Core percolation on complex networks

Yang-Yu Liu

Center for Complex Network Research, Northeastern University.

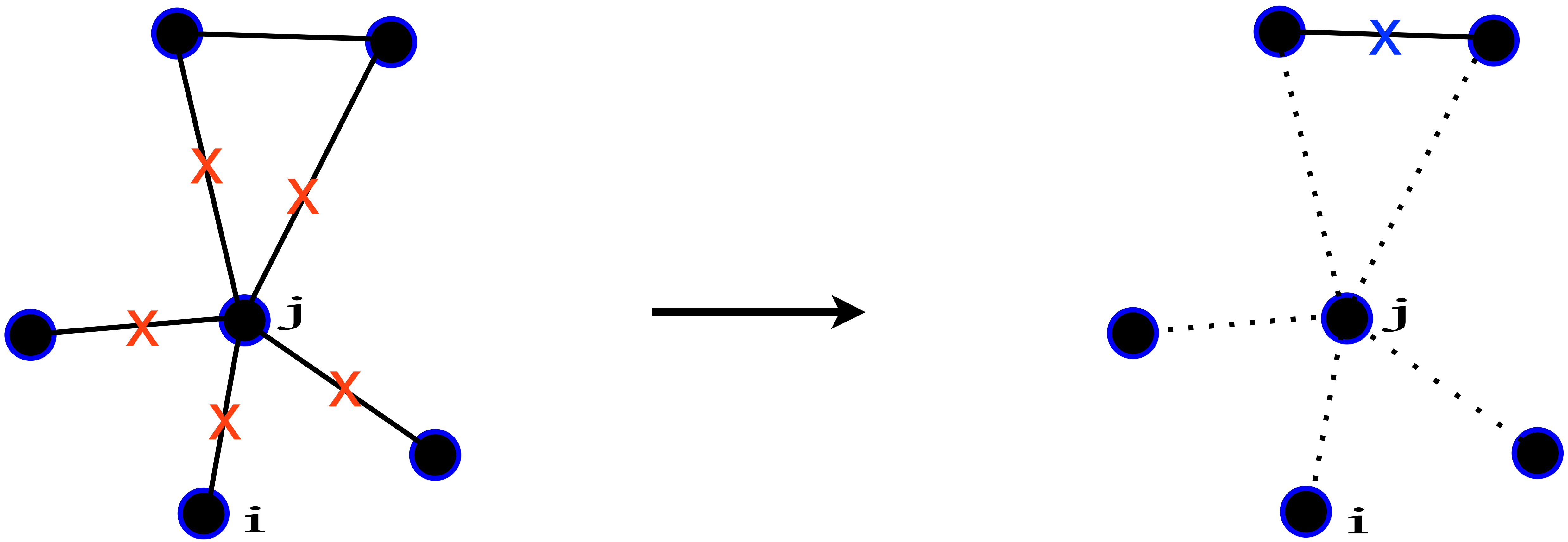
Center for Cancer Systems Biology, Dana-Farber Cancer Institute.

in collaboration with Endre Csóka, Haijun Zhou and Márton Pósfai

Phys. Rev. Lett. 109, 205703 (2012)

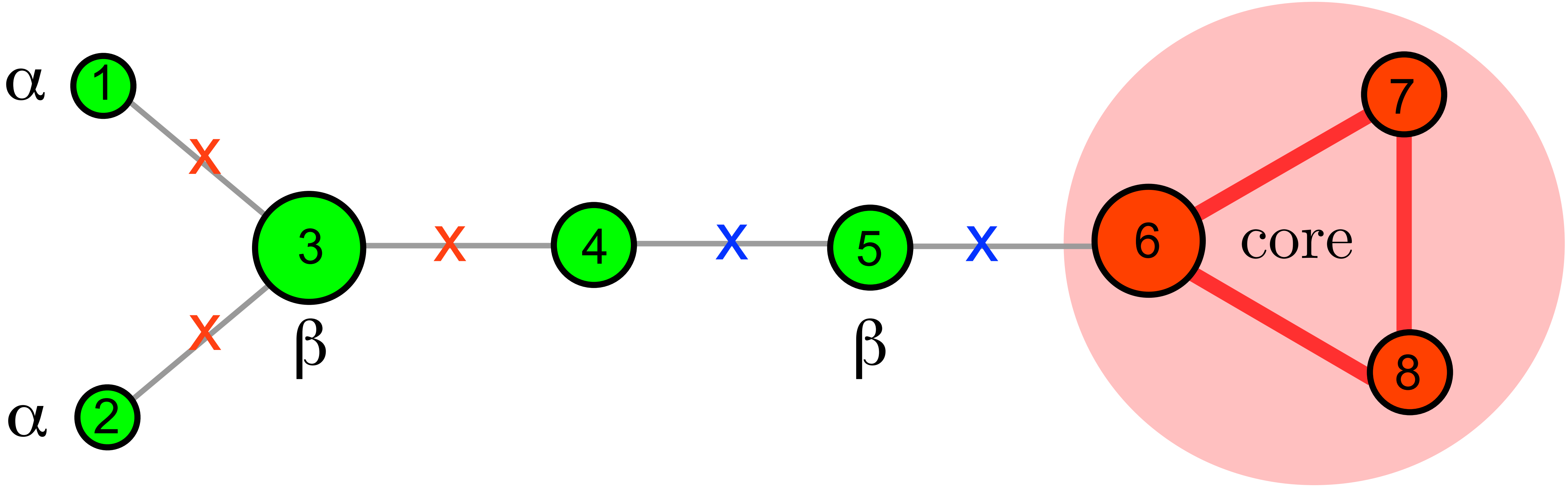
Greedy Leaf Removal

Iteratively remove leaf vertex *i*, its neighbor *j*, as well as all the edges incident with *j*.



Greedy Leaf Removal

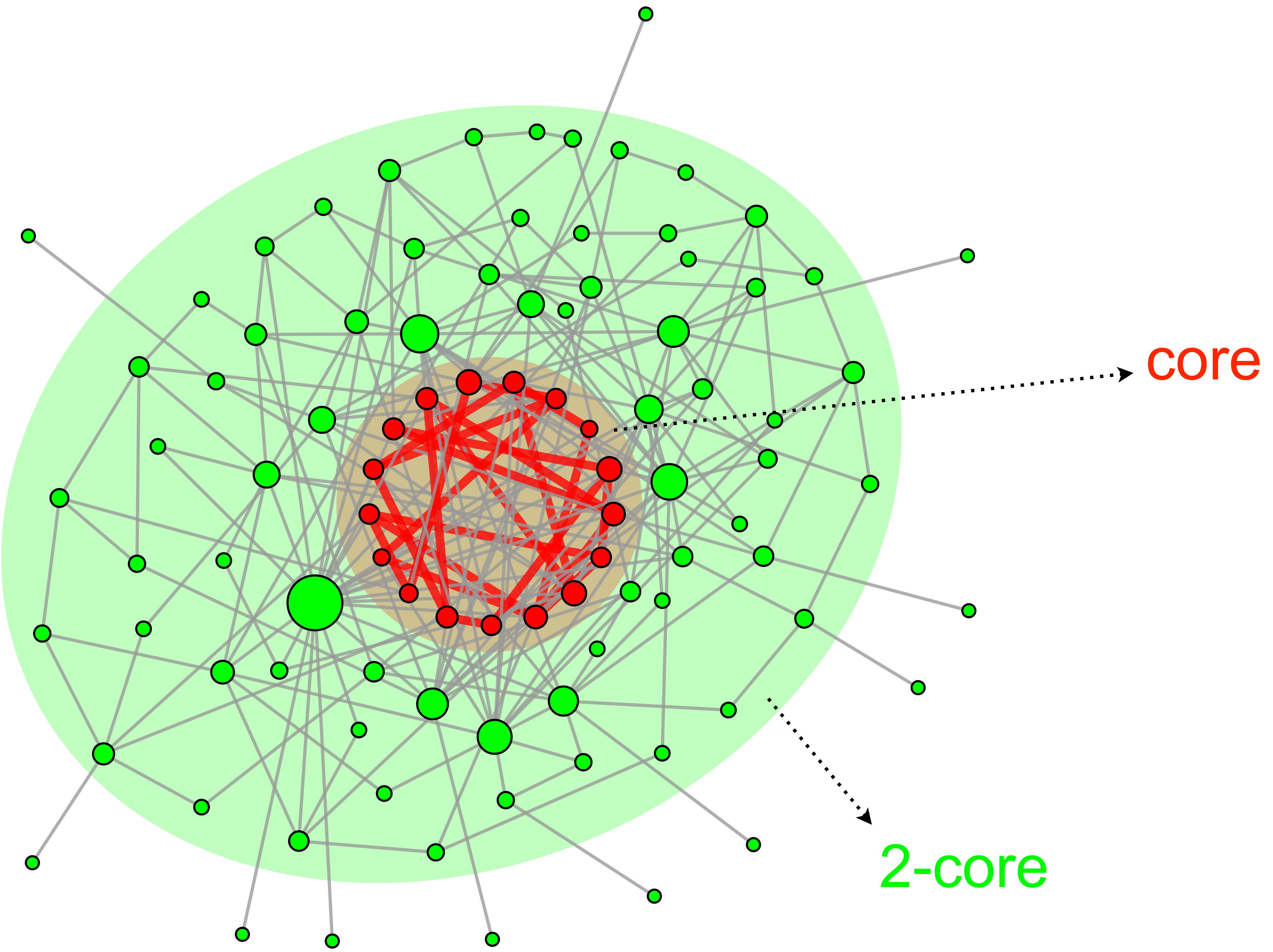
Iteratively remove leaf vertex *i*, its neighbor *j*, as well as all the edges incident with *j*.



The core does not depend on the history of the greedy leaf removal.

Bauer et al. *EPJB* (2001)

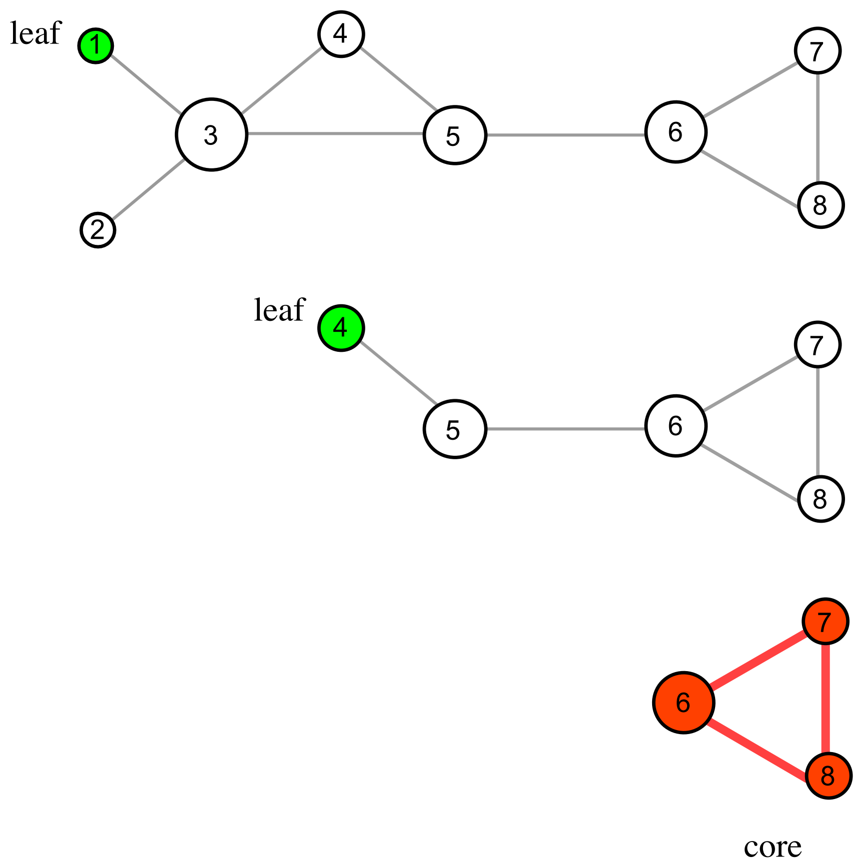
Core ≠ k-core

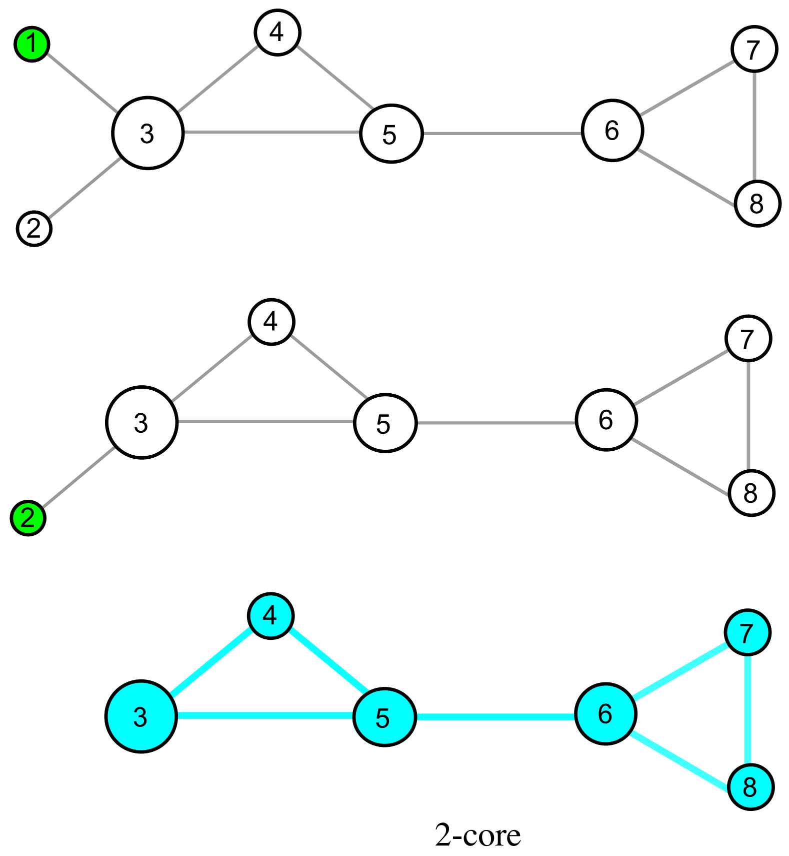


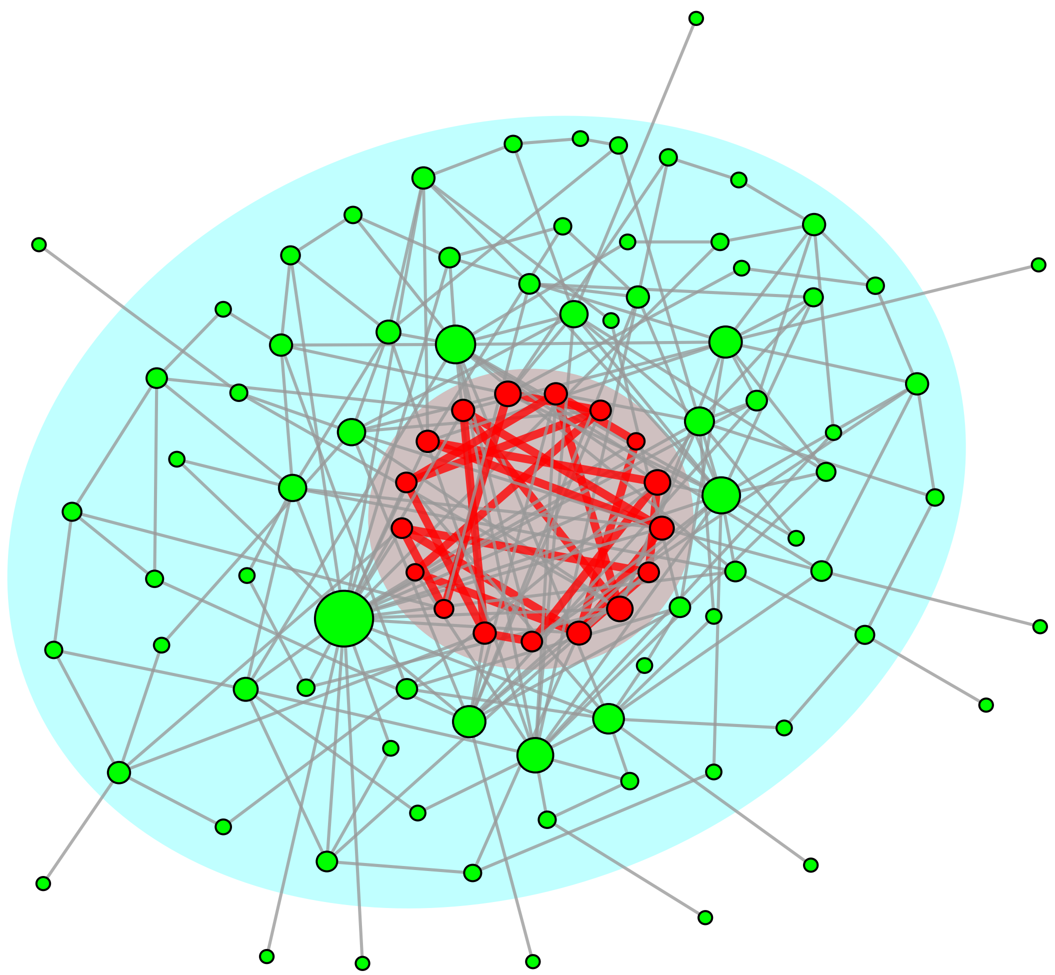
Core ≠ *k*-core

Nodes with degree k=1 and their neighbors are removed iteratively ==> **Core**

Nodes with degree less than k are removed iteratively ==> ***k*-core**



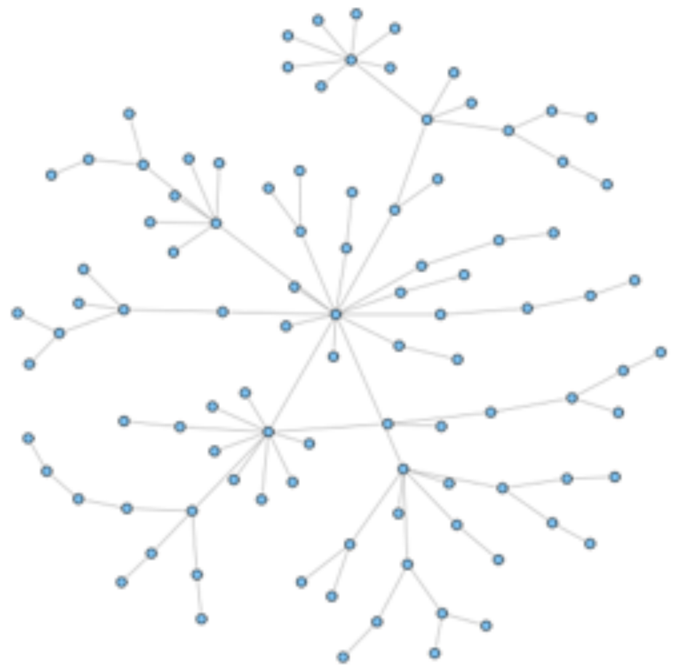




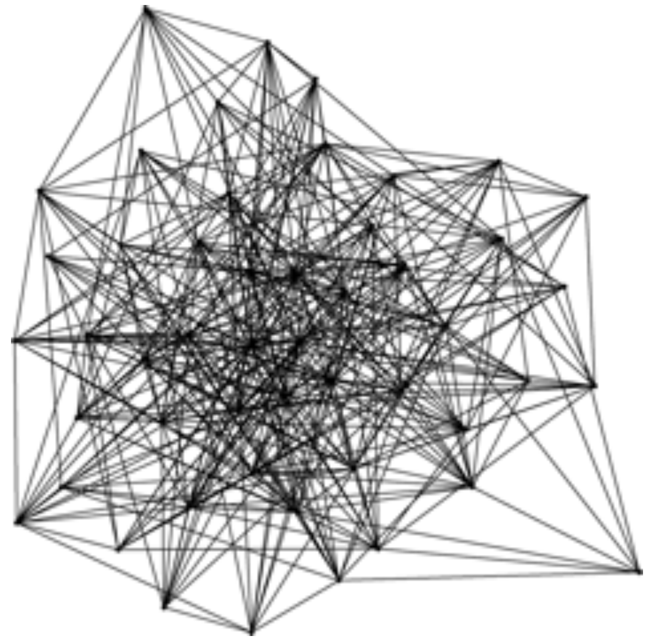
When does the core exist?

Two special cases

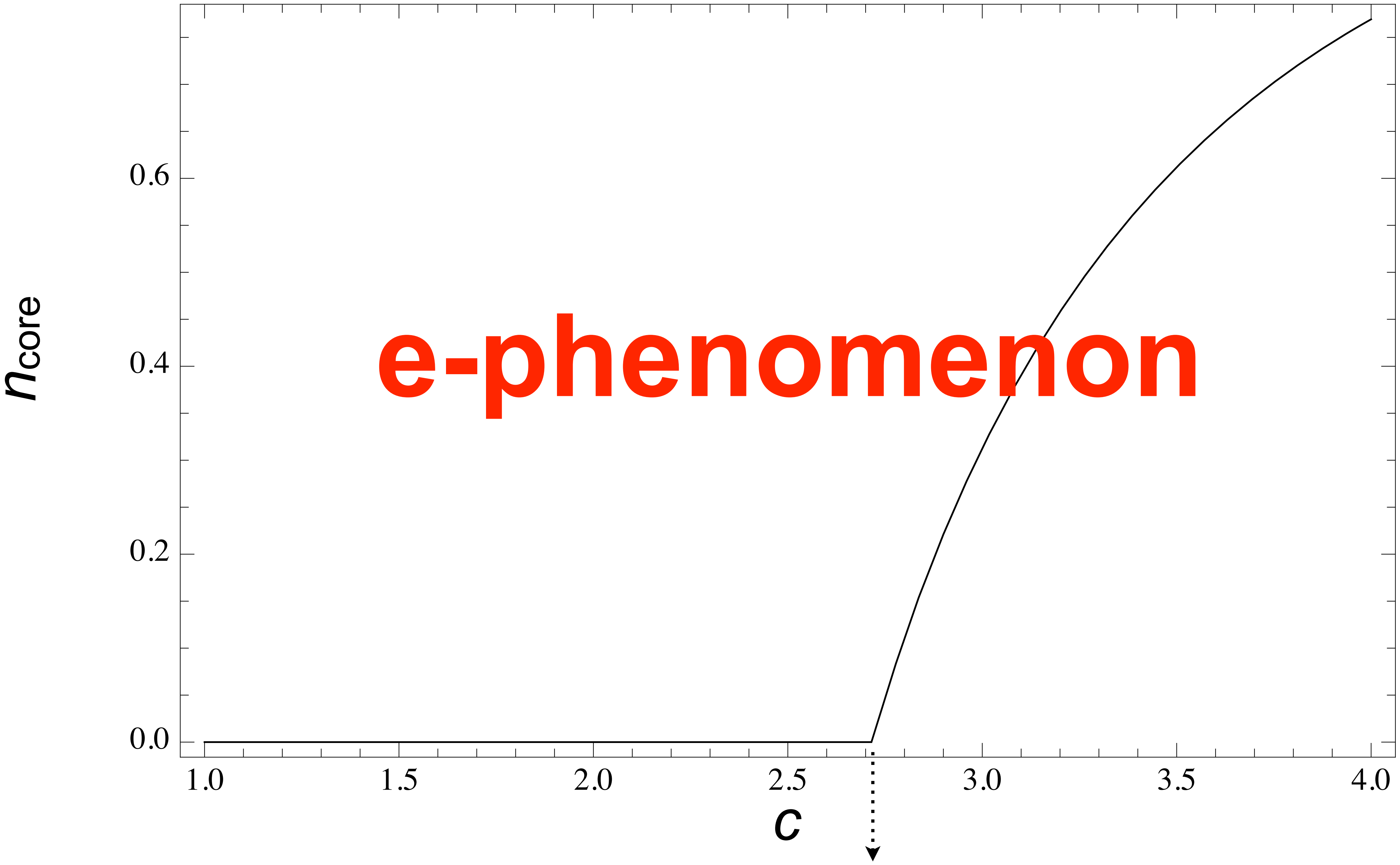
- If a network has no cycles, i.e. a tree or a forest, then no core, e.g. Barabási-Albert (BA) model with parameter m=1.



- If a network has no leaf nodes, then all nodes belong to the core, e.g. regular graphs with all nodes having the same degree k>1 or the BA model with m>1.



Core percolation on ER random graph



Core percolation threshold: $c^* = e = 2.7182818...$

Karp et al. Proc. IEEE (1981)
Bauer et al. *EPJB* (2001)

So what?

Motivation

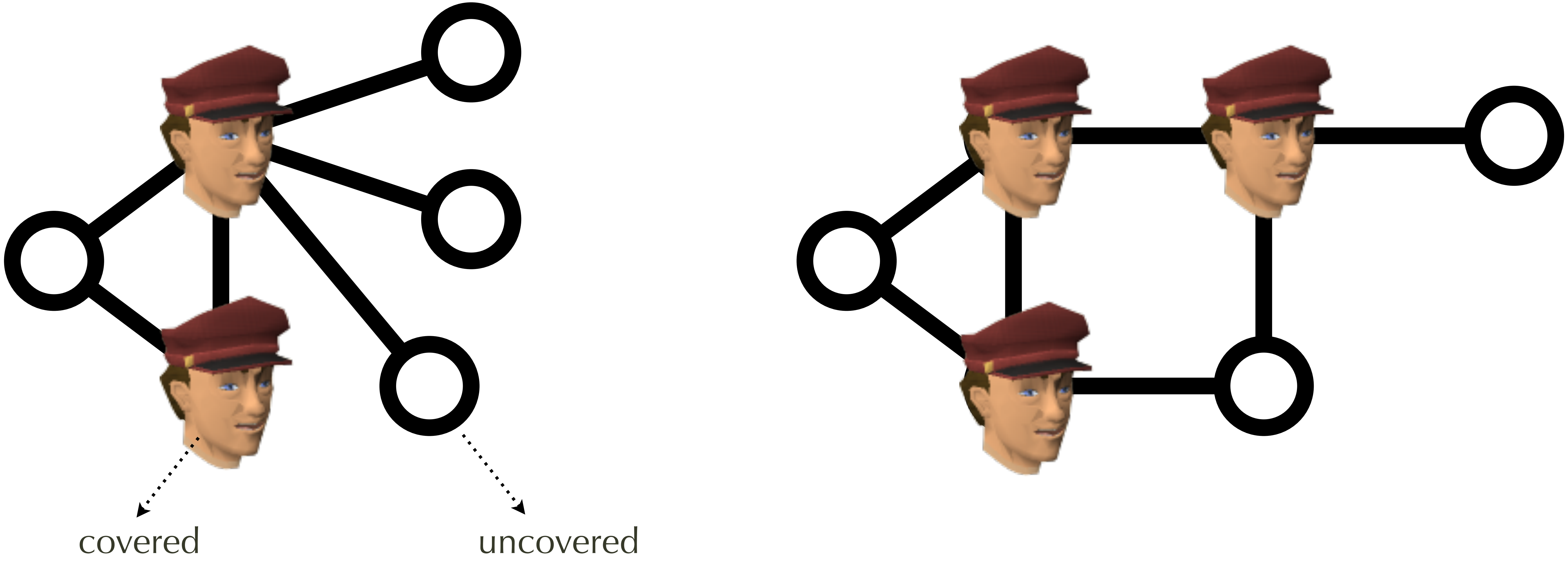
1. Minimum vertex cover (MVC) problem

- For ER, core percolation coincides with the changes of the solution-space structure of MVC problem --- **replica symmetry breaking**.
- For ER, if $c \leq e$ the typical running time of an algorithm for finding MVC is **polynomial**, while for $c > e$, one needs typically **exponential** running time --- "**easy-hard transition**" of typical computational complexity.

Weigt et al. *PRL* (2000)
Hartmann et al. *JPC* (2008)

Minimum Vertex Cover

Minimum set *C* of vertices such that each edge of *G* is incident to at least one vertex in *C*.

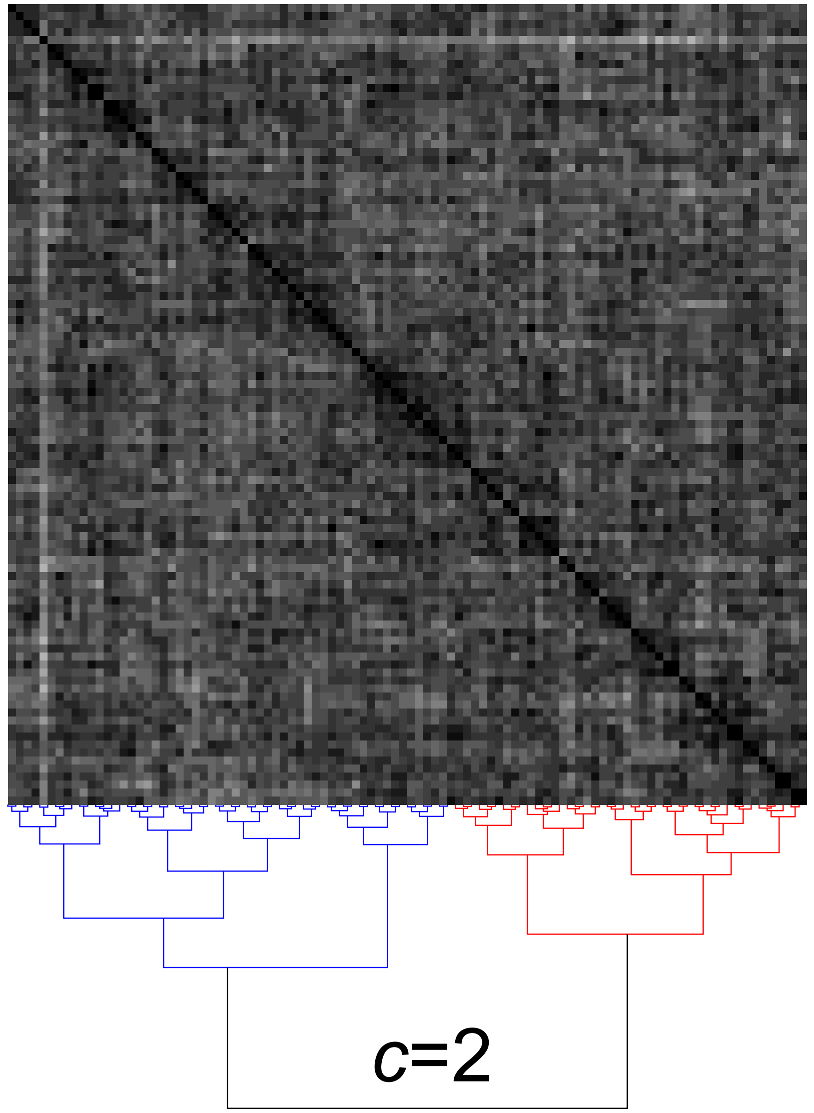


"Minimum Number of Guards Needed by a Museum"

Weigt et al. *PRL* (2000)

Core Percolation vs. Minimum Vertex Cover

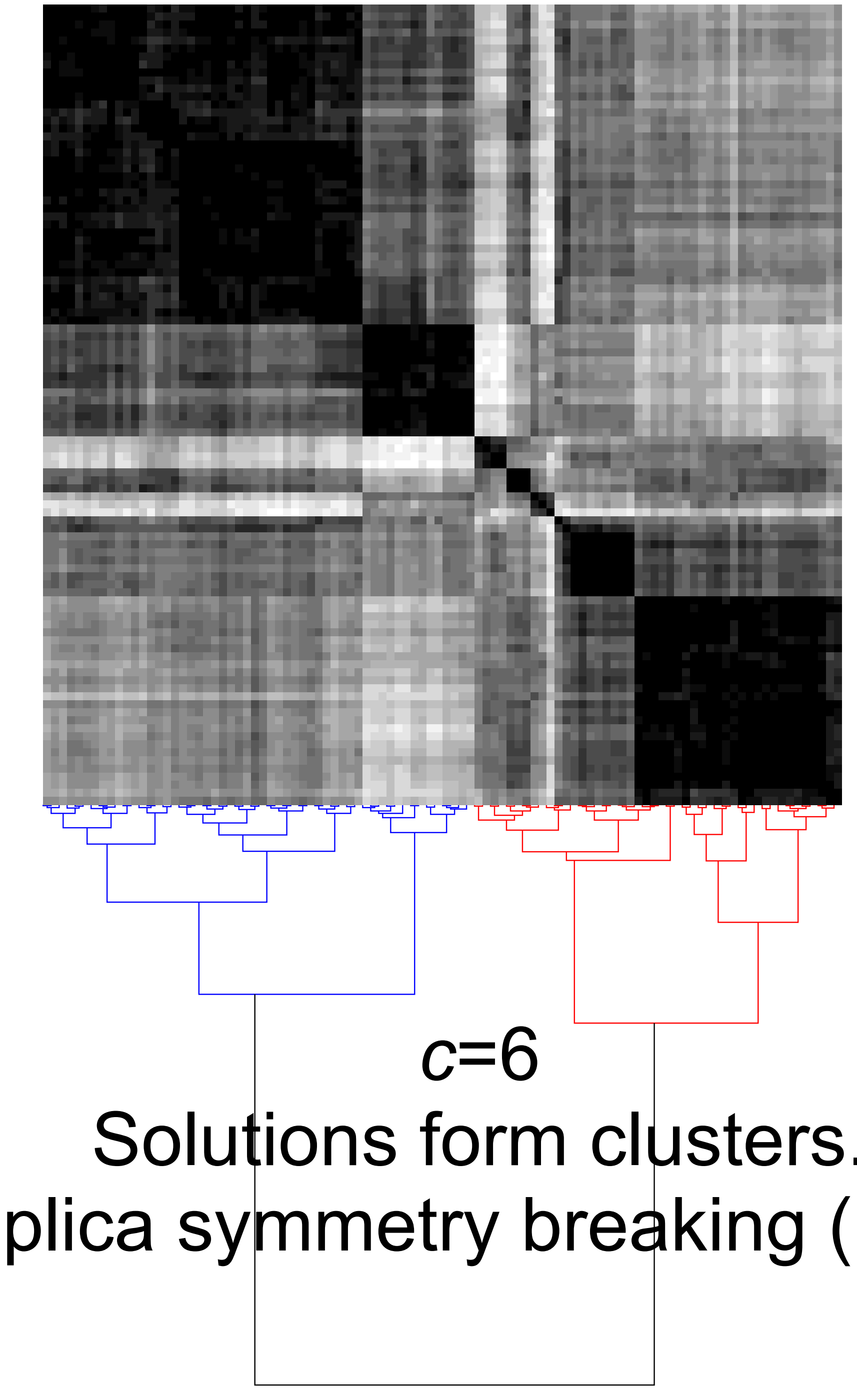
Sample dendrograms of 100 VC solutions for a graph with 400 vertices.



c=2

No structure.

Replica symmetric (RS).



c=6

Solutions form clusters.

Replica symmetry breaking (RSB).

Hartmann et al. *JPC* (2008)

Motivation

1. Minimum vertex cover (MVC) problem

- For ER, core percolation coincides with the changes of the solution-space structure of MVC problem --- **replica symmetry breaking**.
- For ER, if $c \le e$ the typical running time of an algorithm for finding MVC is **polynomial**, while for $c > e$, one needs typically **exponential** running time --- "**easy-hard transition**" of typical computational complexity.

Weigt et al. *PRL* (2000)
Hartmann et al. *JPC* (2008)

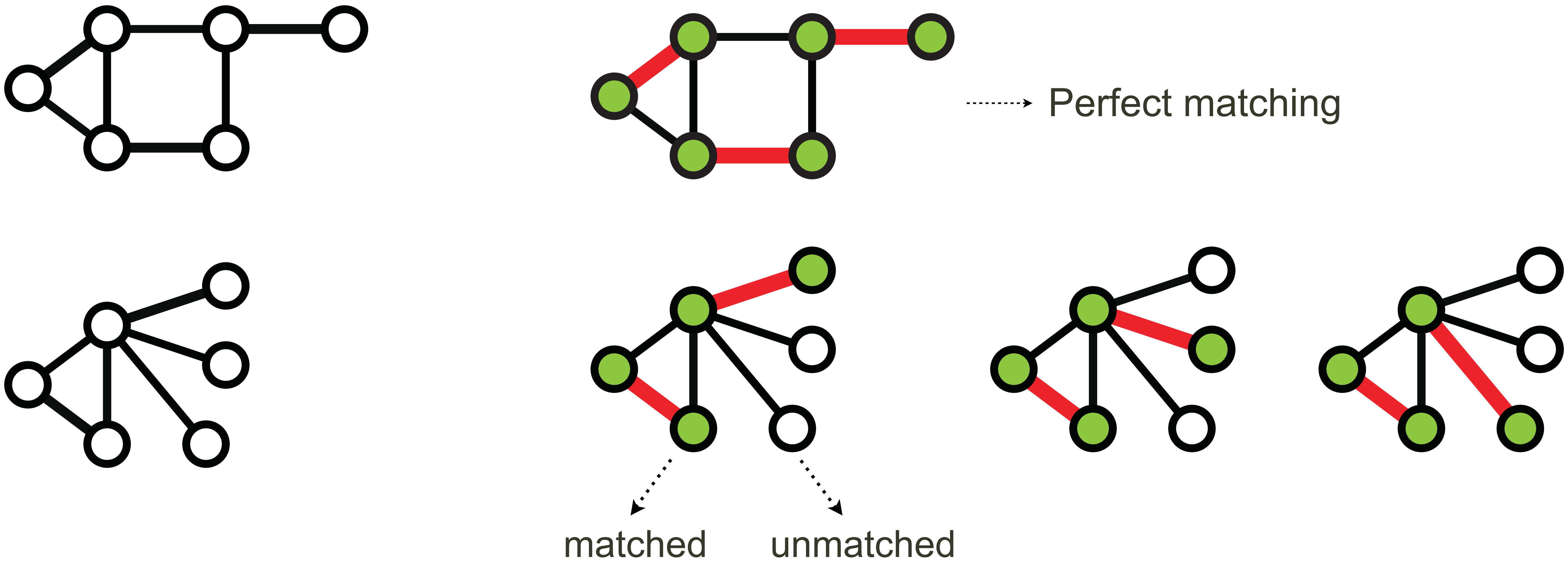
2. Maximum matching (MM) and structural controllability

- Core percolation coincides with the change of the **ground state entropy** in the MM problem.
- **Robustness of structural controllability**: Both link category and node category (critical, redundant, ordinary) are related to core percolation.

Zdeborová et al. *JSM* (2006)
Liu et al. *Nature* (2011)

Maximum Matching

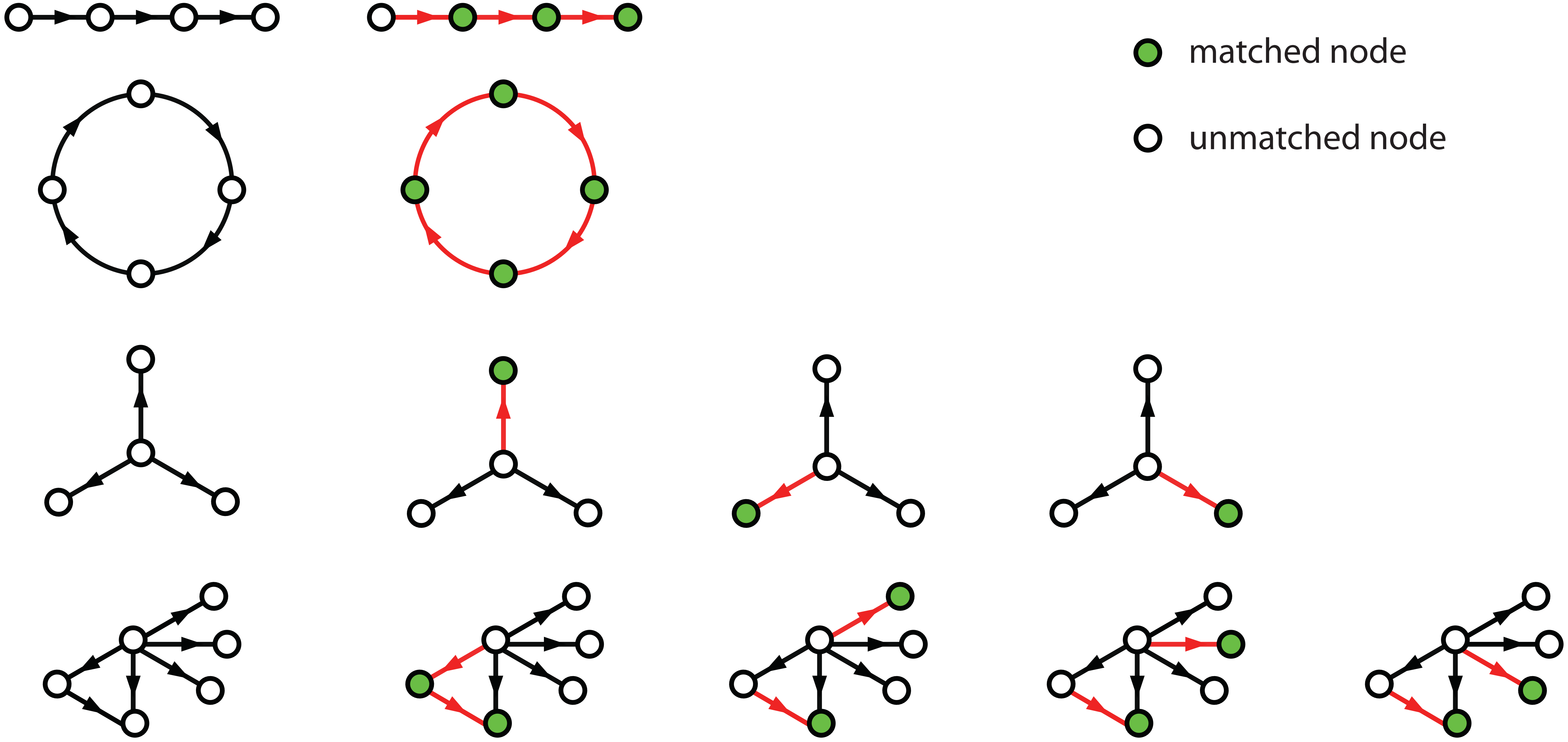
Maximum set *M* of edges without common vertices.



Lovász & Plummer, *Matching Theory*

Maximum Matching in Digraph

Maximum set *M* of edges without common heads or tails.



Liu et al. *Nature* (2011)

Linear Control System

- **Linear Time-Invariant System**

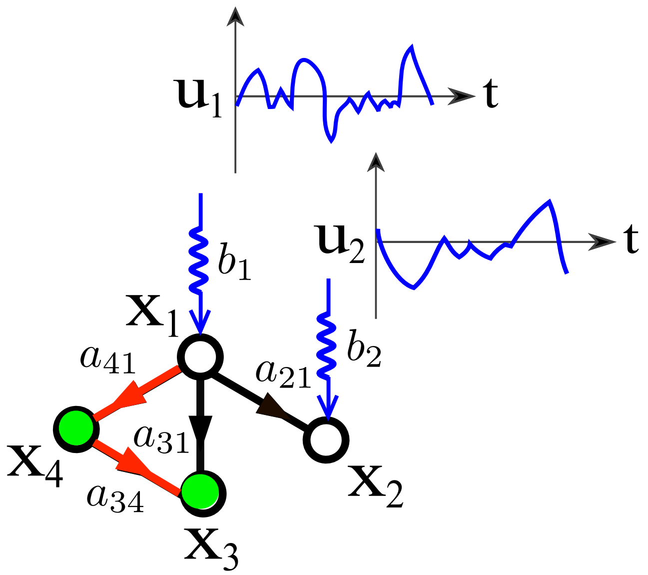
$$\dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t)$$

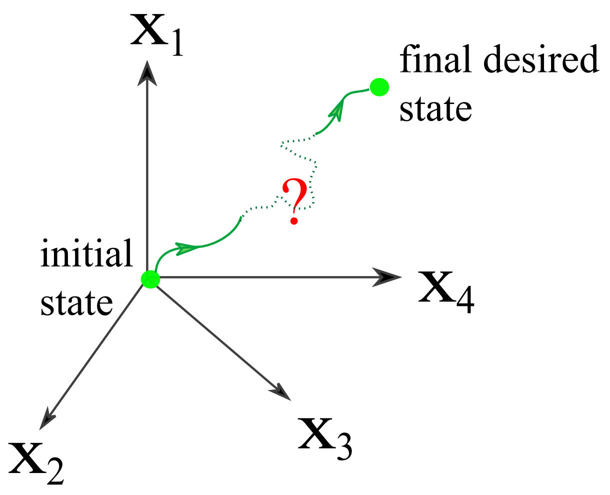
$\mathbf{x}(t) \in \mathbb{R}^{N\times 1}$: state vector.

$\mathbf{u}(t) \in \mathbb{R}^{M\times 1}$: input signals ($M \leq N$).

$\mathbf{A} \in \mathbb{R}^{N\times N}$: state matrix (weighted wiring diagram).

$\mathbf{B} \in \mathbb{R}^{N\times M}$: input matrix ($\Rightarrow$ control configuration).





$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ a_{31} & 0 & 0 & a_{34} \\ a_{41} & 0 & 0 & 0 \end{pmatrix};\ \mathbf{B} = \begin{pmatrix} b_1 & 0 \\ 0 & b_2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix};\ \mathbf{C} = \begin{pmatrix} b_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b_2 & a_{21}b_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_{31}b_1 & 0 & a_{34}a_{41}b_1 & 0 & 0 & 0 \\ 0 & 0 & a_{41}b_1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$N = 4,\ M = 2,\ \mathrm{rank}(\mathbf{C}) = 4 = N$

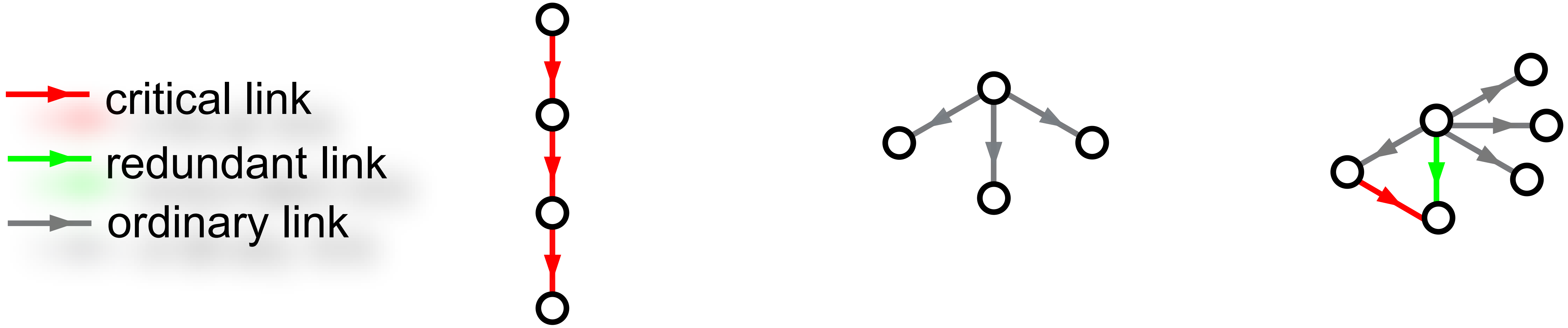
- **Minimum Input Theorem**
  The minimum number of driver nodes (inputs) we need to fully a control network is given by

$$N_D = \max\{1, N_{\text{unmatched}}\}$$

Liu *et al.*, *Nature* **473**, 167 (2011)

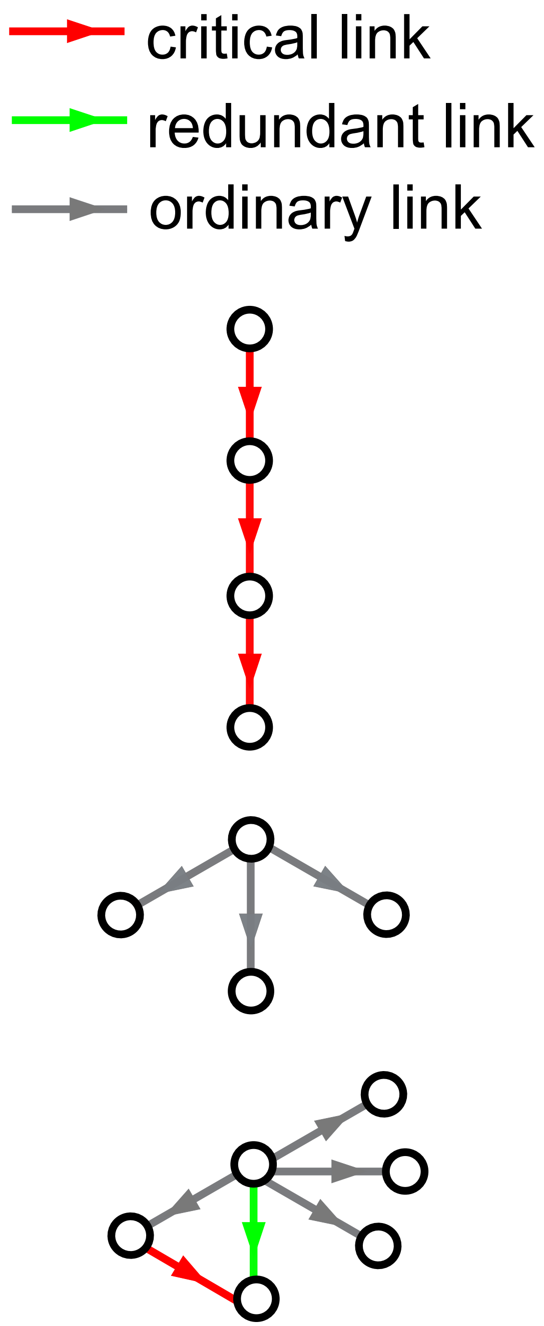
Link category

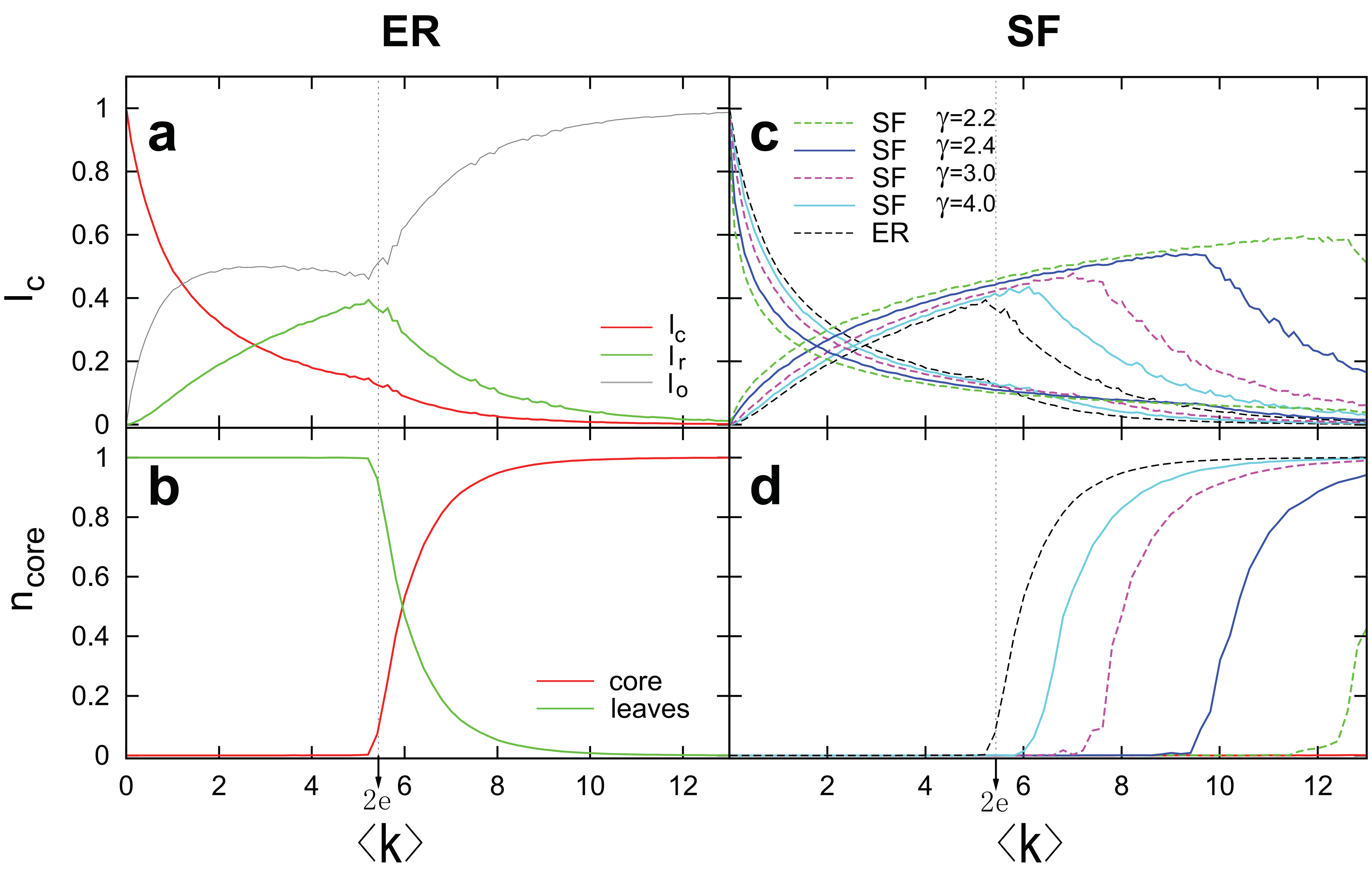
1. critical link ($\in$ all maximum matchings): removal will increase N_{D}.
2. redundant link ($\notin$ any maximum matchings): removal will not change the current driver set and N_{D}.
3. ordinary link ($\in$ some but not all maximum matchings): neither critical nor redundant.



Liu et al. *Nature* (2011)

Core Percolation vs. Link Category



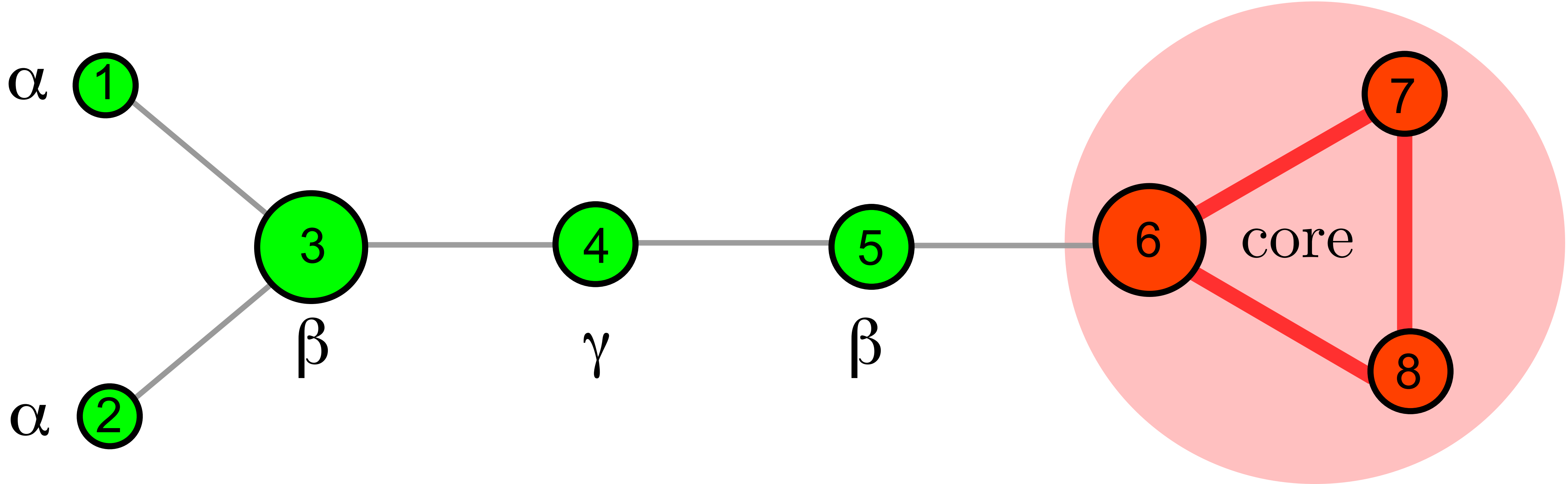


Liu et al. *Nature* (2011)

Theory

Greedy Leaf Removal

Iteratively remove leaf vertex *i*, its neighbor *j*, as well as all the edges incident with *j*.



(1) **α-removable**: nodes that can become isolated without directly removing itself (e.g., v_1 and v_2);
(2) **β-removable**: nodes that can become a neighbor of a leaf (e.g., v_3 and v_5);
(3) **γ-removable**: nodes that can become leaves but are neither nor removable (e.g., v_4);
(4) **non-removable**: nodes that cannot be removed and hence belong to the core (e.g., v_6 , v_7 and v_8).

Self-consistent equations

We can determine the category of a node *v* in graph *G* by the categories of its neighbors in *G* \ *v*. Let α and β denote the probability that a random neighbor of a random node *v* in a network *G* is α-removable and β-removable in *G* \ *v*, respectively.

(1) **α-removable**: all neighbors are β-removable;

$$\boxed{\alpha = \sum_{k=1}^{\infty} Q(k)\beta^{k-1} = A(1-\beta)}$$

with

$$Q(k) \equiv kP(k)/c$$

$$A(x) \equiv \textstyle\sum_{k=0}^{\infty} Q(k+1)(1-x)^k$$

(2) **β-removable**: at least one neighbor is α-removable;

$$\boxed{1-\beta = \sum_{k=1}^{\infty} Q(k)(1-\alpha)^{k-1} = A(\alpha)}$$

(3) **non-removable**: no neighbor is α-removable, and at least two neighbors are not β-removable.

$$\boxed{n_{\text{core}} = \sum_{k=0}^{\infty} P(k) \sum_{s=2}^{k} \binom{k}{s} \beta^{k-s}(1-\beta-\alpha)^s}$$

n_{core} & l_{core}

$$\left\{\begin{array}{rcl} \alpha & = & A(1-\beta) \\ 1-\beta & = & A(\alpha) \end{array}\right. \quad \Rightarrow \quad \alpha = A(A(\alpha))$$

$$n_{\text{core}} = \sum_{k=0}^{\infty} P(k) \sum_{s=2}^{k} \binom{k}{s} \beta^{k-s} (1-\beta-\alpha)^s$$

$$G(x) \equiv \sum_{k=0}^{\infty} P(k) x^k$$

$$\boxed{\begin{aligned} n_{\text{core}} &= G(1-\alpha) - G(\beta) - c(1-\beta-\alpha)\alpha \\ l_{\text{core}} &= \frac{c}{2}(1-\alpha-\beta)^2 \end{aligned}}$$

Condition for core percolation

$\alpha = A(\alpha) = 1 - \beta$ is always a root of $f(x) \equiv A(A(x)) - x$

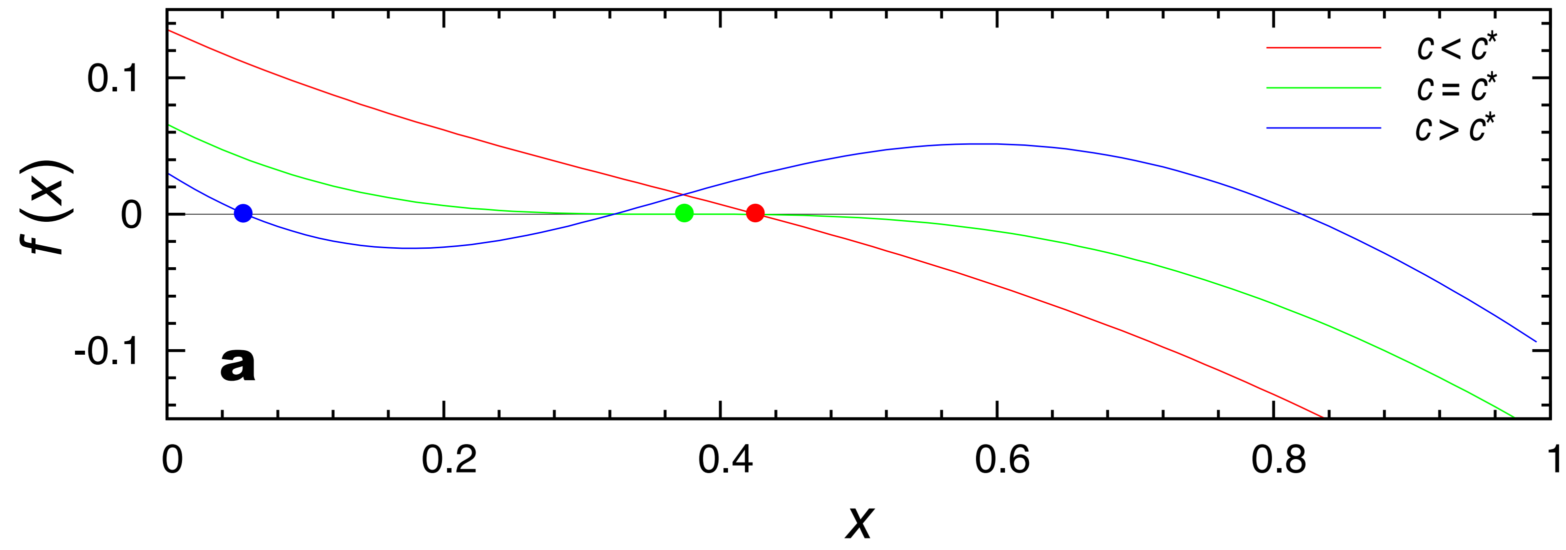
$$\boxed{\begin{aligned} n_{\rm core} &= G(1-\alpha) - G(\beta) - c(1-\beta-\alpha)\alpha \\ l_{\rm core} &= \frac{c}{2}(1-\alpha-\beta)^2 \end{aligned}}$$

But if $\alpha = 1 - \beta$, the core will not exist.

$\boxed{f(x) \text{ must have multiple roots for the core to exist!}}$

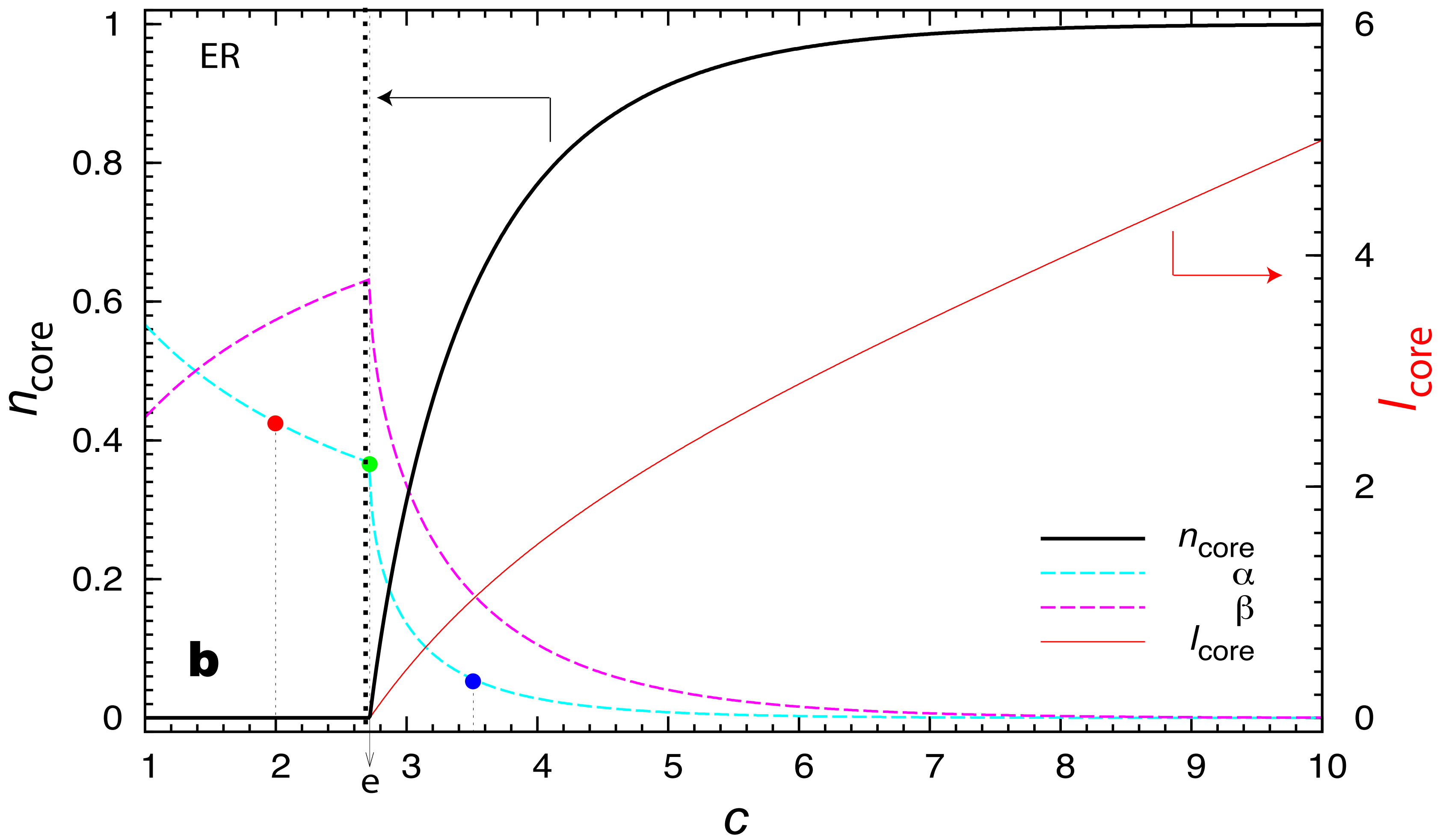
Poisson distributed graphs

$$P(k) = e^{-c}\frac{c^k}{k!}$$

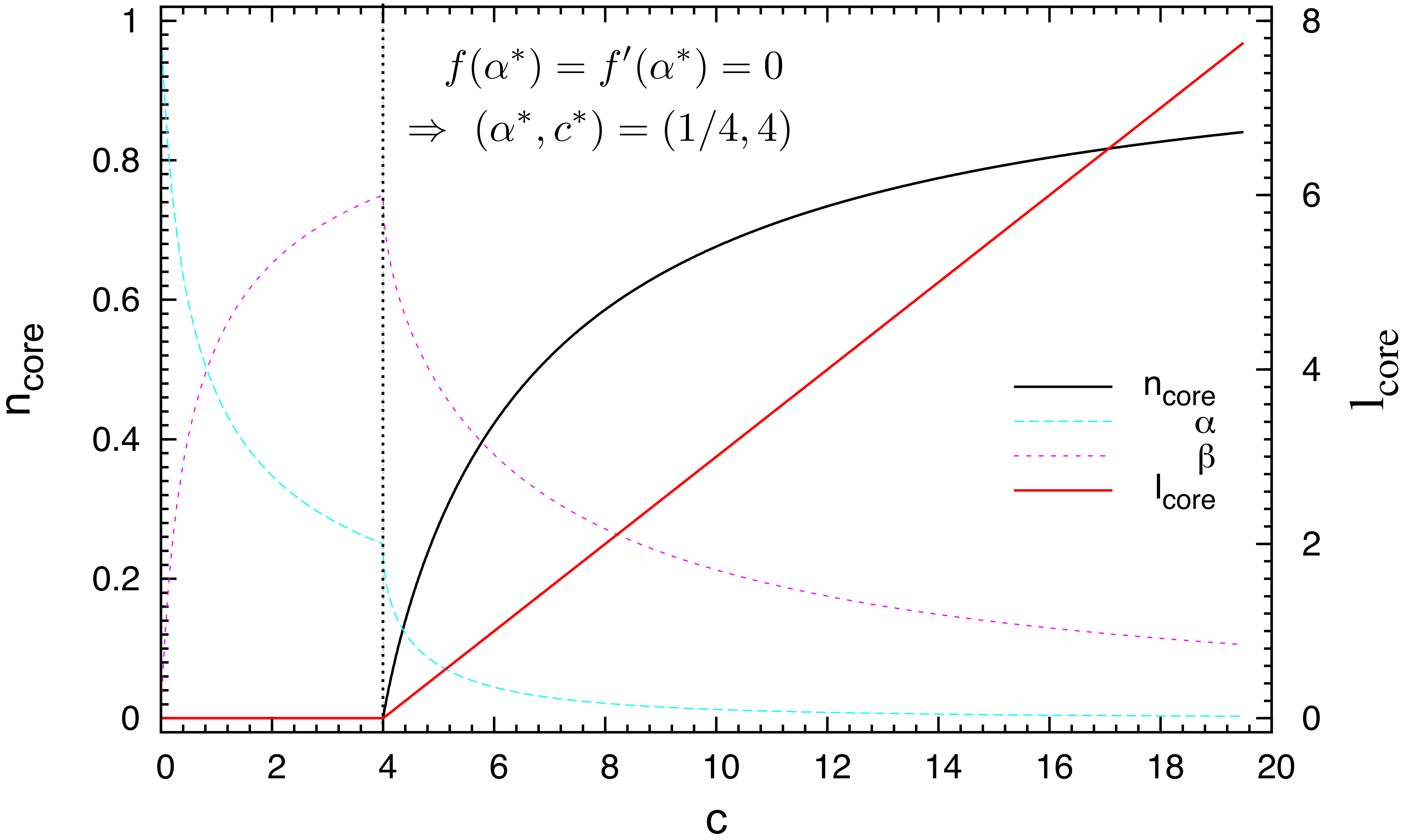


$$f(\alpha^*) = f'(\alpha^*) = 0$$

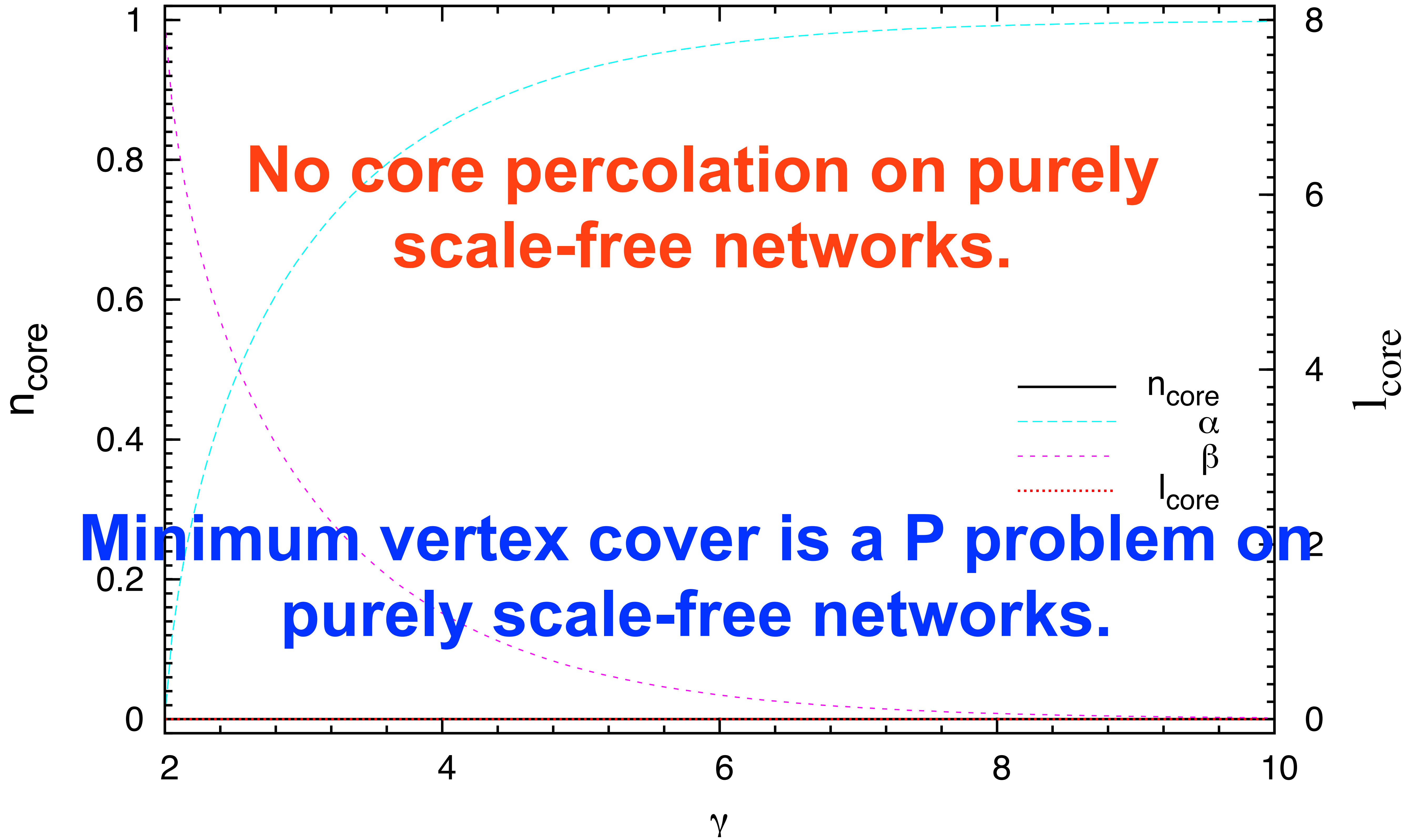
$$\Rightarrow\ (\alpha^*, c^*) = (e^{-1}, e)$$



Exponentially distributed graphs $P(k) = (1 - e^{-1/\kappa}) e^{-k/\kappa}$



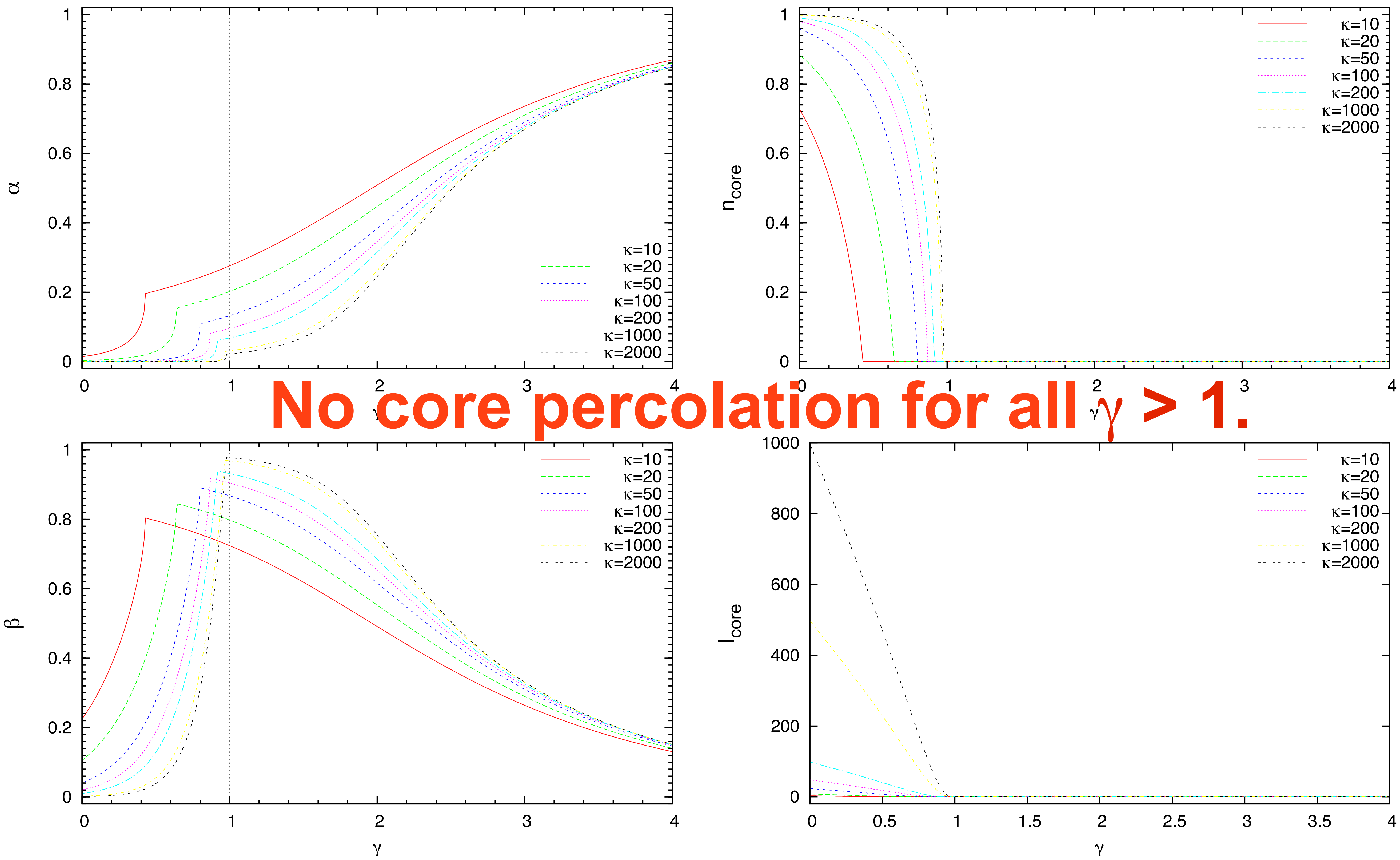
Purely power-law distributed graphs $P(k) = \dfrac{k^{-\gamma}}{\zeta(\gamma)} \quad \text{for} \quad k \geq 1$



No core percolation on purely scale-free networks.

Minimum vertex cover is a P problem on purely scale-free networks.

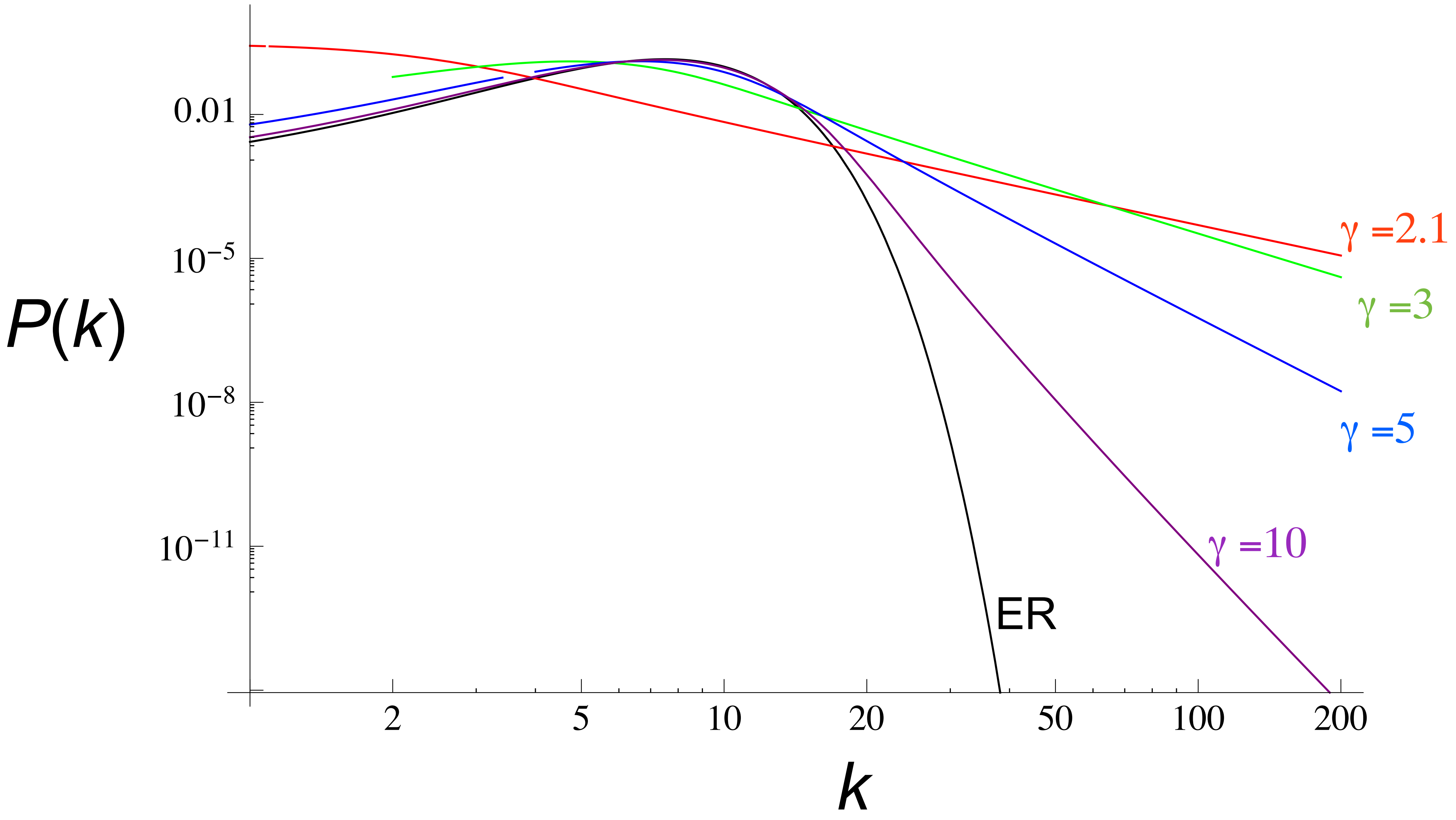
Power-law distribution with exponential cutoff $P(k) = \dfrac{k^{-\gamma}\, e^{-k/\kappa}}{\mathrm{Li}_\gamma(e^{-1/\kappa})}$ for $k \geq 1$



No core percolation for all γ > 1.

Asymptoticly scale-free network: static model

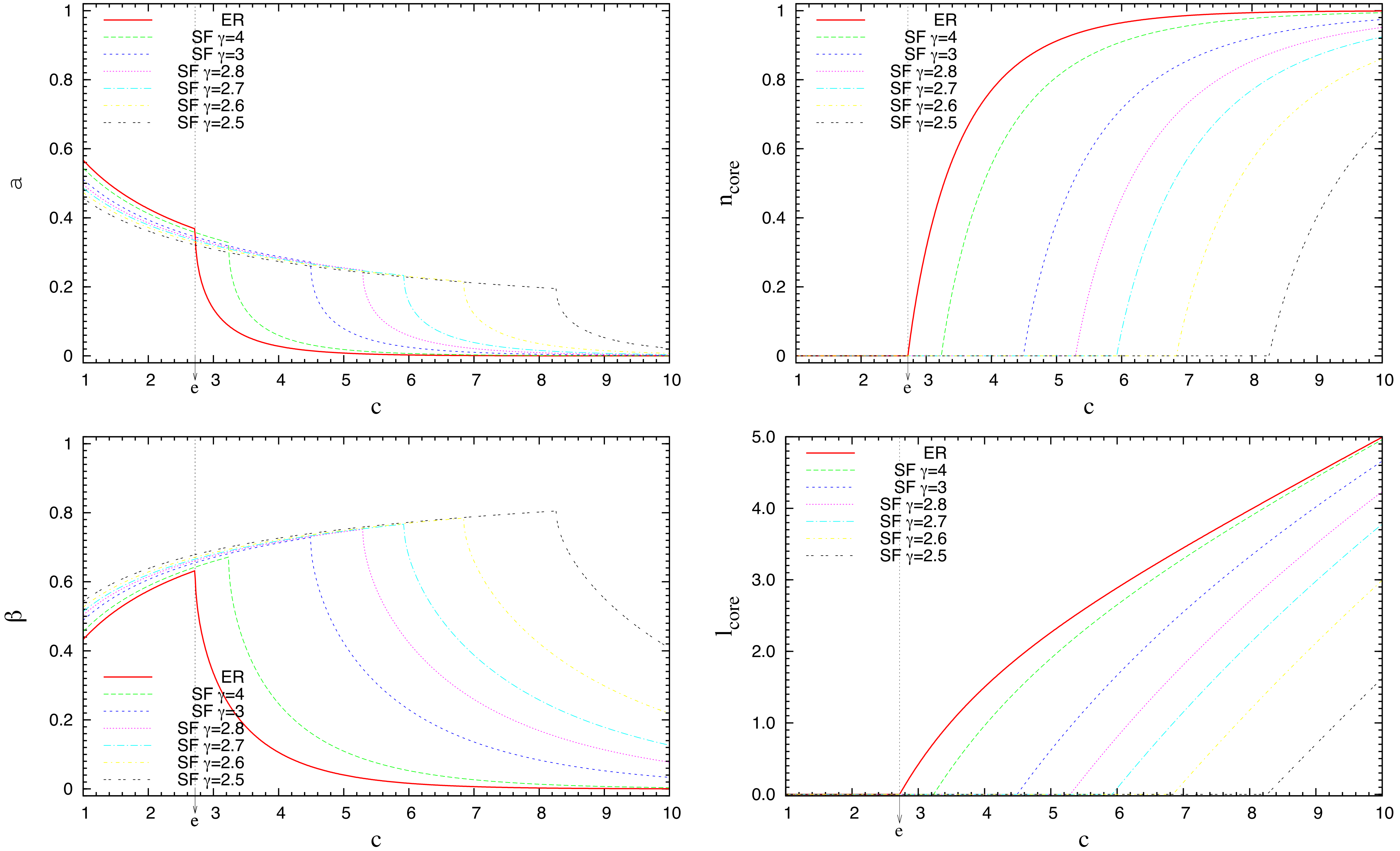
$$P(k) = \frac{[m(1-\xi)]^{1/\xi}}{\xi} \frac{\Gamma(k - 1/\xi, m[1-\xi])}{\Gamma(k+1)} \sim k^{-(1+\frac{1}{\xi})} = k^{-\gamma} \text{ where } \gamma = 1 + \frac{1}{\xi}.$$



Goh et al. PRL (2001)
Catanzaro et al. EPJB (2005)
Lee et al. EPJB (2006)

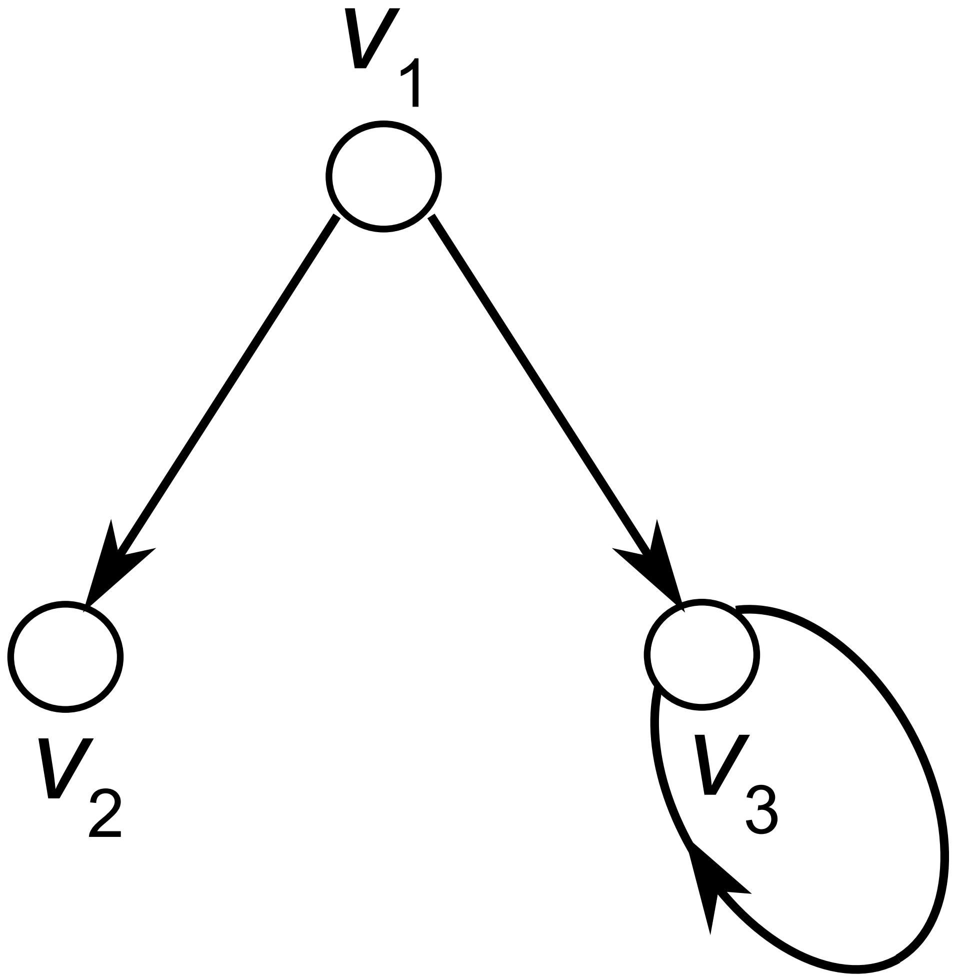
Asymptoticly scale-free network: static model

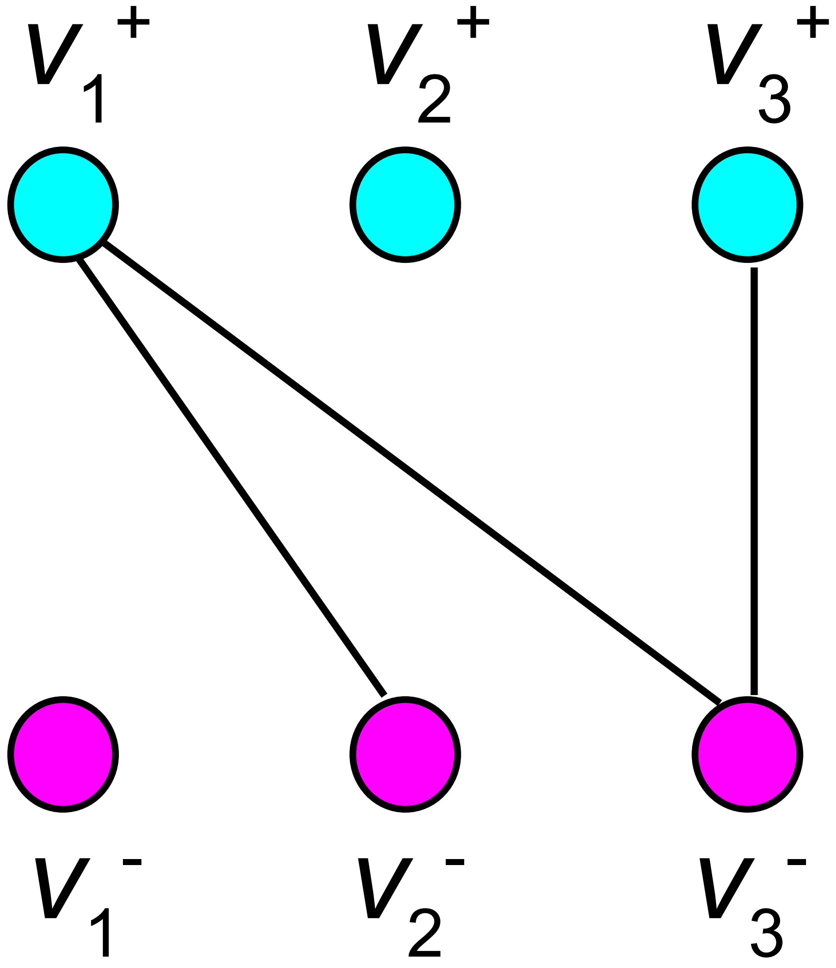
$$P(k) = \frac{[m(1-\xi)]^{1/\xi}}{\xi} \frac{\Gamma(k - 1/\xi, m[1-\xi])}{\Gamma(k+1)} \sim k^{-(1+\frac{1}{\xi})} = k^{-\gamma} \text{ where } \gamma = 1 + \frac{1}{\xi}.$$



How about directed networks?

Core percolation on digraphs





$$
\begin{aligned}
Q^{\pm}(k) &\equiv kP^{\pm}(k)/c \\
A^{\pm}(x) &\equiv \sum_{k=0}^{\infty} Q^{\pm}(k+1)(1-x)^k \\
G^{\pm}(x) &\equiv \sum_{k=0}^{\infty} P^{\pm}(k)x^k
\end{aligned}
$$

Core percolation on digraphs

$$\left\{ \begin{array}{rcl} \alpha^{\pm} & = & A^{\pm}(1-\beta^{\mp}) \\ 1-\beta^{\pm} & = & A^{\pm}(\alpha^{\mp}) \end{array} \right. \quad \Rightarrow \quad \alpha^{\pm} = A^{\pm}(A^{\mp}(\alpha^{\pm}))$$

$$\begin{aligned} n_{\text{core}}^{\pm} &= \sum_{k=0}^{\infty} P^{\pm}(k) \sum_{s=2}^{k} \binom{k}{s} (\beta^{\mp})^{k-s} (1-\beta^{\mp}-\alpha^{\mp})^{s} \\ &= G^{\pm}(1-\alpha^{\mp}) - G^{\pm}(\beta^{\mp}) - c(1-\beta^{\mp}-\alpha^{\mp})\alpha^{\pm} \end{aligned}$$

$$\boxed{\begin{array}{rcl} n_{\text{core}} & = & \frac{1}{2}(n_{\text{core}}^{+} + n_{\text{core}}^{-}) \\ l_{\text{core}} & = & c(1-\beta^{+}-\alpha^{+})(1-\beta^{-}-\alpha^{-}) \end{array}}$$

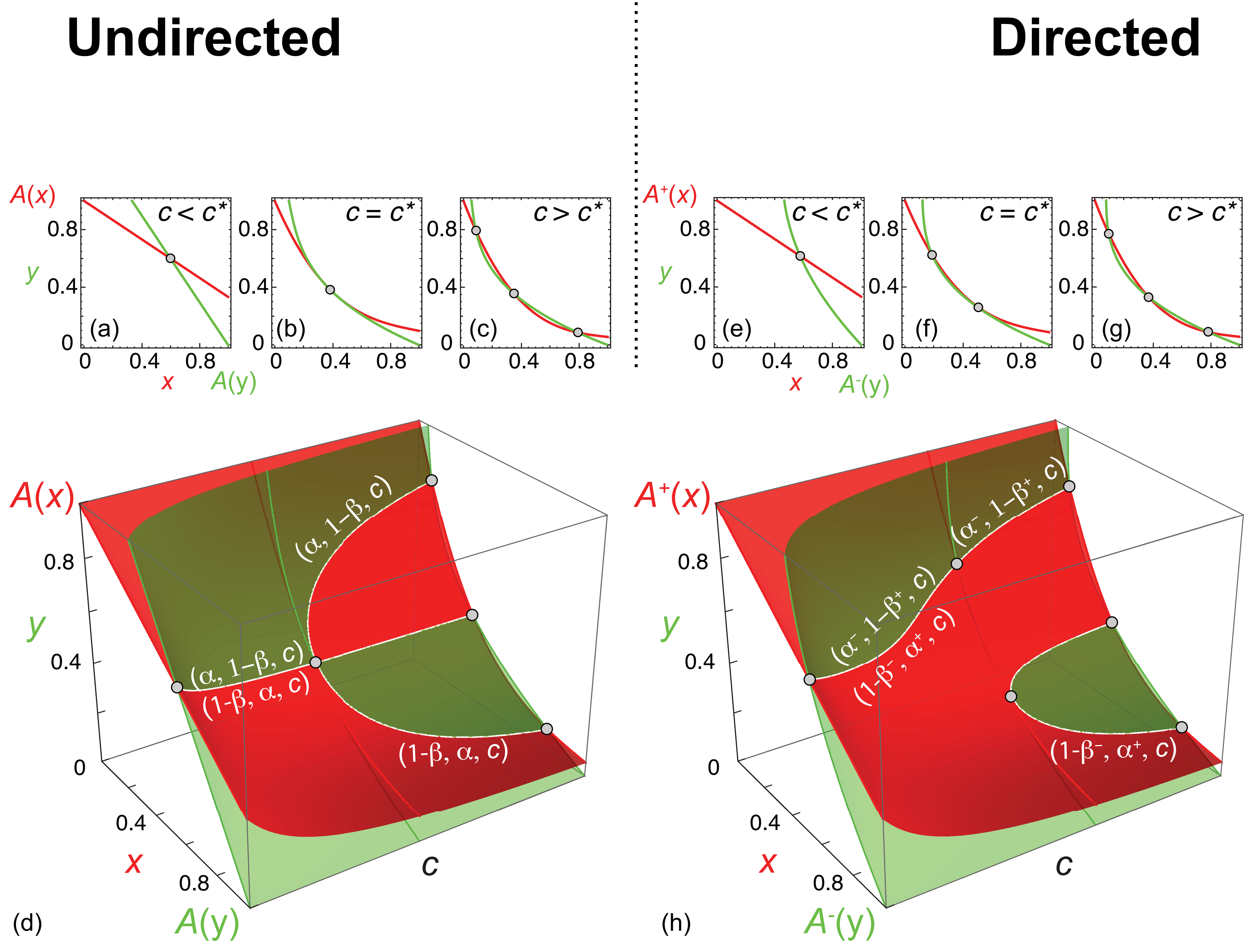
Condition for core percolation

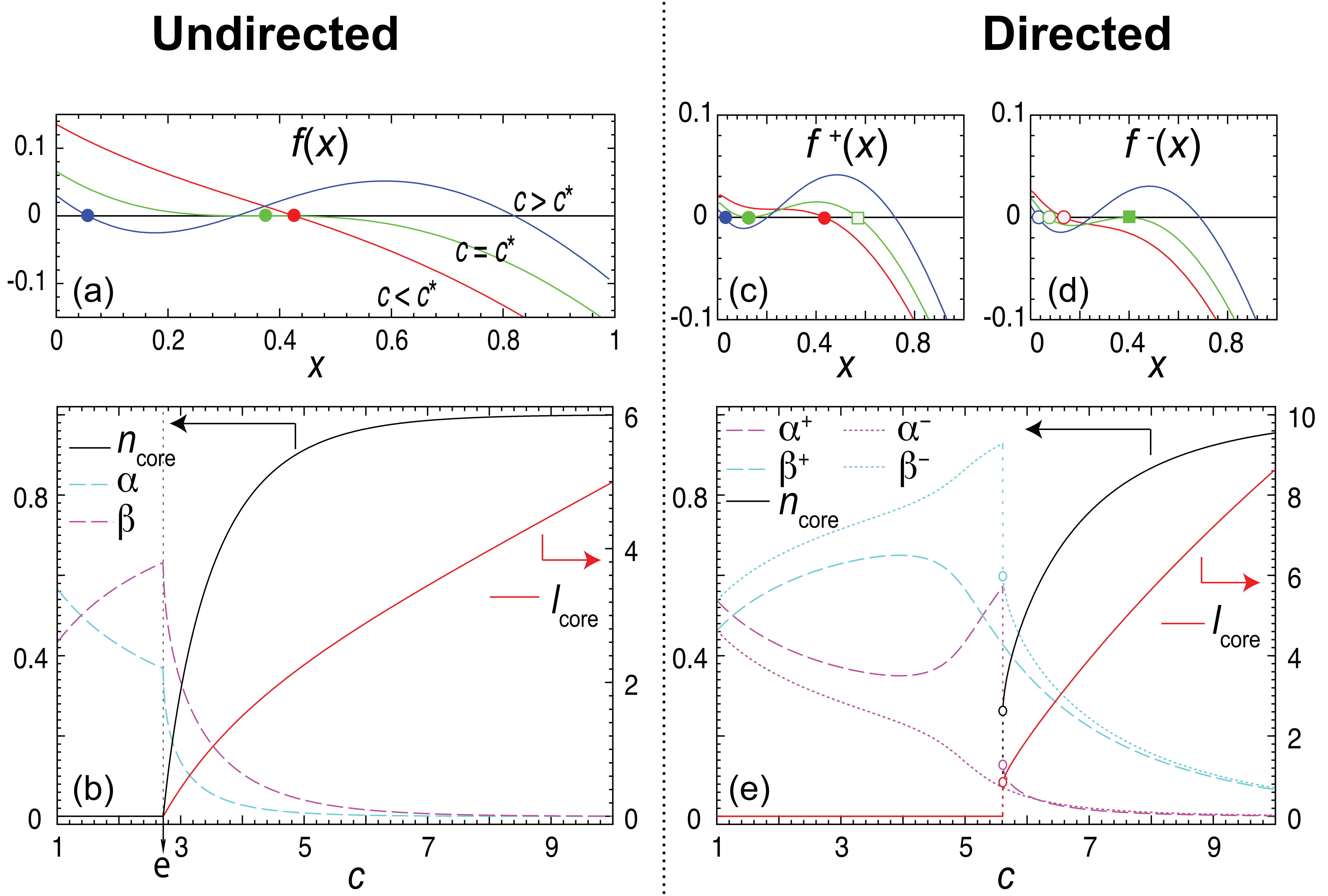
$\alpha^{\pm} = A^{\pm}(\alpha^{\mp}) = 1 - \beta^{\pm}$ is always a root of $f^{\pm}(x) \equiv A^{\pm}(A^{\mp}(x)) - x$

$$\boxed{\begin{aligned} n_{\text{core}}^{\pm} &= G^{\pm}(1-\alpha^{\mp}) - G^{\pm}(\beta^{\mp}) - c(1-\beta^{\mp}-\alpha^{\mp})\alpha^{\pm} \\ l_{\text{core}} &= c(1-\beta^{+}-\alpha^{+})(1-\beta^{-}-\alpha^{-}) \end{aligned}}$$

But if $\alpha^{\pm} = 1 - \beta^{\pm}$, the core will not exist.

$\boxed{f^{\pm}(x) \text{ must have multiple roots for the core to exist!}}$


Undirected
Directed
A(x)
y
x
A(y)
c < c*
c = c*
c > c*
(a)
(b)
(c)
(e)
(f)
(g)
A⁺(x)
A⁻(y)
(α, 1–β, c)
(1-β, α, c)
(α⁻, 1–β⁺, c)
(1-β⁻, α⁺, c)
c
(d)
(h)

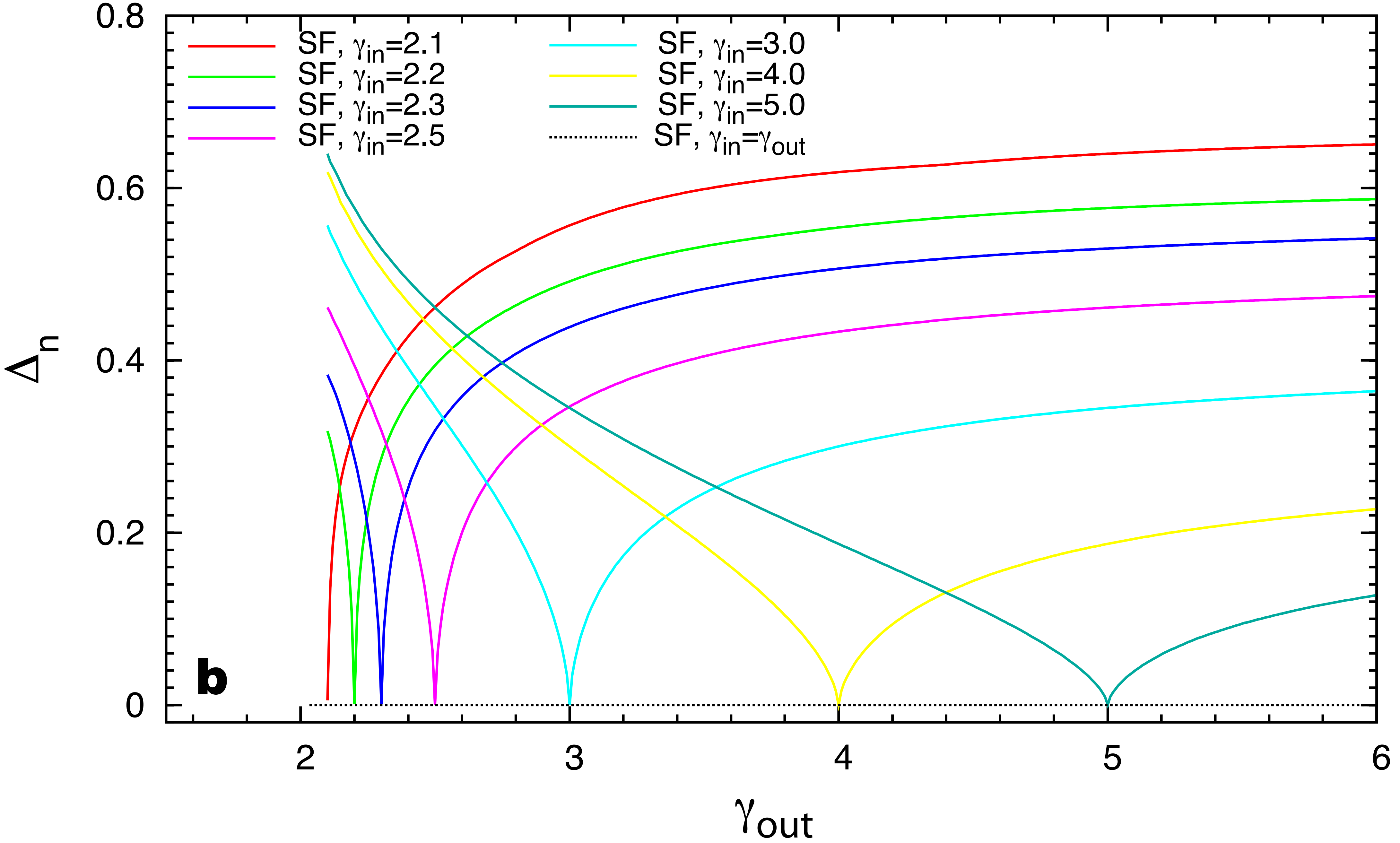


Undirected
Directed
f(x)
c > c*
c = c*
c < c*
(a)
x
f +(x)
(c)
f -(x)
(d)
n core
α
β
l core
(b)
e
c
α+
α−
β+
β−
(e)

Discontinuity

$$\Delta_{\mathrm{n}}^{\pm} \;=\; G^{\pm}(1-\alpha^{\mp,*},c) - G^{\pm}(\beta^{\mp,*},c) - c^{*}\alpha^{\pm,*}\rho^{\mp}$$

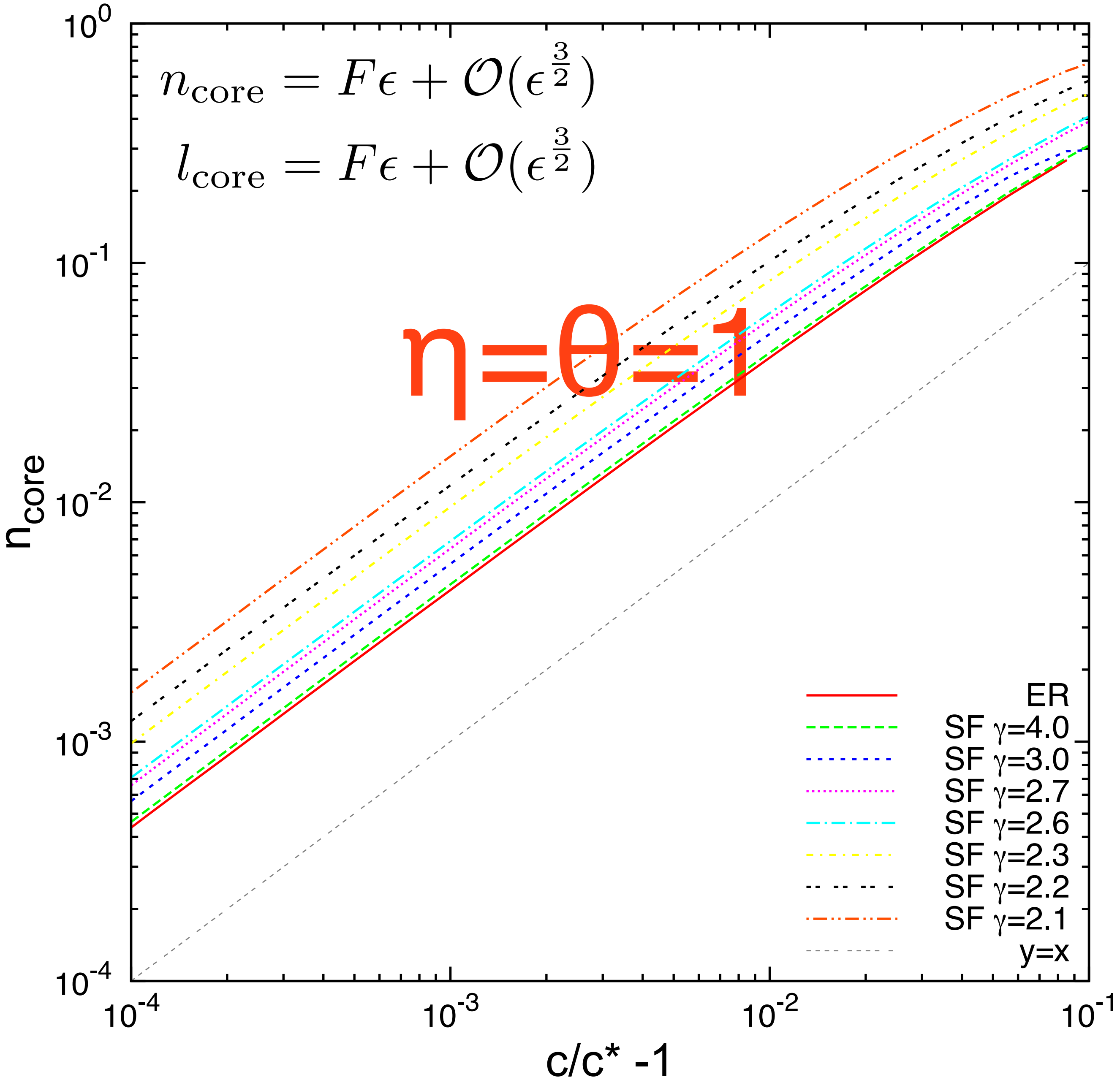
$$\Delta_{\mathrm{n}} \;\equiv\; \tfrac{1}{2}\left(\Delta_{\mathrm{n}}^{+} + \Delta_{\mathrm{n}}^{-}\right)$$



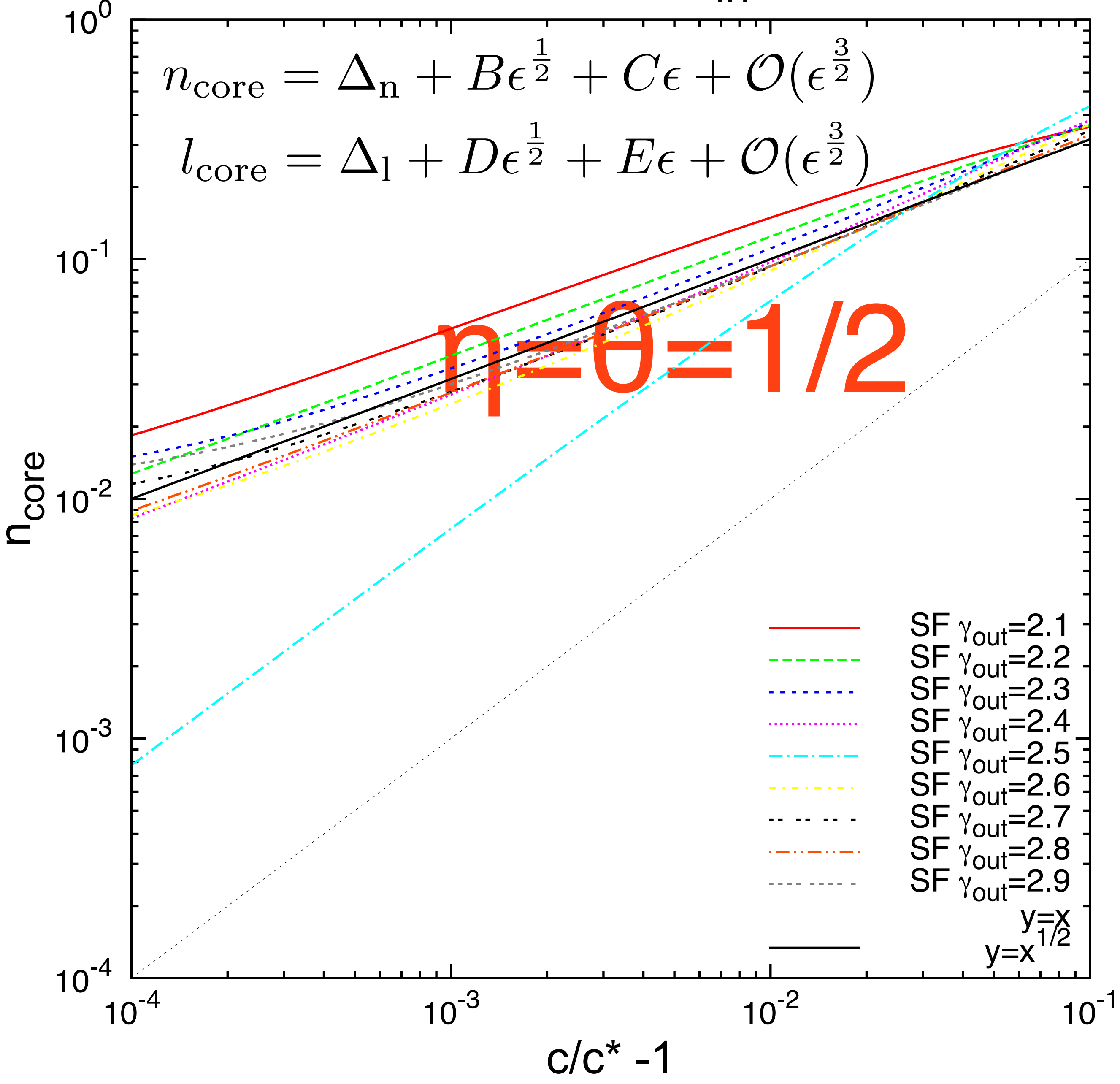
Critical exponents

$$n_{\text{core}} - \Delta_{\text{n}} \sim (c - c^*)^{\eta}$$
$$l_{\text{core}} - \Delta_{\text{l}} \sim (c - c^*)^{\theta}$$

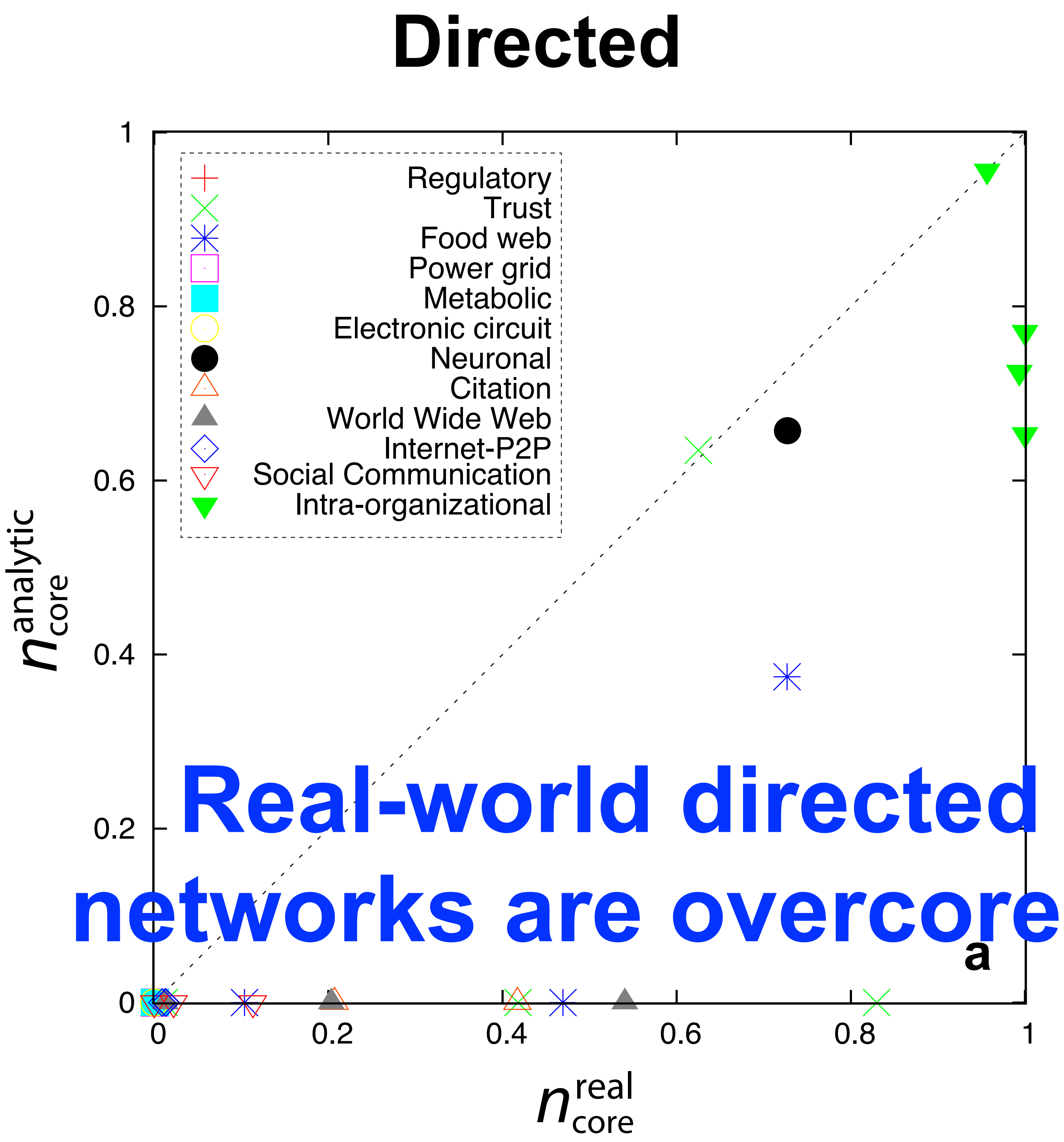
Undirected



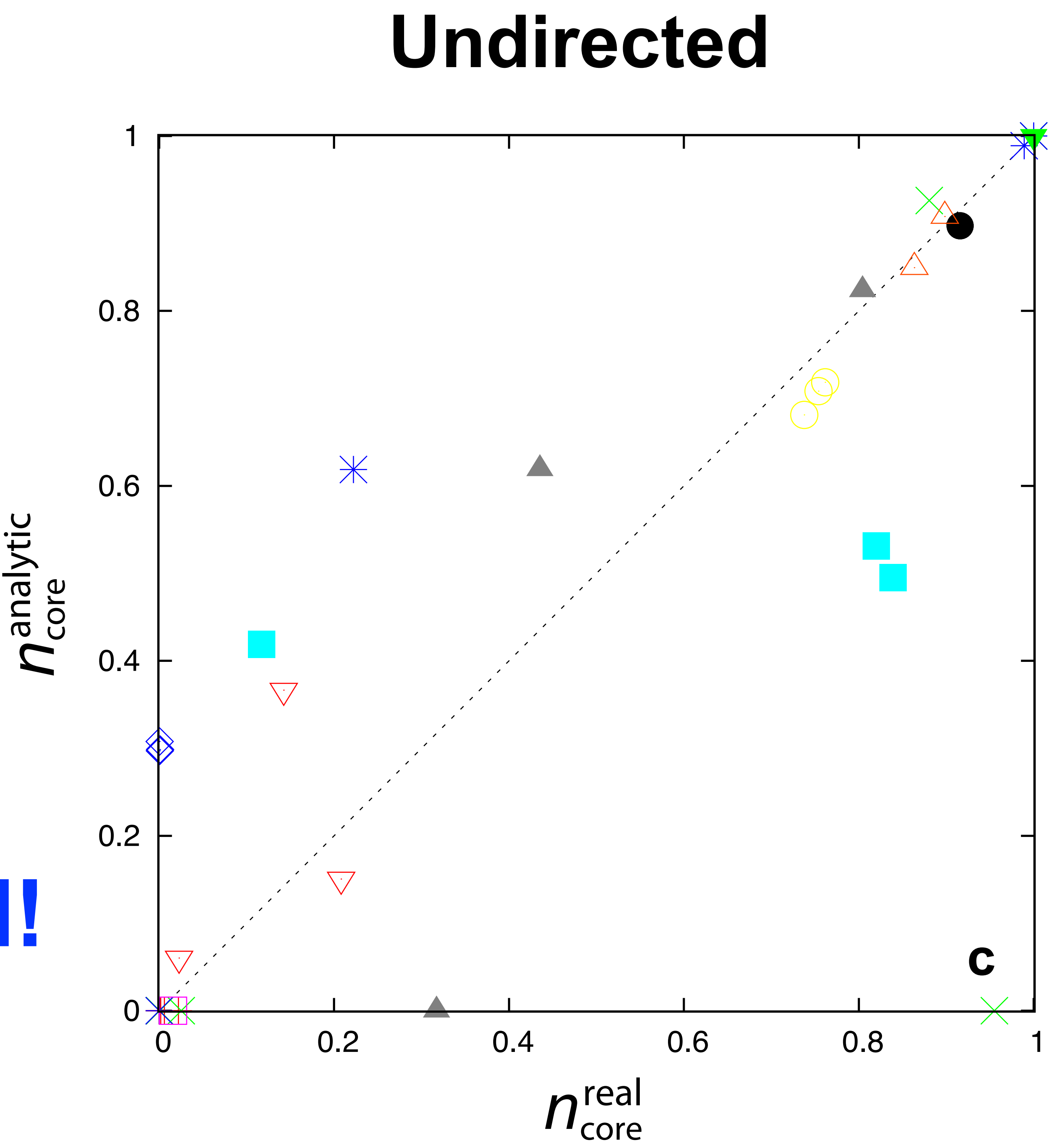
Directed, SF γ_{in}=2.5



Real networks



Real-world directed networks are overcored!



Summary

1. We derive the condition for core percolation on arbitrary complex networks.
    - Purely scale-free networks have no core for any degree exponents.
    - Asymptotically scale-free networks could have core.

2. We show that core percolations in undirected and directed networks have completely different nature.
    - Continuous vs. discontinuous (or hybrid).
    - Different set of critical exponents.

3. Our theory can be used to check if real-world networks have unexpected core sizes.
    - Directed: most real networks have larger cores as compared to random models.
    - Undirected: no systematic deviation.

Acknowledgement

Collaborators: Endre Csóka[1], Haijun Zhou[2], Márton Pósfai[1,3]

1. Eötvös Loránd University
2. Chinese Academy of Science
3. Budapest University of Technology and Economics

ARL ONR JAMES S. MCDONNELL FOUNDATION

1. Nature 473, 167 (2011)
2. Phys. Rev. Lett. 109, 205703 (2012)
3. arXiv:1301.2895, submitted to Phys. Rev. Lett., under review