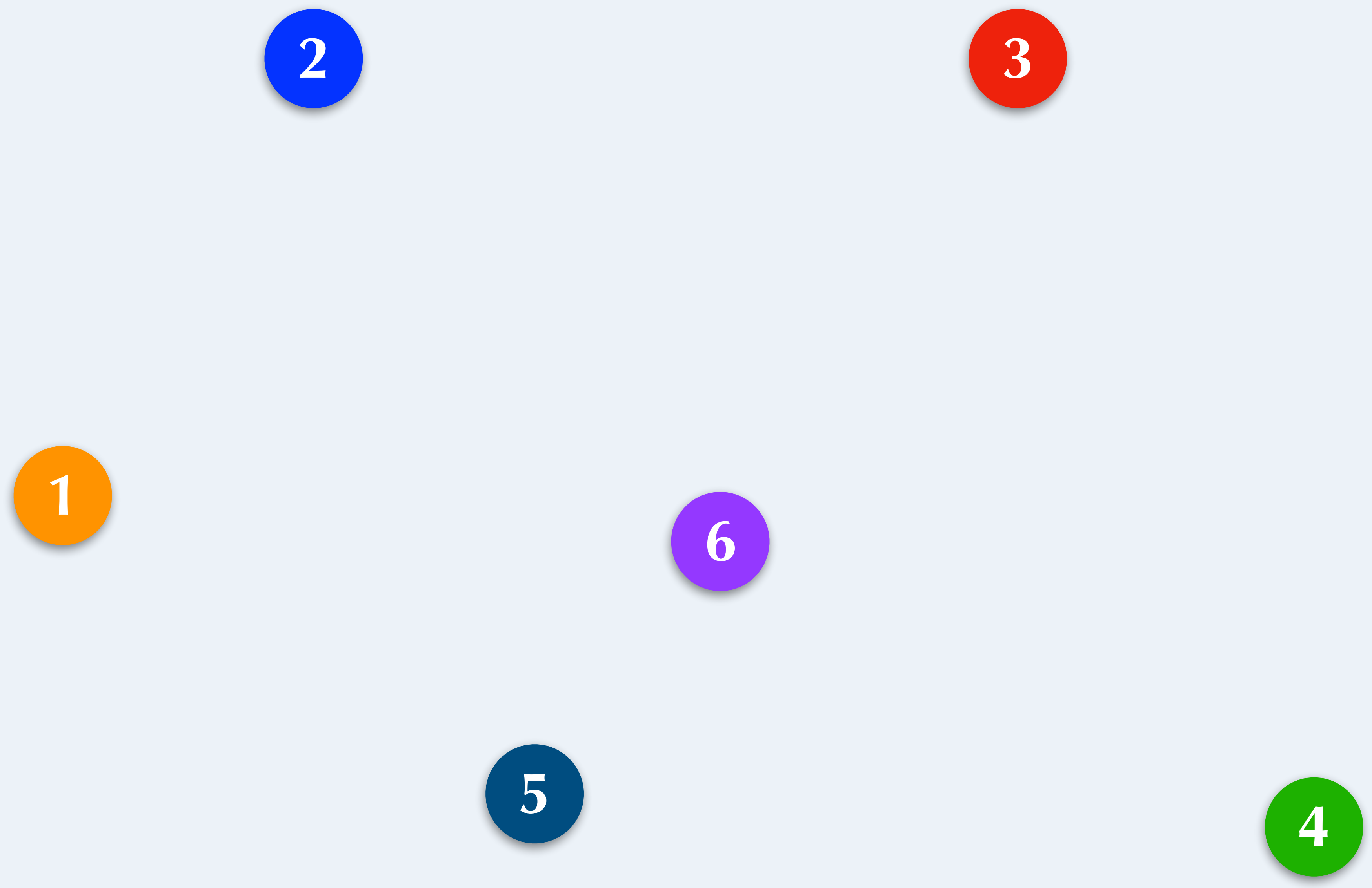
chosen uniformly at random from all n-1 nodes

Edge (i, j) exists if

node i selects node j

or

node j selects node i



Random 2-out graph on 6 nodes

$(K = 2,\ n = 6)$

Random K-out Graphs

$\mathbb{H}(n, K)$

Each node selects K neighboring nodes

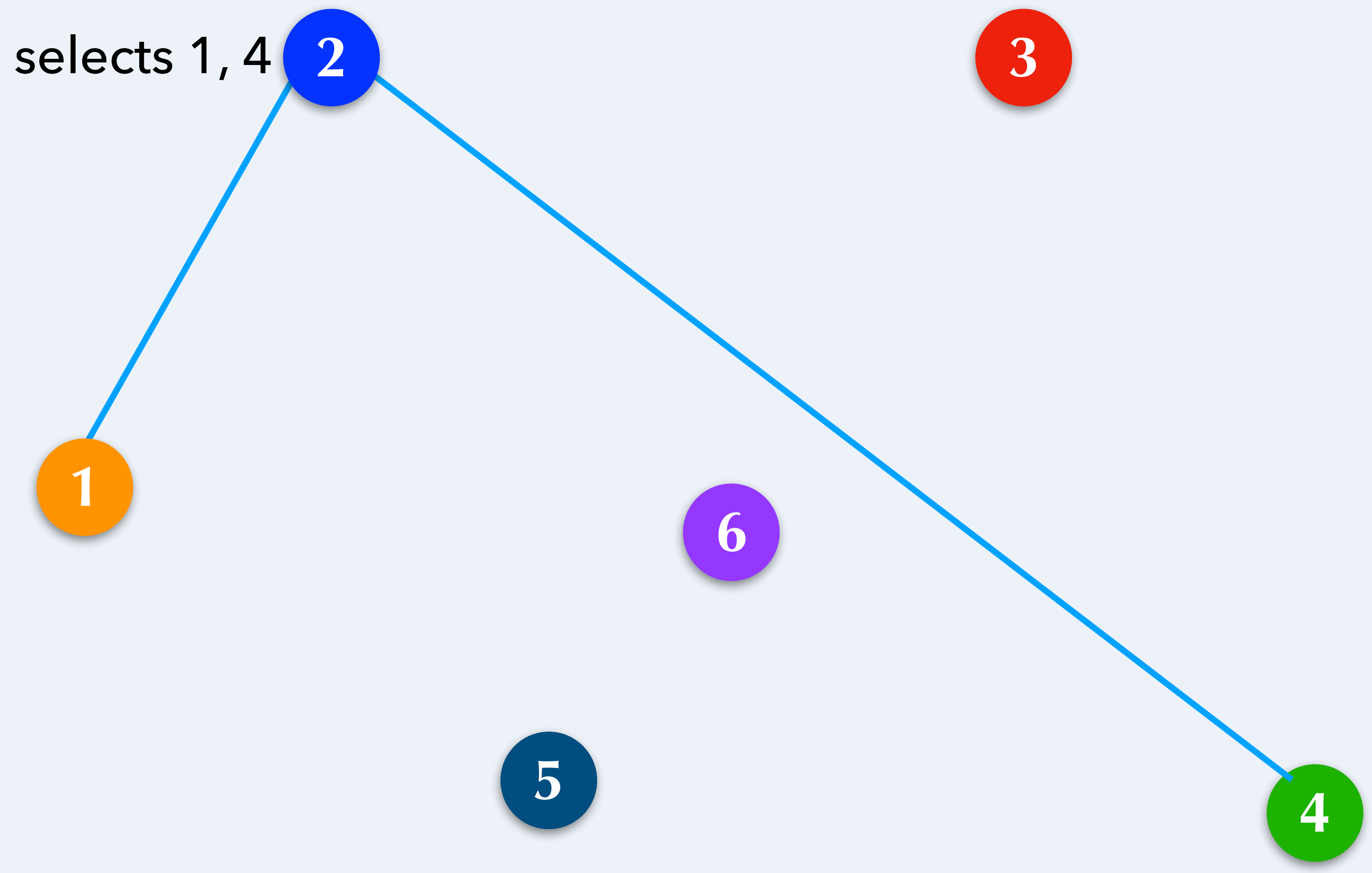
chosen uniformly at random from all n-1 nodes

Edge (i, j) exists if

node i selects node j

or

node j selects node i



Random 2-out graph on 6 nodes

$(K = 2,\ n = 6)$

Random K-out Graphs

$\mathbb{H}(n, K)$

Each node selects K neighboring nodes

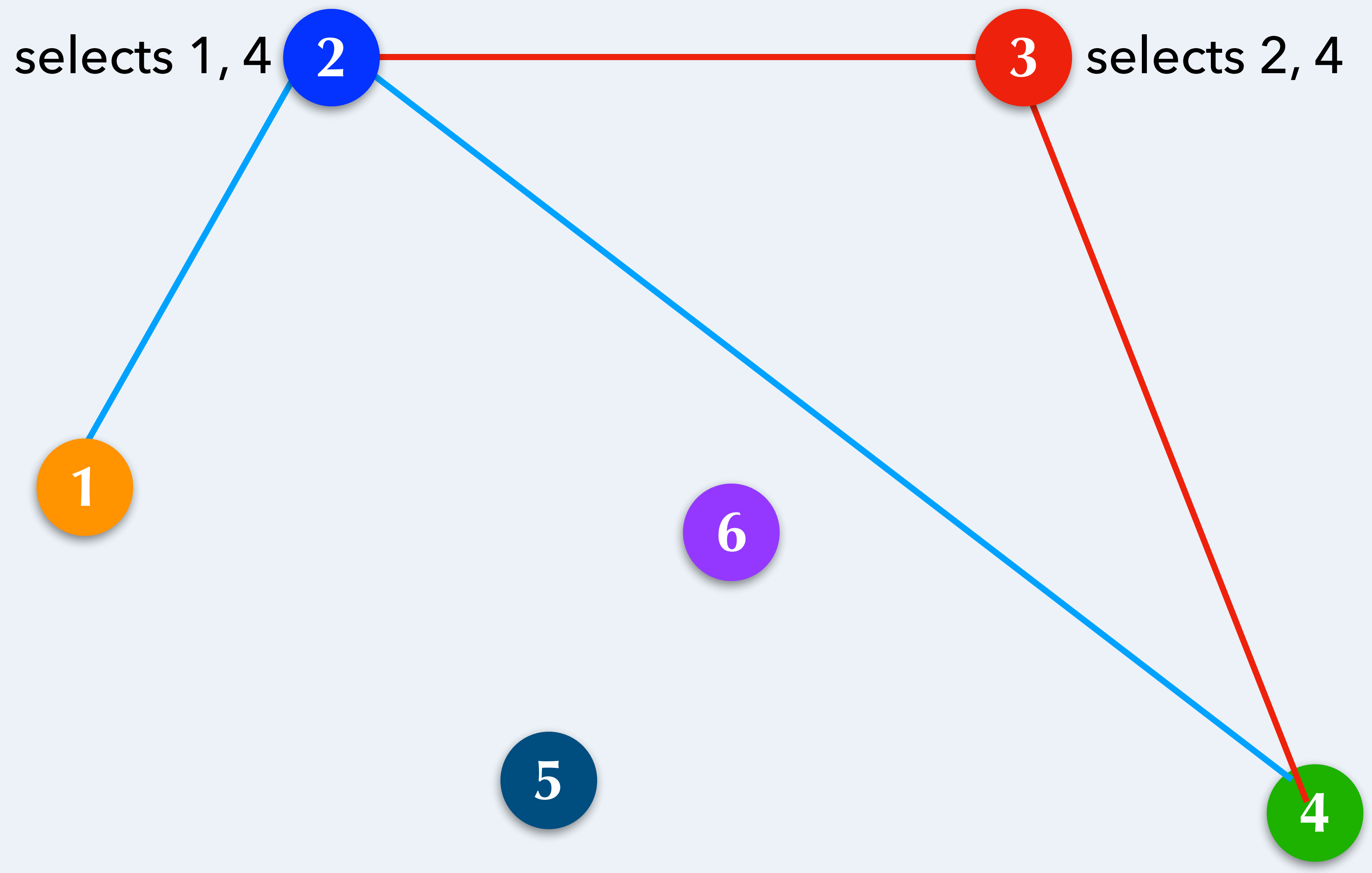
chosen uniformly at random from all n-1 nodes

Edge (i, j) exists if

node i selects node j

or

node j selects node i



Random 2-out graph on 6 nodes

$(K = 2,\ n = 6)$

Random K-out Graphs

$\mathbb{H}(n, K)$

Each node selects K neighboring nodes

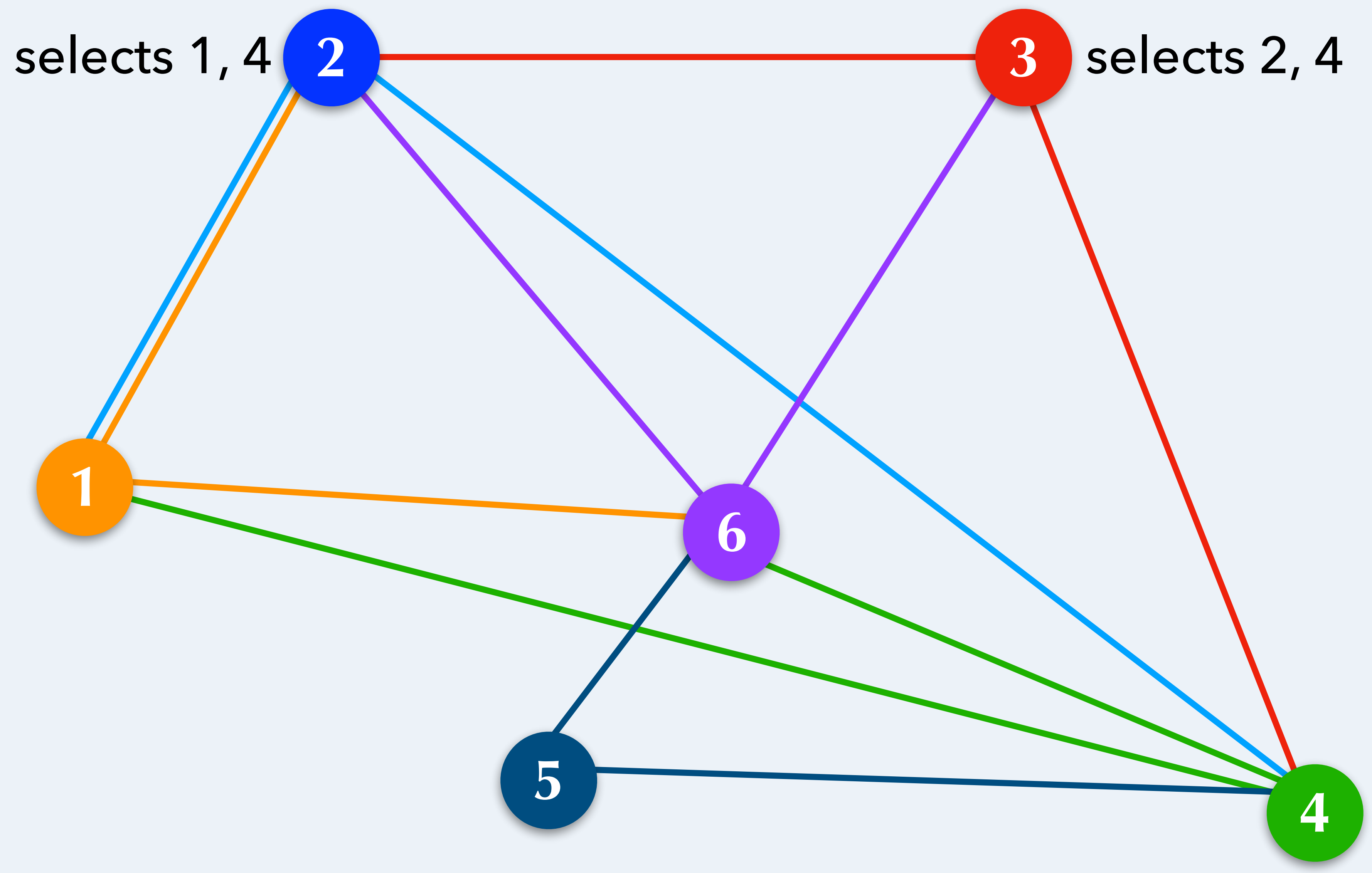
chosen uniformly at random from all n-1 nodes

Edge (i, j) exists if

node i selects node j

or

node j selects node i



Random 2-out graph on 6 nodes

$(K = 2,\ n = 6)$

Random K-out Graphs

$\mathbb{H}(n, K)$

Each node selects K neighboring nodes

chosen uniformly at random from all n-1 nodes

Edge (i, j) exists if

node i selects node j

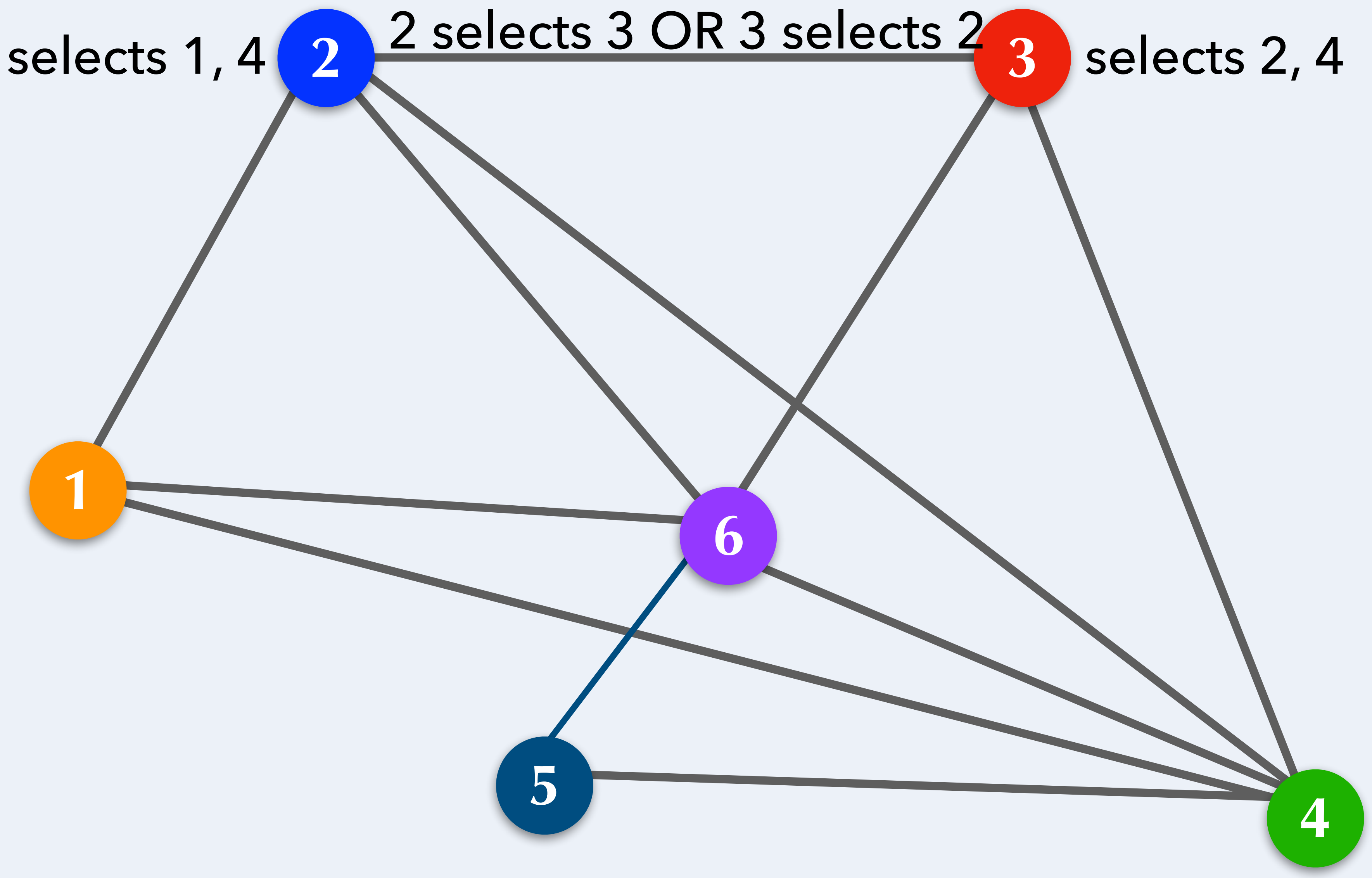
or

node j selects node i

Edge between 2 & 3

iff

2 selects 3 OR 3 selects 2



Random 2-out graph on 6 nodes

$(K = 2,\ n = 6)$

Connectivity

[Fenner and Frieze '82]

Random K-out Graphs $\mathbb{H}(n, K)$

For $K \geq 2$, *connected* with high probability (with probability $\rightarrow 1$ as # nodes $\rightarrow \infty$).

*(for $K = 1$, dis*connected *with high probability)*

Connectivity

[Fenner and Frieze '82]

Random K-out Graphs $\mathbb{H}(n, K)$

For $K \geq 2$, *connected* with high probability (with probability $\rightarrow 1$ as # nodes $\rightarrow \infty$).

(for $K = 1$, *disconnected* with high probability)

With average degree ~ 4, we get connectivity whp

Connectivity

[Fenner and Frieze '82]

Random K-out Graphs $\mathbb{H}(n, K)$

For $K \geq 2$, *connected* with high probability (with probability $\rightarrow 1$ as # nodes $\rightarrow \infty$).

(for $K = 1$, *disconnected* with high probability)

With average degree ~ 4, we get connectivity whp

Erdos Renyi Random Graphs $\mathbb{G}(n, p)$

In contrast Erdos Renyi random graphs require average degree ~ $\log n$ for connectivity whp

scales with n

Connectivity

Random K-out Graphs $\mathbb{H}(n, K)$

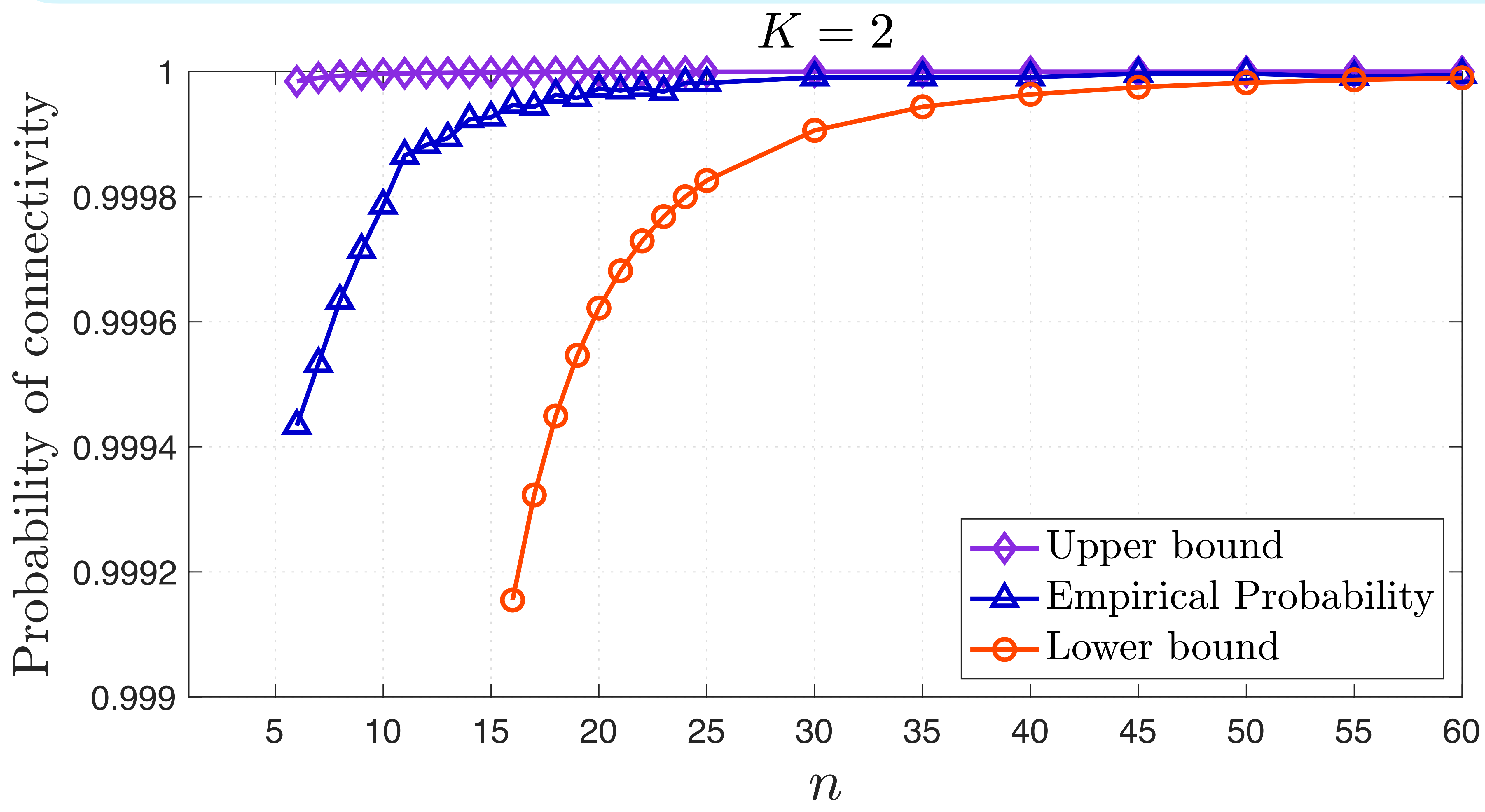
[Fenner and Frieze '82]

For $K \geq 2$, *connected* with high probability (with probability $\rightarrow 1$ as # nodes $\rightarrow \infty$).

(for $K = 1$, *disconnected* with high probability)

Theorem [Sood and Yagan, ICC'21*]

$$\mathbb{P}[\ \mathbb{H}(n, K) \text{ is } \textit{connected}\ ] = 1 - \Theta(1/n^{K^2-1}),\ K \geq 2$$



*Best Paper Award

What if K is not same for all nodes?

So far…

For (homogeneous) random K-out graphs, $p_{\text{connectivity}} = 1 - \Theta(1/n^{K^2-1}),\ K \geq 2$

What if some nodes make fewer than 2 selections?

Inhomogeneous Random K-out Graphs

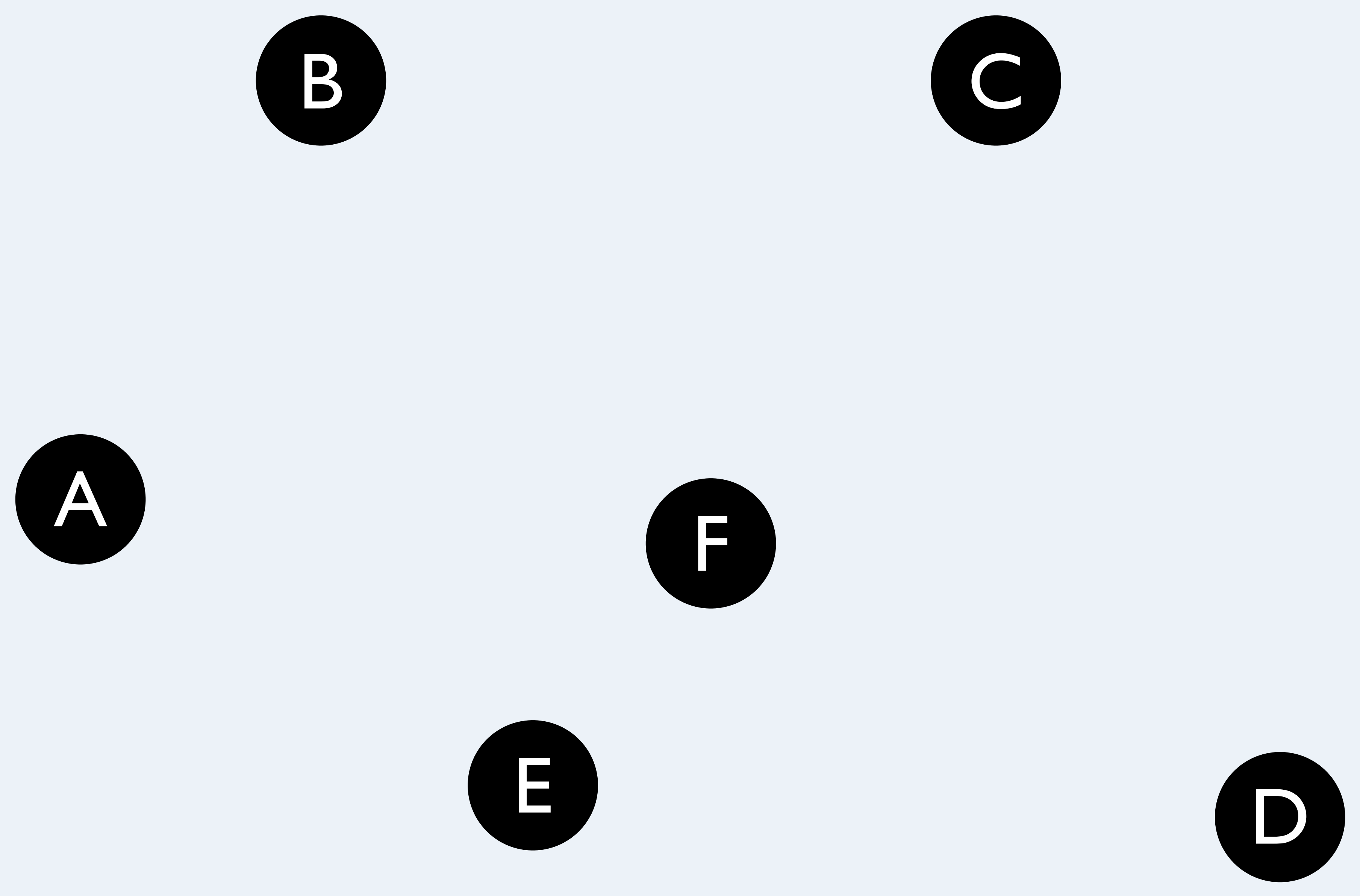
- Each node is assigned a type which determines the number of selections
- Nodes can make fewer than 2 selections

Inhomogeneous Random K-out Graphs

$\mathbb{H}(n, \mu, K_n)$

$\boldsymbol{n}$: number of nodes

Label nodes independently as
Type-I wp $\boldsymbol{\mu}$ (>0), Type-II wp $\mathbf{1}$-$\boldsymbol{\mu}$



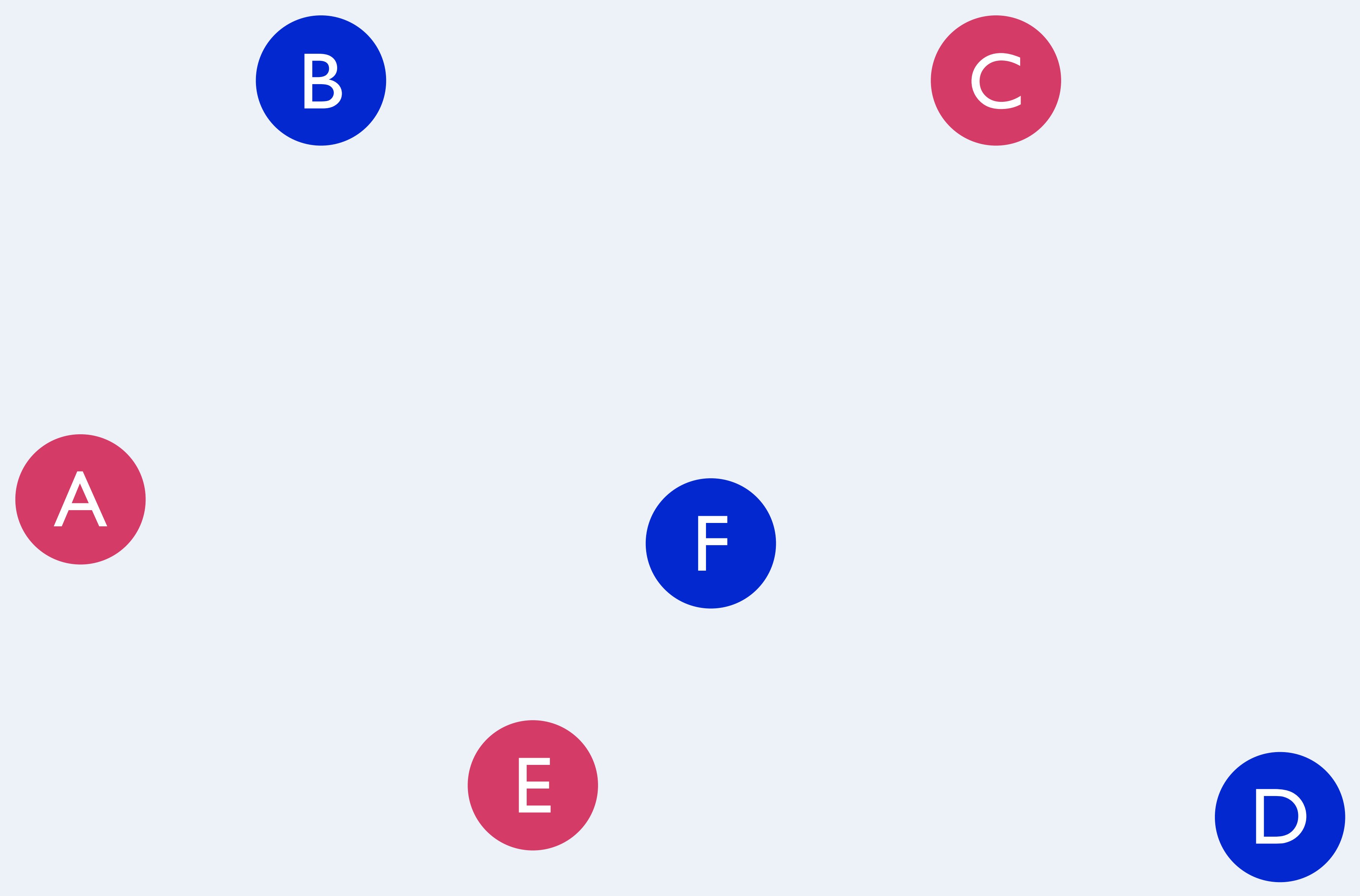
Inhomogeneous K-out Random graph (n = 6, $\boldsymbol{K_n}$ = 3)

Inhomogeneous Random K-out Graphs

$\mathbb{H}(n, \mu, K_n)$

$\boldsymbol{n}$: number of nodes

Label nodes independently as
Type-I wp $\boldsymbol{\mu}$ (>0), Type-II wp $\mathbf{1}$-$\boldsymbol{\mu}$



Inhomogeneous K-out Random graph (n = 6, $\boldsymbol{K_n}$ = 3)

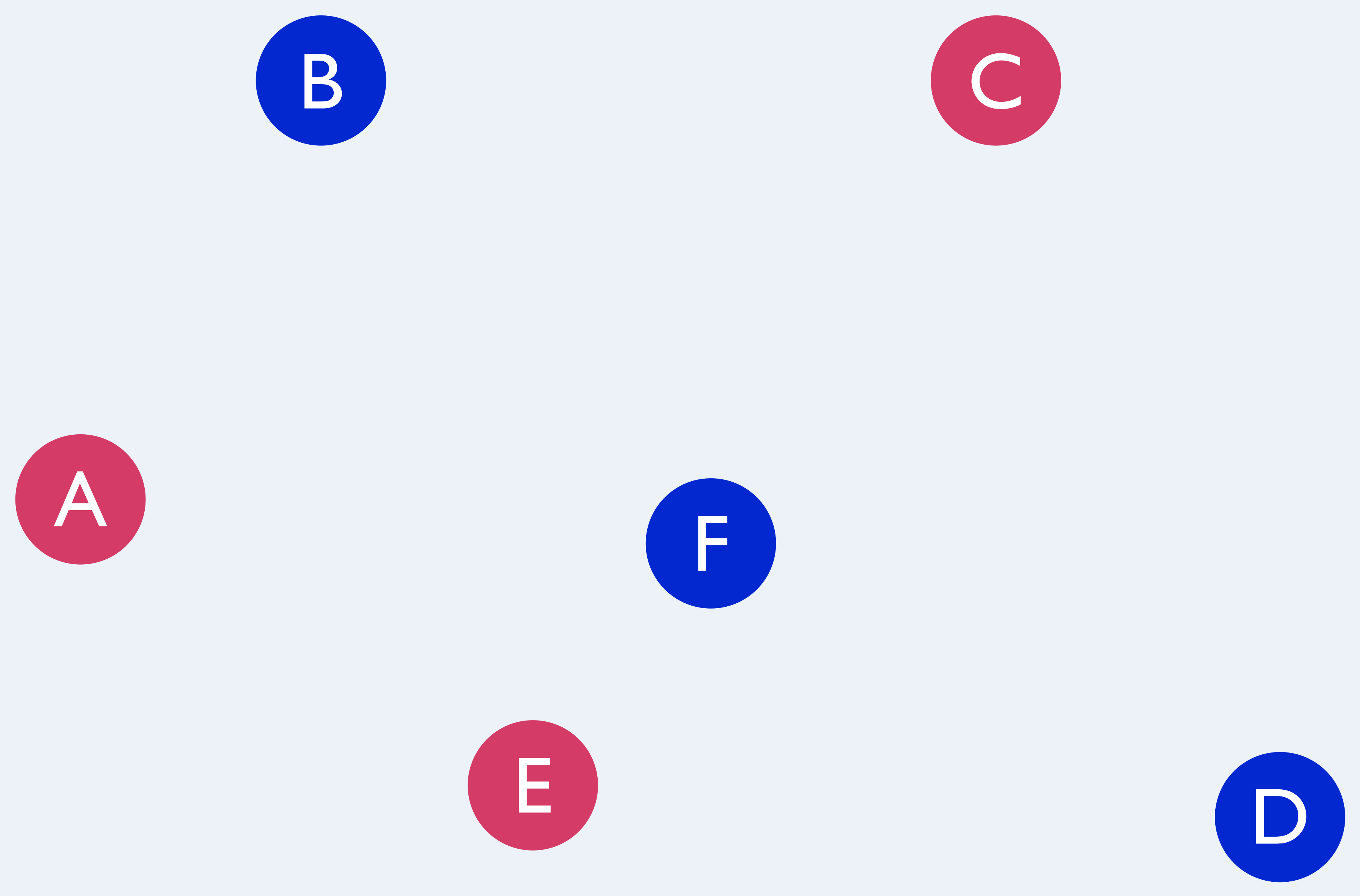
Inhomogeneous Random K-out Graphs

$\mathbb{H}(n, \mu, K_n)$

$\boldsymbol{n}$: number of nodes

Label nodes independently as
Type-I wp $\boldsymbol{\mu}$ (>0), Type-II wp $\mathbf{1}$-$\boldsymbol{\mu}$

Type-I nodes select $\mathbf{1}$ node,
Type-II nodes select K_n (≥2) nodes
(uniformly at random from all n-1 nodes)



Inhomogeneous K-out Random graph (n = 6, $\boldsymbol{K_n}$ = 3)

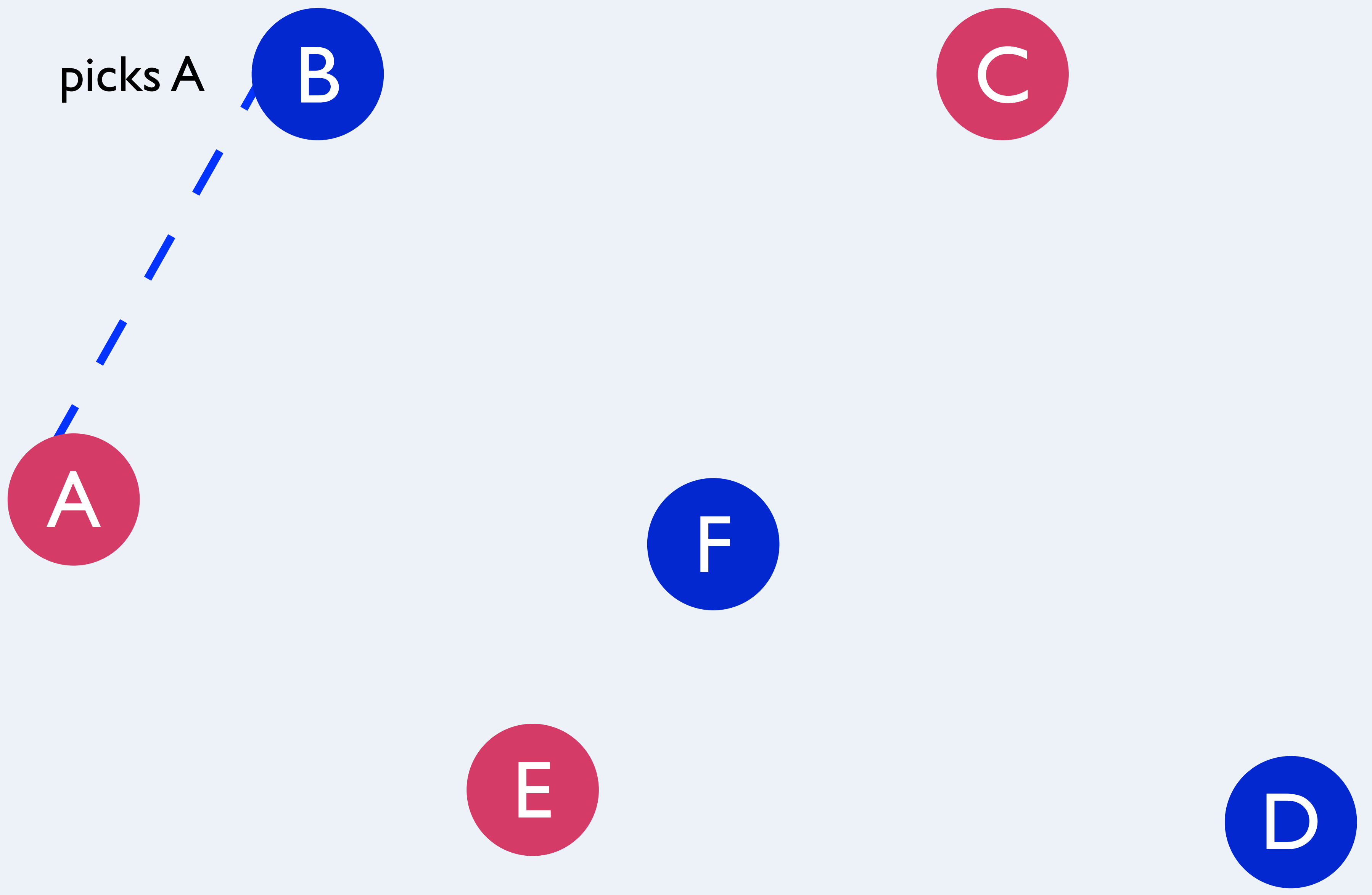
Inhomogeneous Random K-out Graphs

$\mathbb{H}(n, \mu, K_n)$

$\boldsymbol{n}$: number of nodes

Label nodes independently as
Type-I wp $\boldsymbol{\mu}$ (>0), Type-II wp $\mathbf{1}$-$\boldsymbol{\mu}$

Type-I nodes select $\mathbf{1}$ node,
Type-II nodes select K_n (≥2) nodes
(uniformly at random from all n-1 nodes)



Inhomogeneous K-out Random graph (n = 6, $\boldsymbol{K_n}$ = 3)

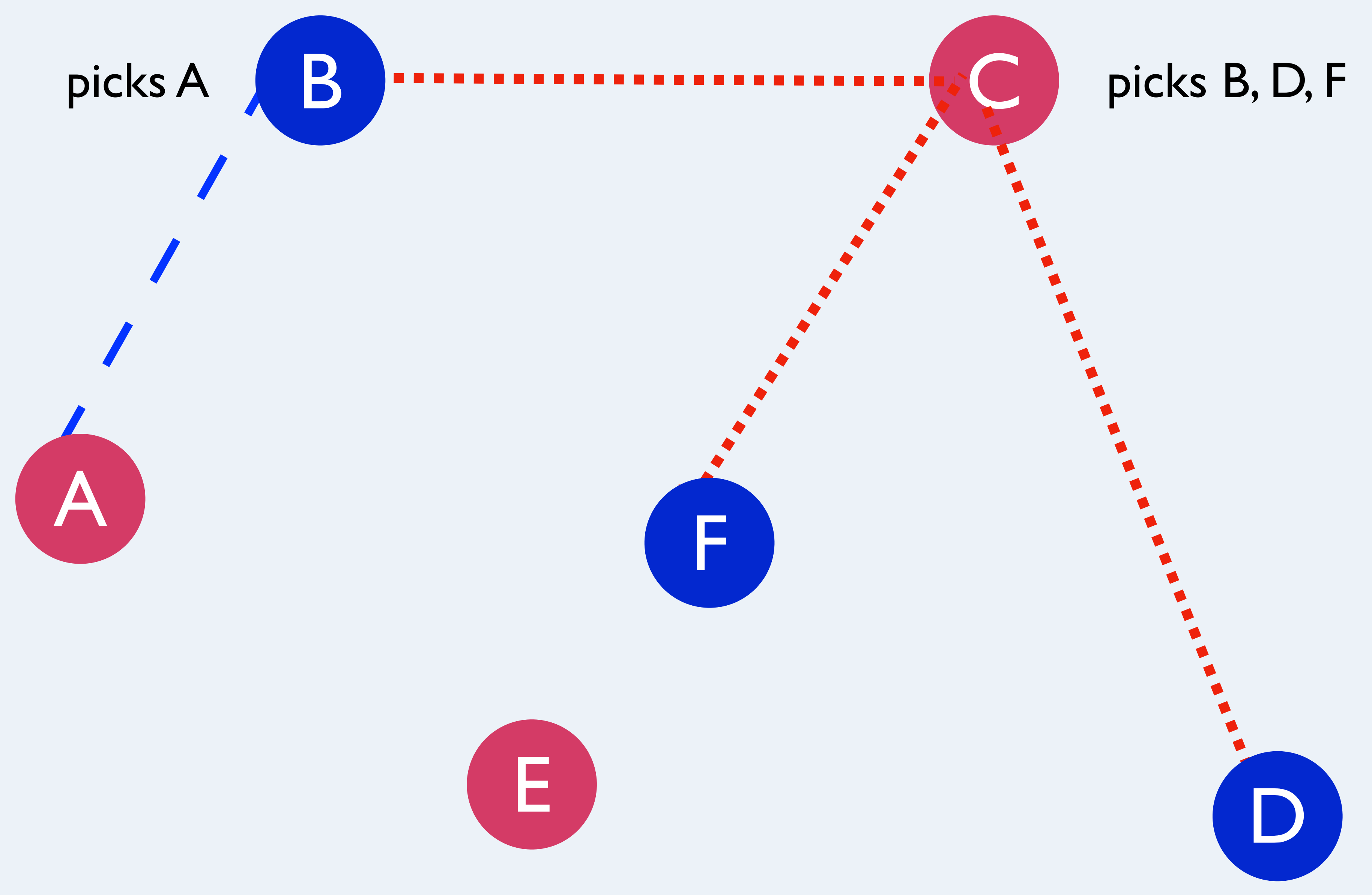
Inhomogeneous Random K-out Graphs

$\mathbb{H}(n, \mu, K_n)$

$\boldsymbol{n}$: number of nodes

Label nodes independently as
Type-I wp $\boldsymbol{\mu}$ (>0), Type-II wp $\mathbf{1}$-$\boldsymbol{\mu}$

Type-I nodes select $\mathbf{1}$ node,
Type-II nodes select K_n (≥2) nodes
(uniformly at random from all n-1 nodes)



Inhomogeneous K-out Random graph (n = 6, $\boldsymbol{K_n}$ = 3)

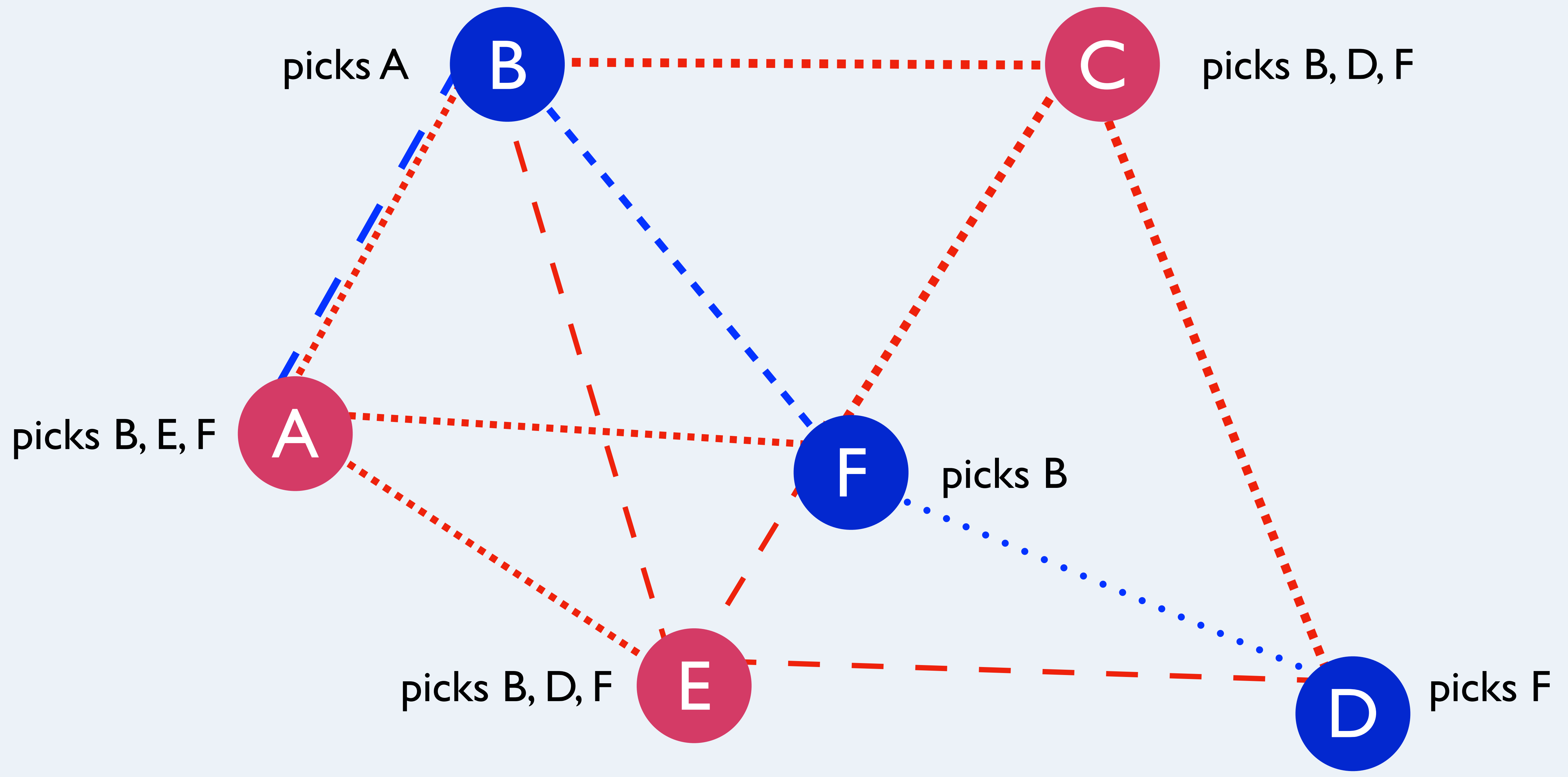
Inhomogeneous Random K-out Graphs

$\mathbb{H}(n, \mu, K_n)$

$\boldsymbol{n}$: number of nodes

Label nodes independently as
Type-I wp $\boldsymbol{\mu}$ (>0), Type-II wp $\mathbf{1}$-$\boldsymbol{\mu}$

Type-I nodes select $\mathbf{1}$ node,
Type-II nodes select K_n (≥2) nodes
(uniformly at random from all n-1 nodes)



Inhomogeneous K-out Random graph (n = 6, $\boldsymbol{K_n}$ = 3)

Inhomogeneous Random K-out Graphs

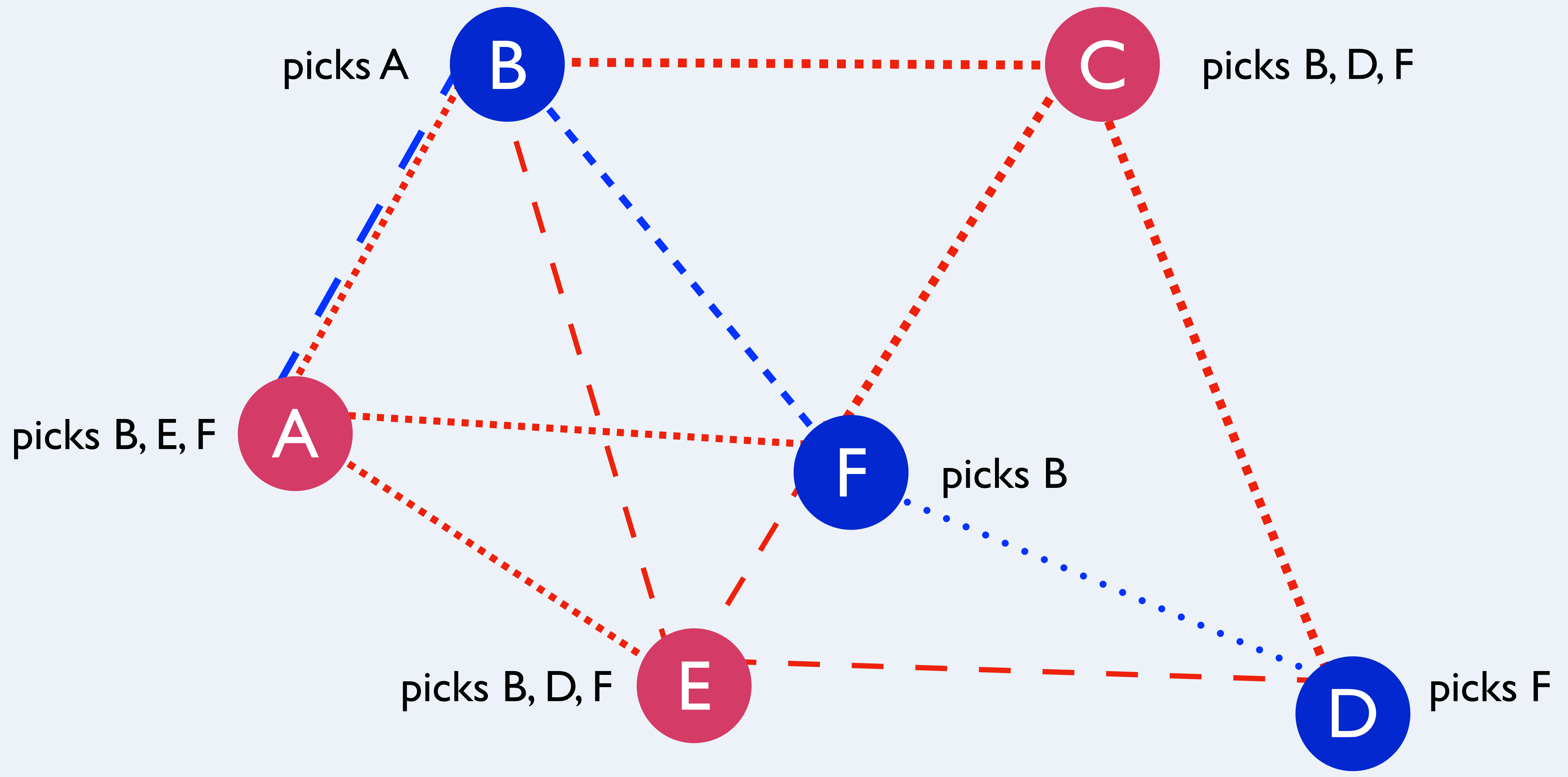
$\mathbb{H}(n, \mu, K_n)$

$\boldsymbol{n}$: number of nodes

Label nodes independently as
Type-I wp $\boldsymbol{\mu}$ (>0), Type-II wp $\mathbf{1}$-$\boldsymbol{\mu}$

Type-I nodes select $\mathbf{1}$ node,
Type-II nodes select K_n (≥2) nodes
(uniformly at random from all n-1 nodes)

Edge (i, j) exists if
node i selects node j
or
node j selects node i



Inhomogeneous K-out Random graph (n = 6, $\boldsymbol{K_n}$ = 3)

Inhomogeneous Random K-out Graphs

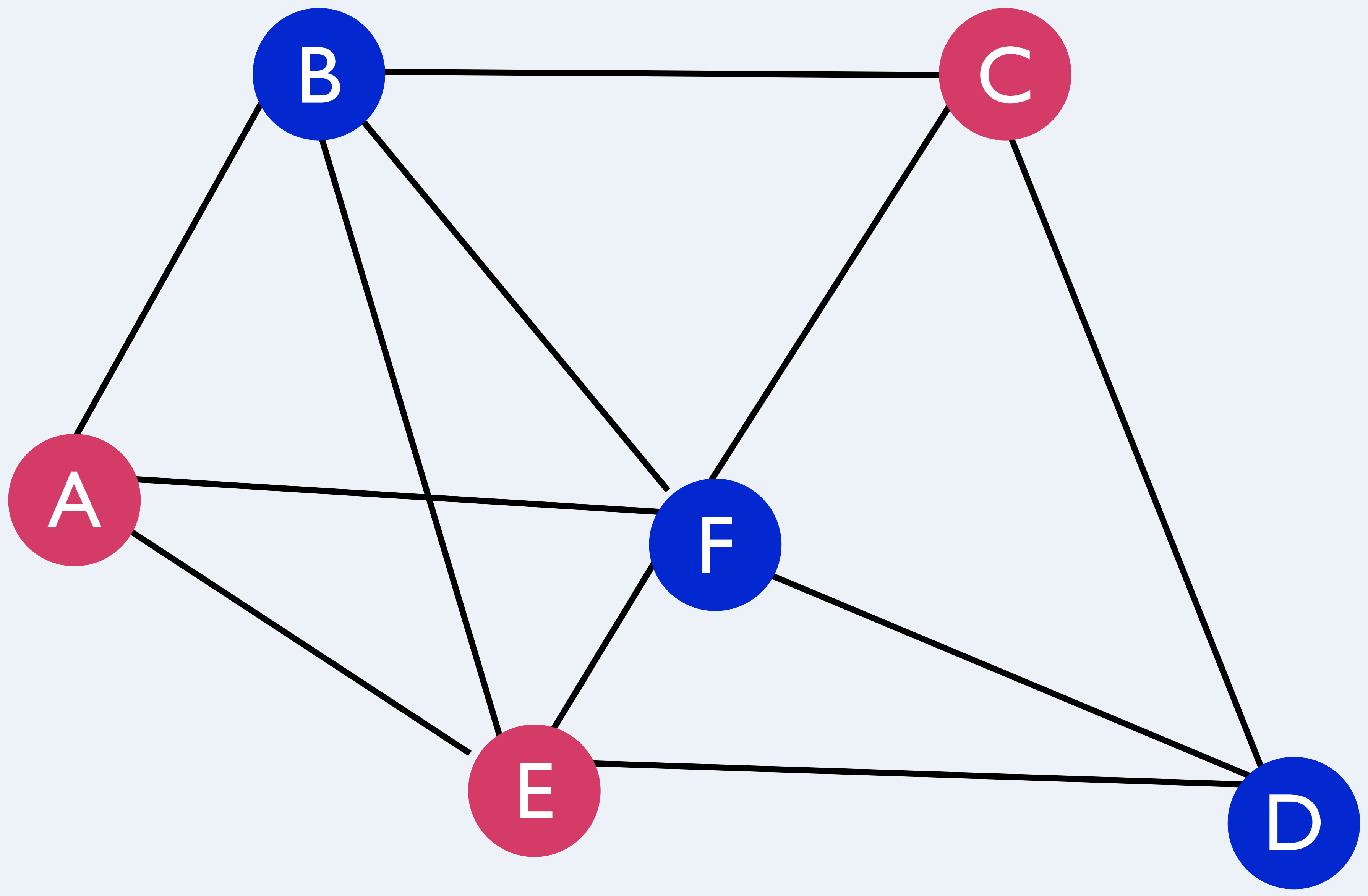
$\mathbb{H}(n, \mu, K_n)$

n: number of nodes

Label nodes independently as
Type-I wp μ (>0), Type-II wp $1-\mu$

Type-I nodes select 1 node,
Type-II nodes select K_n (≥2) nodes
(uniformly at random from all n-1 nodes)

Edge (i, j) exists if
node i selects node j
or
node j selects node i



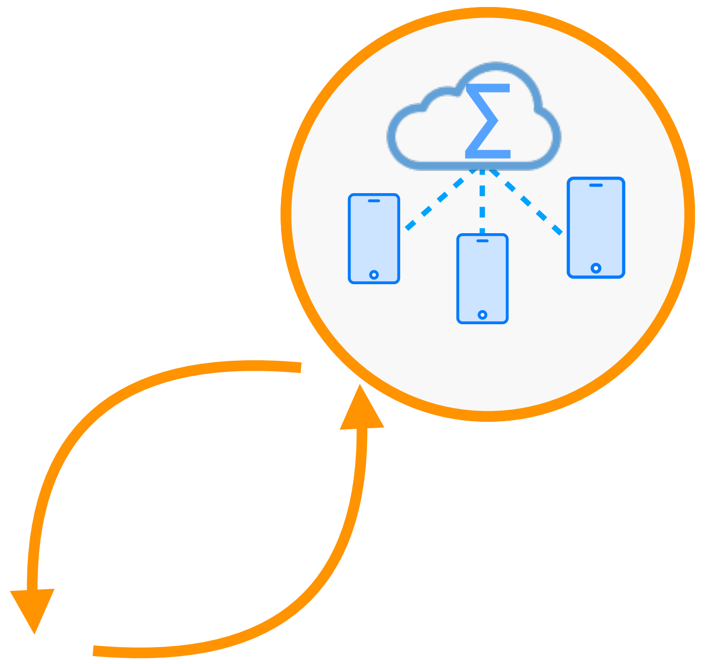
Inhomogeneous K-out Random graph (n = 6, K_n = 3)

Research Overview

Theme *How can we leverage network structure to better understand and design socio-technical systems?*

Thrusts

Network design and performance analysis for reliable inference in distributed systems



Contributions

Formal characterization of strength of connectivity of 'random K-out graphs'

IEEE Transactions on Information Theory '21, '23
IEEE ICC '21 (Best Paper Award)
ISIT '21, '20, CDC '20, Globecom '19

(Joint work with O. Yagan and E. C. Elumar)

How to quantify strength of connectivity?

Increasing *strength* of connectivity

a 'giant' component

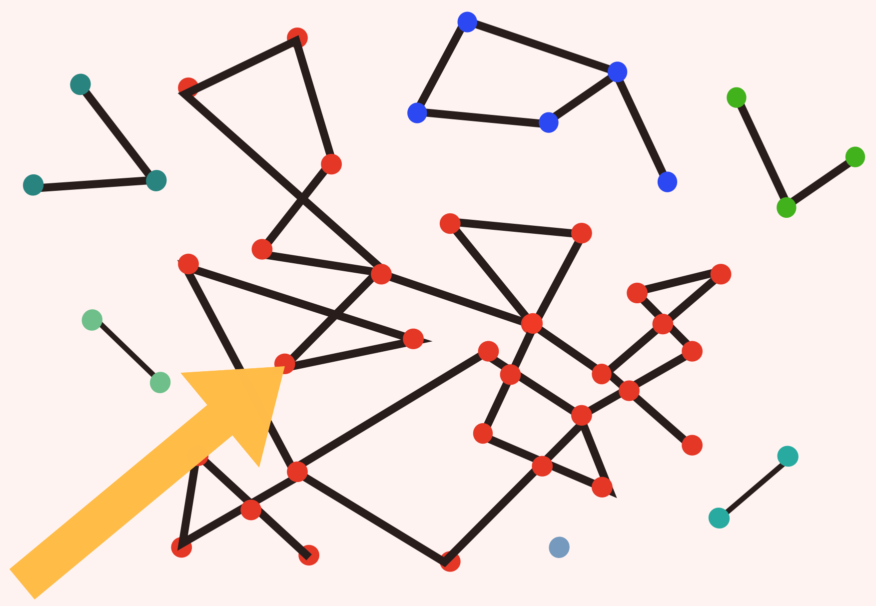
$|C_{max}| = \Omega(n)$

1-connectivity

$|C_{max}| = n$

***k*-connectivity**

ensures connectivity despite any k-1 node failures



C_{max}

Related work

Increasing *strength* of connectivity →

Weakest → Strongest

	a 'giant' component $\lvert C_{max} \rvert = \Omega(n)$	**1-connectivity** $\lvert C_{max} \rvert = n$	***k*-connectivity** $(k \geq 2)$
Inhomogeneous Random *K*-out Graphs $\mathbb{H}(n, \mu, K_n)$	?	$K_n = \omega(1)$ [Eletreby & Yagan, '19]	?
Homogeneous Random *K*-out Graphs $\mathbb{H}(n, K_n)$	?	$K_n \geq 2$ [Fenner & Frieze '82] ? Tighter bounds? what if a random subset of nodes fail?	$K_n \geq 2k$ [Fenner & Frieze '82]

Key Contributions

Increasing *strength* of connectivity →

Weakest — Strongest

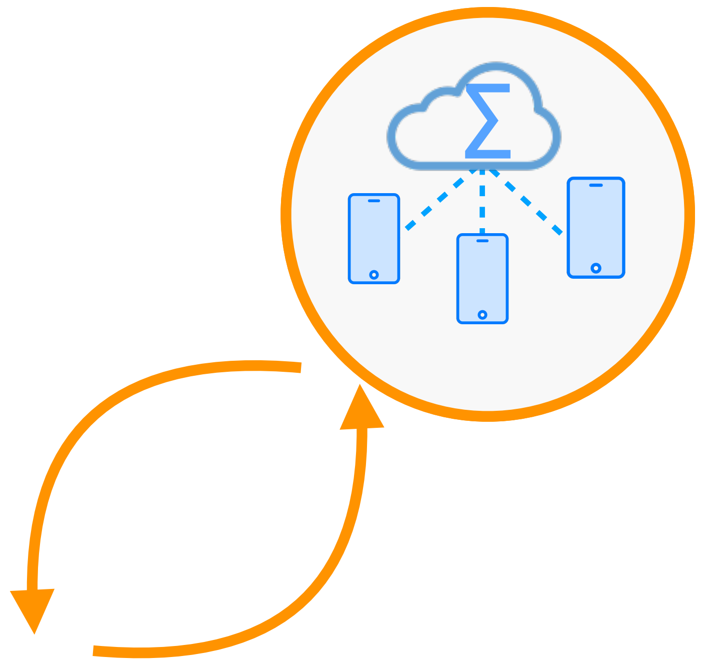
	a 'giant' component $\lvert C_{\max} \rvert = \Omega(n)$	**1-connectivity** $\lvert C_{\max} \rvert = n$	***k*-connectivity** $(k \geq 2)$
Inhomogeneous Random K-out Graphs $\mathbb{H}(n, \mu, K_n)$	✓	$K_n = \omega(1)$ [Eletreby & Yagan, '19]	✓
Homogeneous Random K-out Graphs $\mathbb{H}(n, K_n)$	✓	$K_n \geq 2$ [Fenner & Frieze '82] Tighter bounds? ✓ what if a random subset of nodes fail?	$K_n \geq 2k$ [Fenner & Frieze '82]

Research Overview

Theme *How can we leverage network structure to better understand and design socio-technical systems?*

Thrusts

Network design and performance analysis for reliable inference in distributed systems



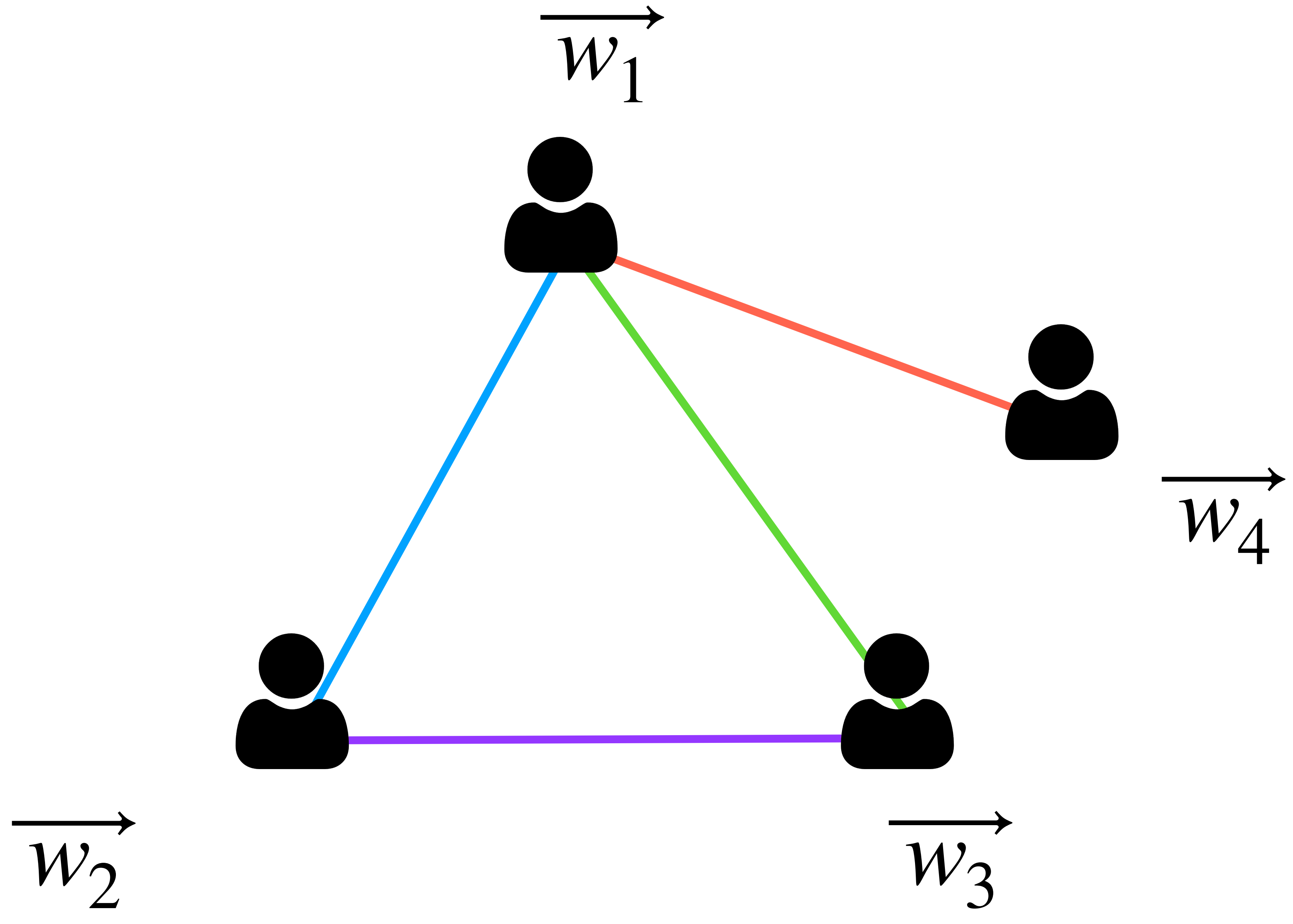
Contributions

Formal characterization of strength of connectivity of 'random K-out graphs'

(Joint work with O. Yagan and E. C. Elumar)

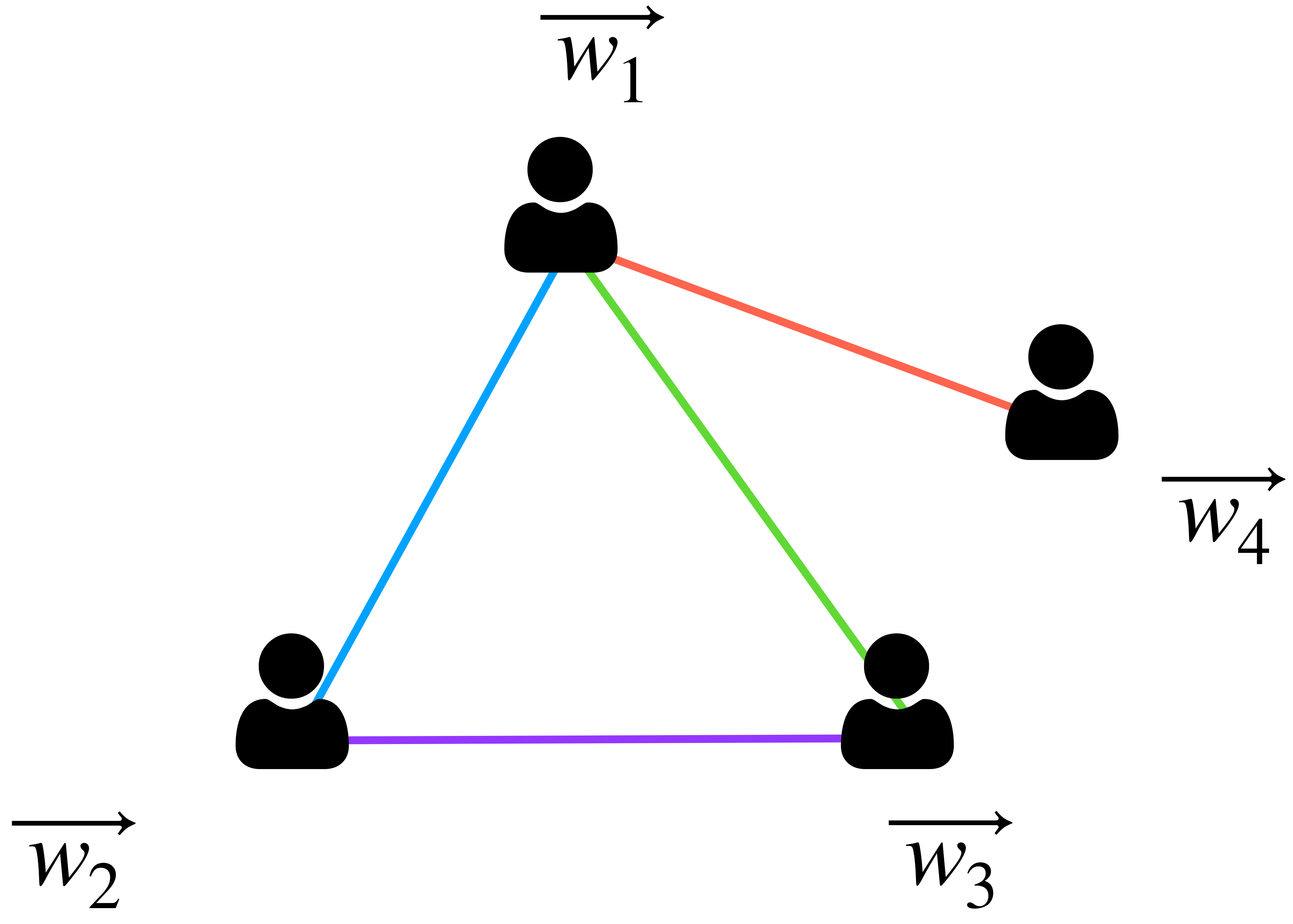
K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$
without revealing w_i



K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$
without revealing w_i

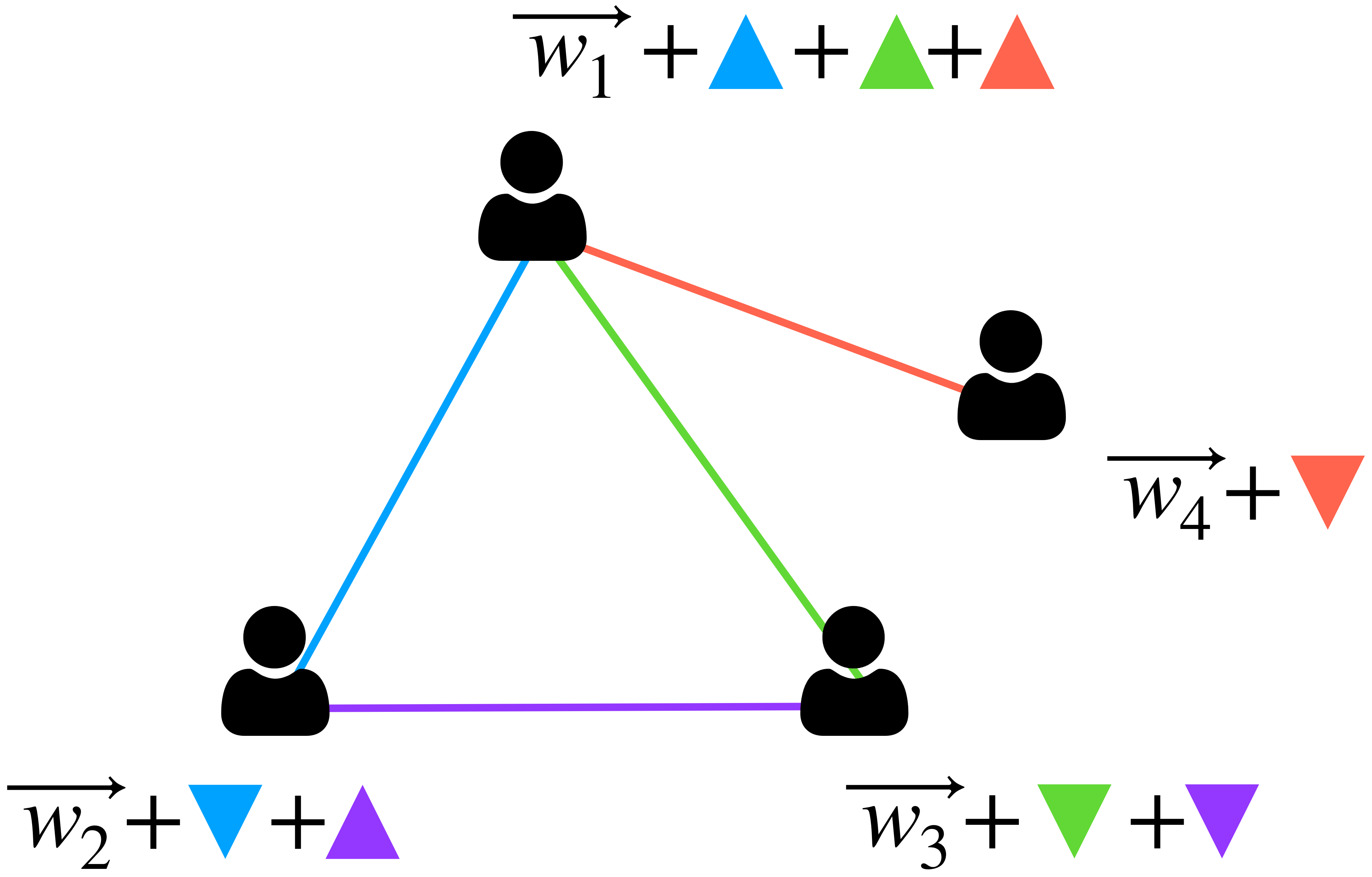


Approach: add pairwise masks
that cancel in aggregate

[Sabater et al. '20], [Bell et al. '20], ...

K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$ without revealing w_i

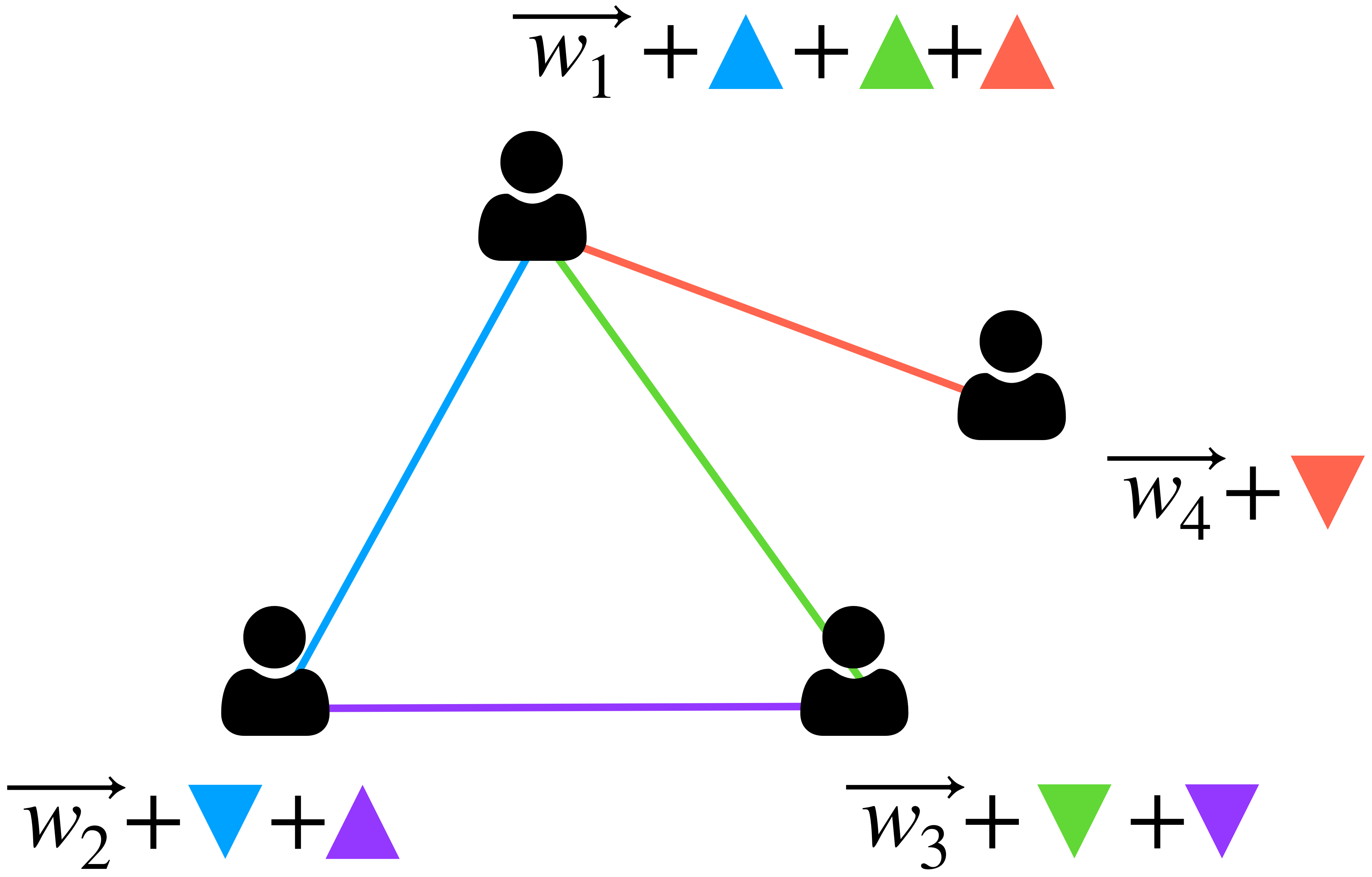


Approach: add pairwise masks that cancel in aggregate

[Sabater et al. '20], [Bell et al. '20], ...

K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$
without revealing w_i



Approach: add pairwise masks
that cancel in aggregate

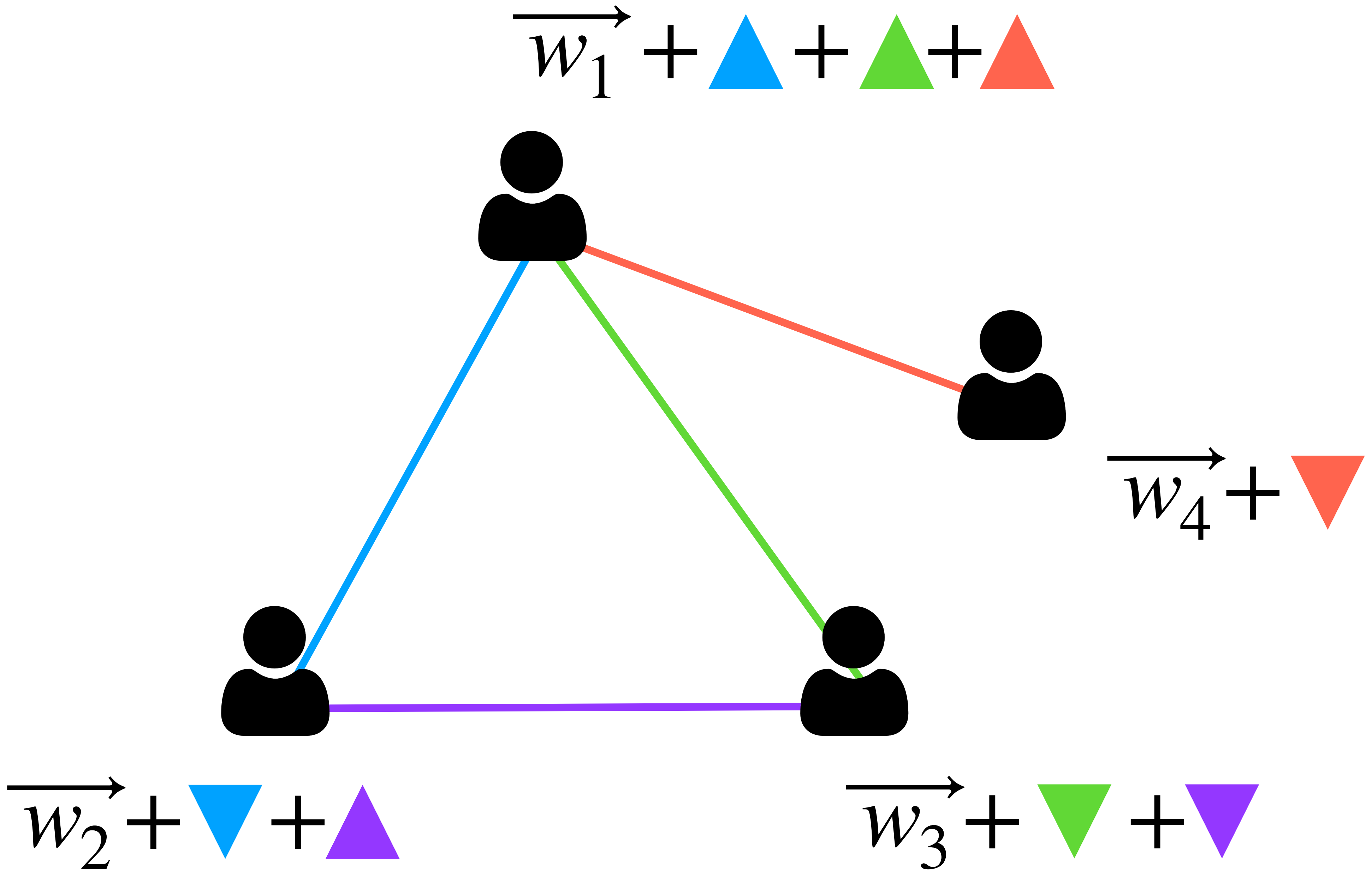
[Sabater et al. '20], [Bell et al. '20], ...

Performance trade-off:

better connectivity $\Longrightarrow$ better masking
but
raises communication costs

K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$
without revealing w_i



Approach: add pairwise masks that cancel in aggregate

[Sabater et al. '20], [Bell et al. '20], …

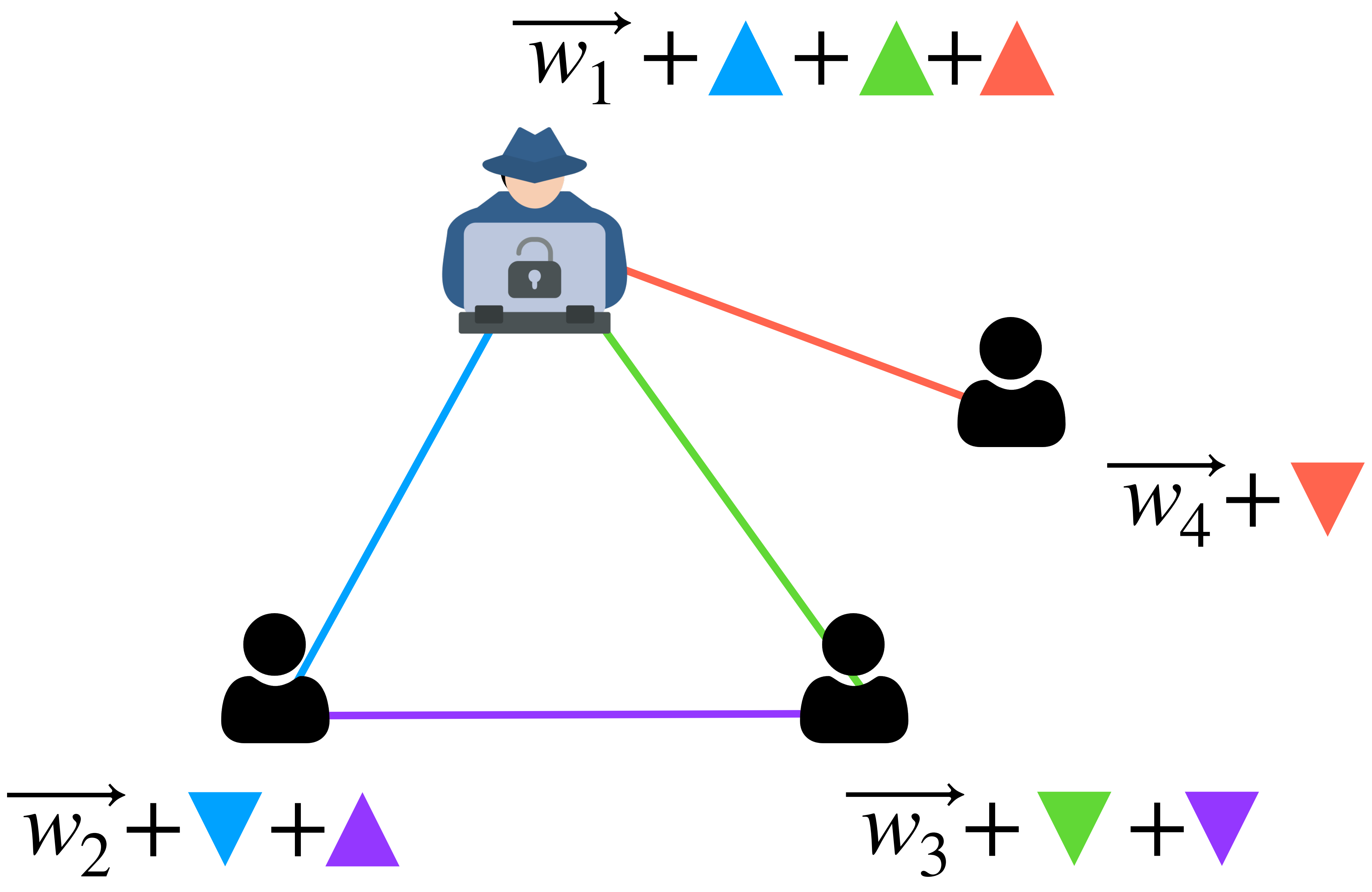
Performance trade-off:

better connectivity $\Longrightarrow$ better masking but raises communication costs

Random K-out graphs have been proposed to balance sparsity with connectivity

K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$ without revealing w_i



Approach: add pairwise masks that cancel in aggregate

[Sabater et al. '20], [Bell et al. '20], …

Performance trade-off:

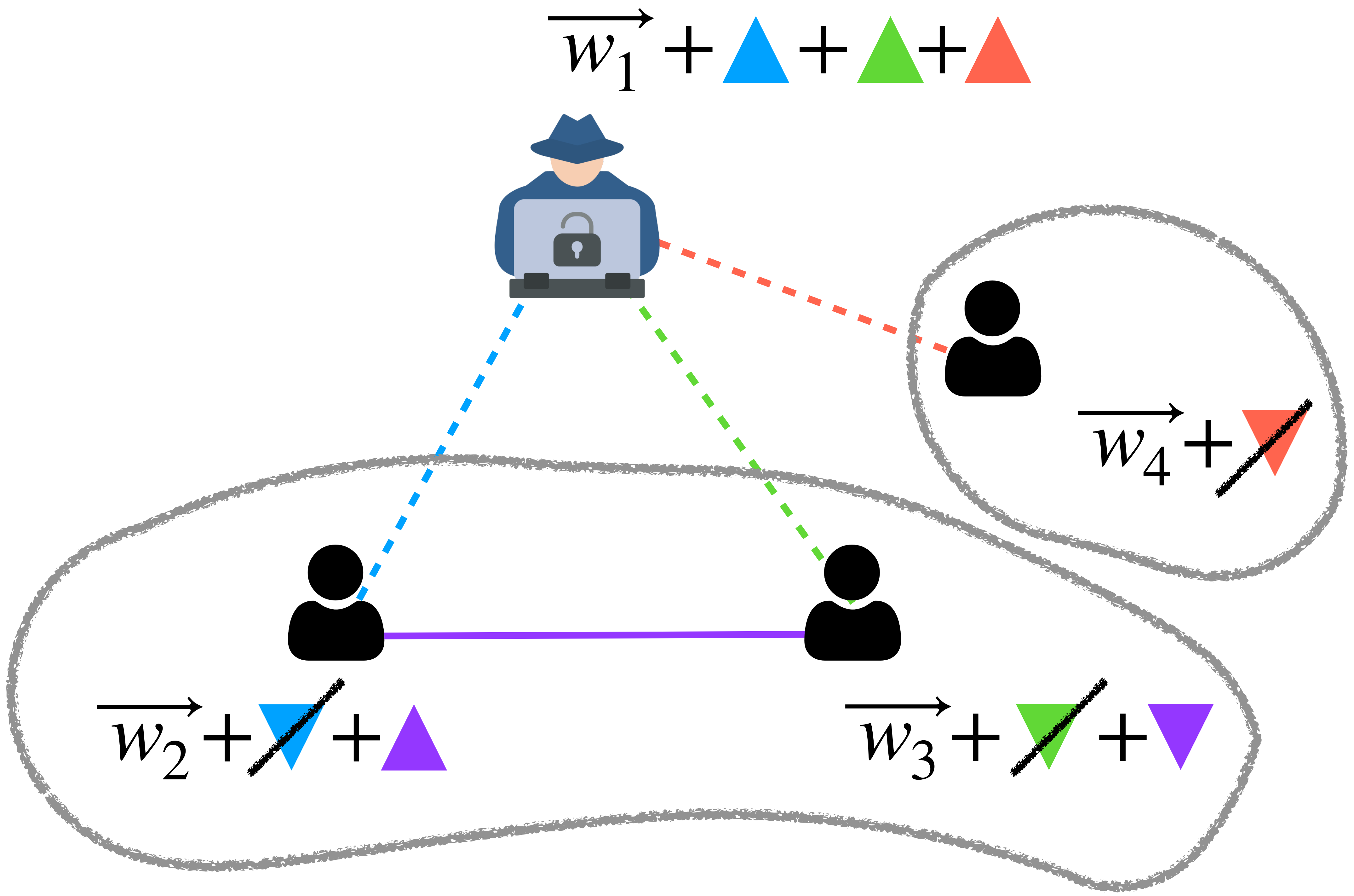
better connectivity $\Longrightarrow$ better masking but raises communication costs

Random K-out graphs have been proposed to balance sparsity with connectivity

What if there are multiple corrupt nodes?

K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$
without revealing w_i



Approach: add pairwise masks
that cancel in aggregate

[Sabater et al. '20], [Bell et al. '20], …

Performance trade-off:

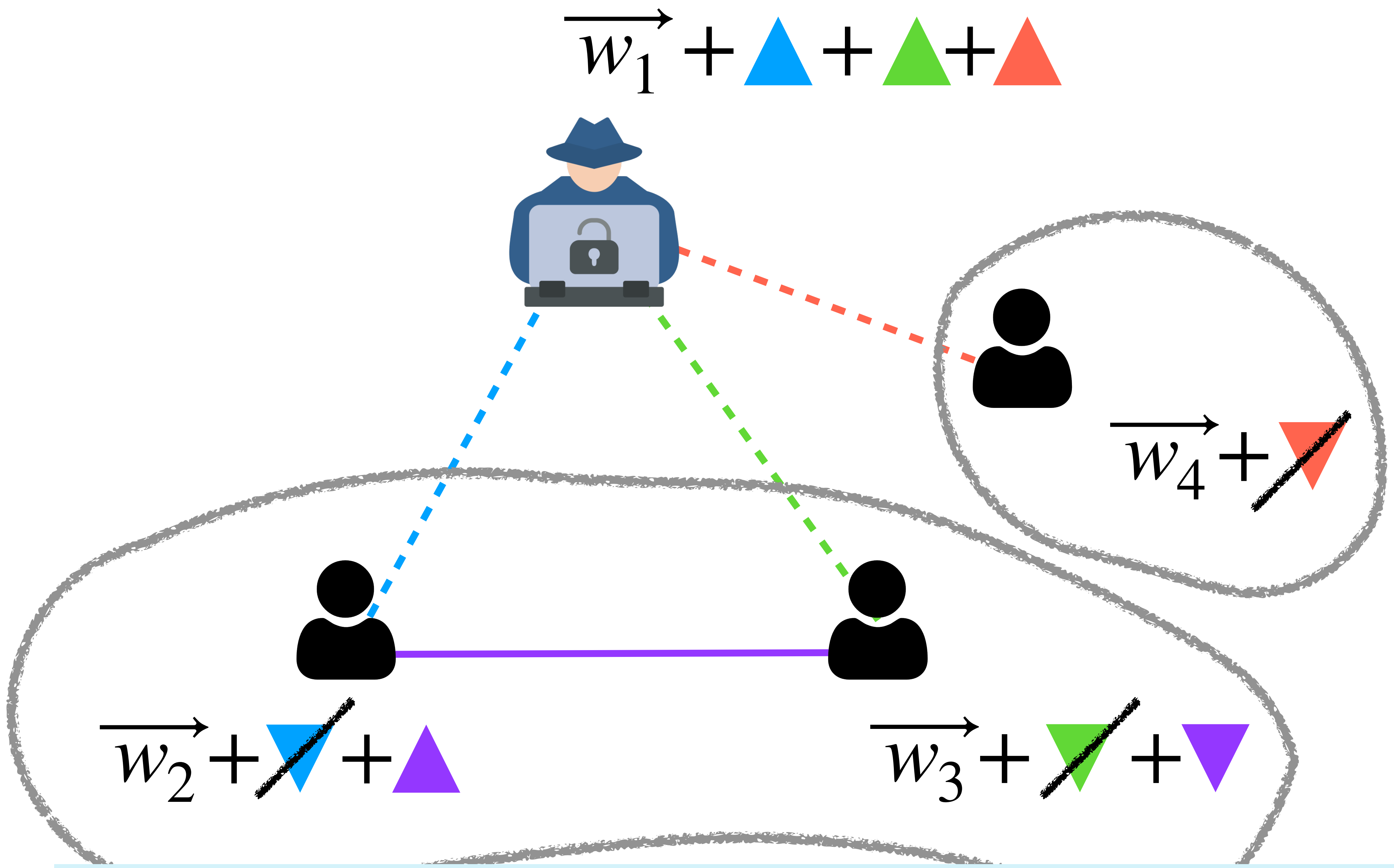
better connectivity $\implies$ better masking
but
raises communication costs

Random K-out graphs have been proposed
to balance sparsity with connectivity

What if there are multiple corrupt nodes?

K-out graphs in action: Distributed pairwise masking

Setting: compute $\Sigma_i w_i$
without revealing w_i



Approach: add pairwise masks
that cancel in aggregate

[Sabater et al. '20], [Bell et al. '20], ...

Performance trade-off:

better connectivity $\Longrightarrow$ better masking
but
raises communication costs

Random K-out graphs have been proposed
to balance sparsity with connectivity

What if there are multiple corrupt nodes?
Is the subgraph of honest nodes connected?
Can we charaterize the size of connected subgraphs of honest nodes?
....

Our results in action: Distributed pairwise masking

What if there are multiple corrupt nodes?

How to select K_n to ensure privacy properties for the subgraph of honest nodes?

Suppose δ_n nodes chosen uniformly at random from $\mathbb{H}(n, K_n)$ are corrupt

Let $\mathbb{S}(n, K_n, \delta_n)$ denote the subgraph of honest nodes

Our results in action: Distributed pairwise masking

What if there are multiple corrupt nodes?

How to select K_n to ensure privacy properties for the subgraph of honest nodes?

Suppose δ_n nodes chosen uniformly at random from $\mathbb{H}(n, K_n)$ are corrupt

Let $\mathbb{S}(n, K_n, \delta_n)$ denote the subgraph of honest nodes

How to select K_n to ensure as δ_n varies that:

- $\mathbb{S}(n, K_n, \delta_n)$ is connected whp?
- $|C_{\max}(\mathbb{S}(n, K_n, \delta_n))| \geq T_n$ whp?

Our results in action: Distributed pairwise masking

What if there are multiple corrupt nodes?

How to select K_n to ensure privacy properties for the subgraph of honest nodes?

Suppose δ_n nodes chosen uniformly at random from $\mathbb{H}(n, K_n)$ are corrupt

Let $\mathbb{S}(n, K_n, \delta_n)$ denote the subgraph of honest nodes

How to select K_n to ensure as δ_n varies that: $o(n), \omega(1)$

- $\mathbb{S}(n, K_n, \delta_n)$ is connected whp? $K_n = \Omega(\log(\delta_n))$
- $|C_{\max}(\mathbb{S}(n, K_n, \delta_n))| \geq T_n$ whp? $K_n \geq 2$

T_n: $n(1 - o(1))$

Key Contributions

Increasing *strength* of connectivity

Weakest — Strongest

a 'giant' component

$|C_{max}| = \Omega(n)$

1-connectivity

$|C_{max}| = n$

***k*-connectivity**

$(k \geq 2)$

Homogeneous Random K-out Graphs

$\mathbb{H}(n, K_n)$

Provide K_n required to ensure a given $|C_{max}|$ whp as a function of size of random node failures

$K_n \geq 2$ [Fenner & Frieze '82]

$p_{con} = 1 - \Theta(1/n^{K^2-1}),\ K \geq 2$

$p_{con} \to 1$ even after $o(\sqrt{n})$ nodes fail

(additional results for other failure regimes)

$K_n \geq 2k$

[Fenner & Frieze '82]

Key Contributions

Increasing *strength* of connectivity →

Weakest → Strongest

	a 'giant' component $\lvert C_{max} \rvert = \Omega(n)$	**1-connectivity** $\lvert C_{max} \rvert = n$	***k*-connectivity** $(k \geq 2)$
Inhomogeneous Random K-out Graphs $\mathbb{H}(n, \mu, K_n)$	For any $K_n \geq 2$ (whp) $\lvert C_{max} \rvert = n - O(1)$ (even after $O(1)$ nodes fail) $\lvert C_{max} \rvert = n(1 - o(1))$ after $o(n)$ nodes fail	$K_n = \omega(1)$ [Eletreby & Yagan, '19]	$K_n = \Omega(\log n)$
Homogeneous Random K-out Graphs $\mathbb{H}(n, K_n)$	Provide K_n required to ensure a given $\lvert C_{max} \rvert$ whp as a function of size of random node failures	$K_n \geq 2$ [Fenner & Frieze '82] $p_{con} = 1 - \Theta(1/n^{K^2-1}),\ K \geq 2$ $p_{con} \to 1$ even after $o(\sqrt{n})$ nodes fail (additional results for other failure regimes)	$K_n \geq 2k$ [Fenner & Frieze '82]

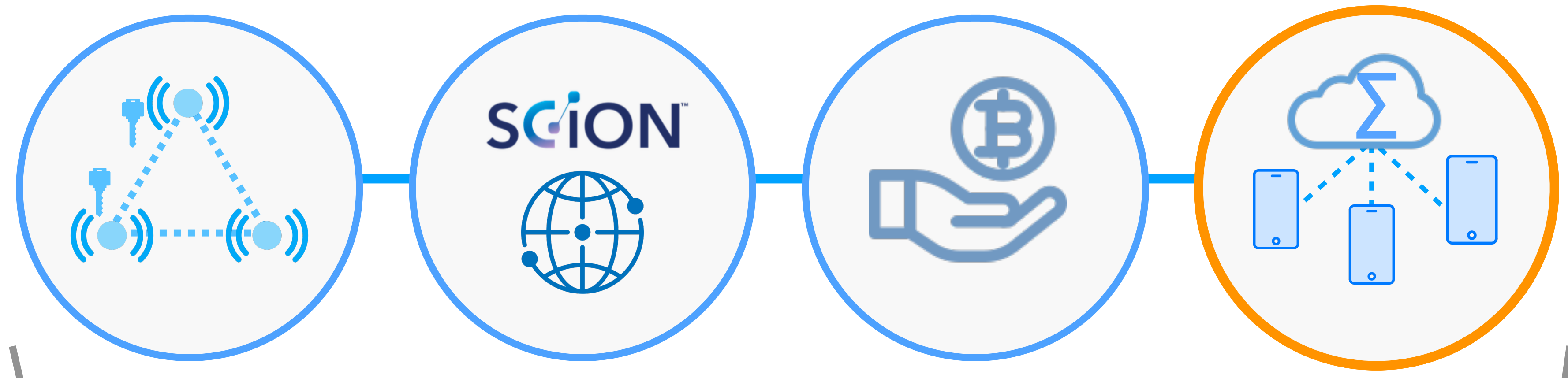
Research Overview

Theme *How can we leverage network structure to better understand and design socio-technical systems?*

Thrusts

Network design and performance analysis for reliable inference in distributed systems

Modeling, analyzing, and controlling spreading processes in social networks



Contributions

Formal characterization of strength of connectivity of 'random K-out graphs'

IEEE Transactions on Information Theory '21, '23
IEEE ICC '21 (Best Paper Award)
ISIT '21, '20, CDC '20, Globecom '19

(Today's focus)

(Joint work with O. Yagan and E. C. Elumar)

Analyzing spreading processes trigerred by evolving contagions

Proceedings of the National Academy of Sciences, '23
IEEE ICC '23

Research Overview

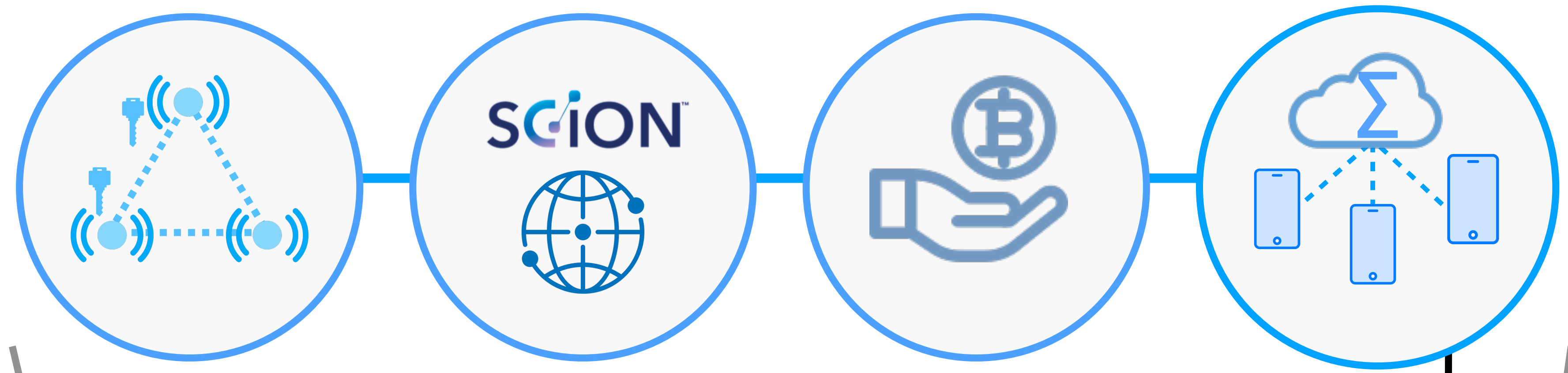
Theme

How can we leverage network structure to better understand and design socio-technical systems?

Thrusts

Network design and performance analysis for reliable inference in distributed systems

Modeling, analyzing, and controlling spreading processes in social networks



Contributions

Formal characterization of strength of connectivity of 'random K-out graphs'

IEEE Transactions on Information Theory '21, '23
IEEE ICC '21 (Best Paper Award)
ISIT '21, '20, CDC '20, Globecom '19

Analyzing spreading processes trigerred by evolving contagions

Proceedings of the National Academy of Sciences, '23
IEEE ICC '23

Ongoing, future work

Privacy-scalability frontiers in distributed & decentralized learning

Decentralized contentent moderation
Cross-platform interactions & information spread

Thanks

Mansi Sood

Carnegie Mellon University

msood@andrew.cmu.edu

https://www.mansisood.com/

References / Links

- TI Fenner, AM Frieze, **On the connectivity of random m-orientable graphs and digraphs**, Combinatorica 1982
- M Sood, O Yagan, **On the Minimum Node Degree and k-connectivity in Inhomogeneous Random K-out Graphs**, IEEE Transactions on Information Theory 2021
- M Sood, O Yagan, **Existence and Size of the Giant Component in Inhomogeneous Random K-out Graphs,** , IEEE Transactions on Information Theory 2023
- M Sood, O Yagan, **Tight Bounds for the Probability of Connectivity in Random K-out Graphs**, IEEE International Conference on Communications (ICC) 2021
- C Sabater, A Bellet, J Ramon, **An Accurate, Scalable and Verifiable Protocol for Federated Differentially Private Averaging**, Machine Learning 2022
- JH Bell, KA Bonawitz, A Gascón, T Lepoint, M Raykova, **Secure single-server aggregation with (poly) logarithmic overhead**, ACM SIGSAC Conference on Computer and Communications Security (CCS) 2020
- E Elumar, M Sood, O Yagan, **On the Connectivity and Giant Component Size of Random K-out Graphs Under Randomly Deleted Nodes**, IEEE International Symposium on Information Theory (ISIT) 2021
- M Sood, A Sridhar, R Eletreby, CW Wu, SA Levin, HV Poor, O Yagan, **Spreading Processes with Mutations over Multi-layer Networks**, Proceedings of the National Academy of Sciences (PNAS) 2023