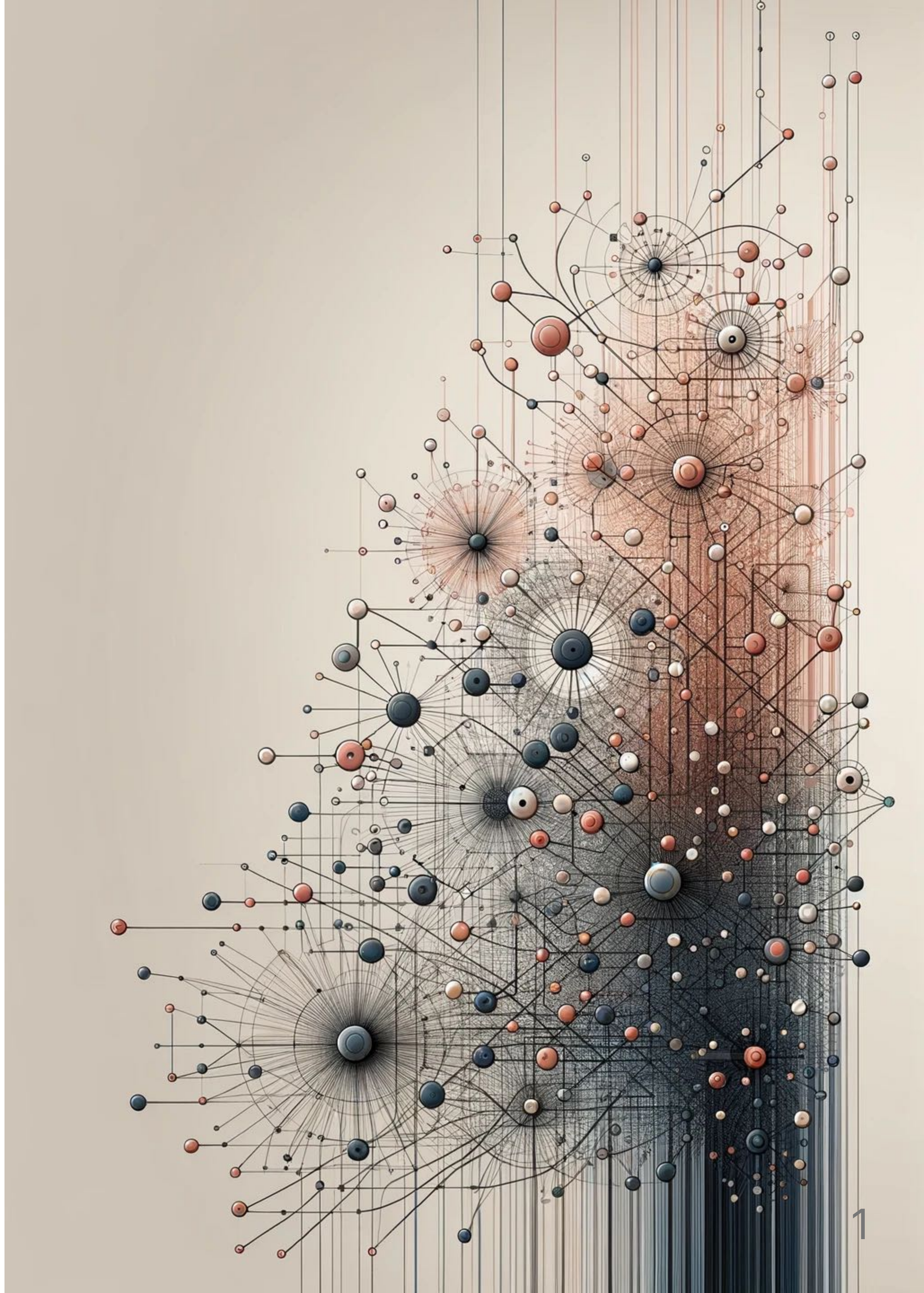
Analisi e Visualizzazione delle Reti Complesse

NS18 - Cascading behaviors in networks

Prof. Rossano Schifanella



Agenda

- **Diffusion in Networks:** How innovations and behaviors spread socially
- **Modeling Diffusion through a Network:** Formal models of influence
- **Cascades and Clusters:** When spread continues or stops
- **Diffusion, Thresholds, and the Role of Weak Ties:** Simple vs. complex contagions
- **Extensions of the Basic Cascade Model:** Heterogeneous thresholds, collective action, cascade capacity, and compatibility

The Diffusion of Innovations

- Many innovations—new behaviors, practices, opinions, conventions, technologies—spread person-to-person via social influence. Neighbors or friends adopting something new can influence others to follow.
- *Diffusion of innovation* research (mid-20th century) emphasized both **informational effects** (learning from others) and **network effects** (influence through social ties). Classic studies by Rogers (1962) and others identified adoption categories (innovators, early adopters, etc.).
- We focus on **direct-benefit effects**: situations where adopting a technology or behavior yields increasing benefits as more friends or contacts adopt it too. Examples:
  - Telephones or social media: each friend who joins increases its usefulness (a network externality).
  - Products like fax machines or email: more adopters among acquaintances provide direct benefits for you to adopt.

UNIVERSITÀ DI TORINO

The Diffusion of Innovations

- **Network effects at the local level** are crucial: it's not just global popularity, but whether *your* friends have adopted.
- **Objective:** Formulate a simple model for the spread of an innovation through a social network, capturing how local network structure and incentives drive diffusion.

Diffusion of a New Behavior

- **Key assumption:** Individuals base their adoption decisions on the choices of their neighbors. This is a *local* decision rule focusing on peer influence along network links.
  - We ignore external marketing or media influence here and concentrate on peer-to-peer spread.
- We emphasize **direct-benefit (coordination) effects** rather than pure information or "hype." Each person's payoff for adopting increases as more neighbors adopt the same behavior.
- A natural modeling framework: a **networked coordination game** (first introduced by Morris, 2000).
  - Each node (individual) chooses between two behaviors (or technologies), say A (new behavior) and B (status quo).
  - People have an incentive to align with their neighbors' choices to gain coordination benefits (compatibility, communication, etc.).
- **Reference:** Stephen Morris (2000), *Contagion*, which provided one of the first formal analyses of such network diffusion processes.

Examples of Diffusion in Networks

VHS

vs

Betamax

Blu-ray Disc

vs

HD DVD

A Networked Coordination Game

- To formalize, consider a **coordination game on a network**:
  - Each node chooses either behavior A or behavior B.
  - **Payoffs:** If a node and a neighbor choose the same behavior, both gain a benefit. Let's say:
    - If both use A, each gets payoff *a* from that link.
    - If both use B, each gets payoff *b* from that link.
    - If they differ, payoff is 0 on that link (mismatch yields no benefit).
  - Typically assume *a* and *b* are positive, and one might be larger (A might be superior but new; B is inferior but established).
  - This captures direct-benefit effects: you benefit from aligning choices.

- Each node wants to coordinate with neighbors. Thus, the more neighbors choosing A, the more attractive A becomes.
- This leads to a simple decision rule based on the **fraction of neighbors** on A vs. B and the payoffs *a* and *b*.

$$
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{w} \\
 & & A & B \\
\hline
\multirow{2}{*}{v} & A & a, a & 0, 0 \\
\cline{3-4}
 & B & 0, 0 & b, b \\
\hline
\end{array}
$$

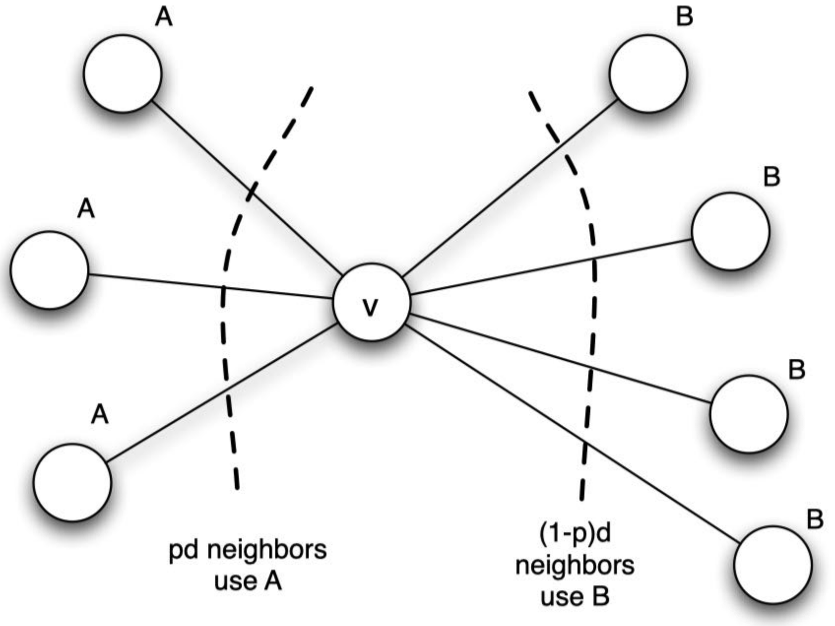
A Networked Coordination Game: Threshold Condition

Let:

- p = *fraction* of a node's neighbors who have adopted A.
- $(1-p)$ = *fraction* of a node's neighbors who have adopted B.
- d = total number of neighbors (node's degree).

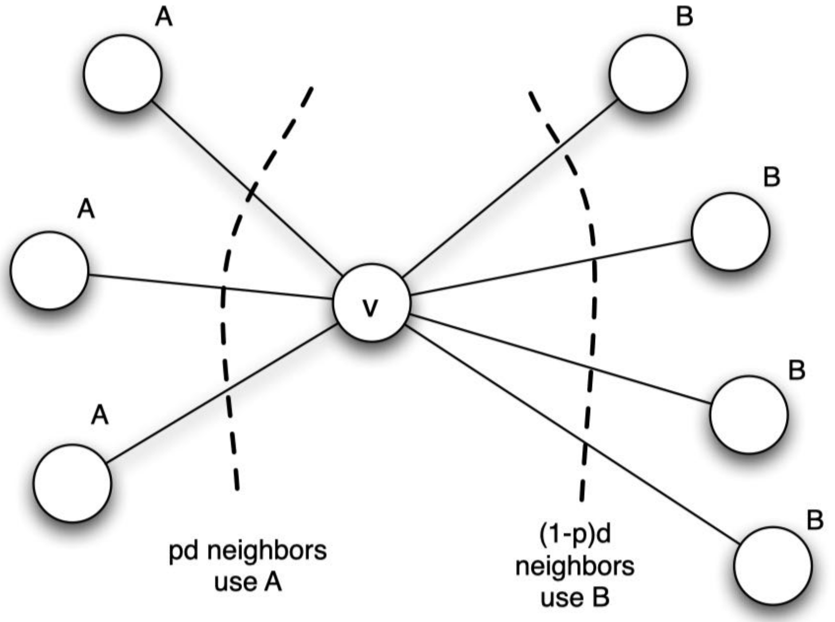
Node v will choose A if:

$$p \cdot d \cdot a \geq (1-p) \cdot d \cdot b$$
$$pa \geq (1-p)b \quad \text{(dividing both sides by } d > 0)$$
$$pa \geq b - pb$$
$$pa + pb \geq b$$



A Networked Coordination Game: Threshold Condition

- We define $q = \frac{b}{a+b}$ as the **threshold fraction** of neighbors adopting A that makes the node indifferent. If at least a fraction q of your neighbors choose A, you prefer A.
- Intuition: If $a > b$, then $q < 0.5$, meaning even a minority of neighbors on A might tip the balance (since A yields higher payoff per neighbor). If $a < b$ (A is lower quality), $q > 0.5$, meaning you need a majority of neighbors on A to make it worthwhile.



UNIVERSITÀ DI TORINO

Threshold Rule

- The adoption decision rule for each node:
  - **Adopt A if** $p \geq q$, i.e., if at least a fraction q of your neighbors have adopted A.
  - Otherwise, stick with B.
- This is a very simple, **myopic threshold model** of decision-making:
  - "Myopic" because each node just looks at the current state of neighbors (one step look-ahead, no strategic long-term planning).
  - All nodes use the same threshold q (for now) derived from payoffs. Later we will consider heterogeneity in q.
- Such threshold models are widely used to study **social contagion** (e.g., Granovetter's threshold model of collective behavior).
- **Research note:** In reality, people might consider more complex strategies, expectations, or long-range outcomes. Here we assume synchronous rounds where each node updates based on neighbors, which is a baseline model. Richer models can incorporate things like inertia, decay, or foresight.

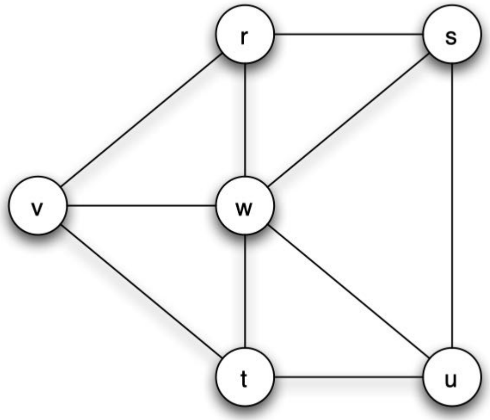
Cascading Behavior

- Given this threshold rule, the network can exhibit multiple **equilibria** (stable outcomes):
  - Everyone adopts A (all nodes switched to the new behavior).
  - Everyone sticks with B (no one adopts the innovation).
  - These correspond to two stable **conventions** or **norms**.
- Key questions:
  - Under what conditions will the all-A equilibrium emerge starting from a few initial adopters? How easy is it to get a full cascade to A?
  - Can there be *partial* adoption equilibria (some mix of A and B stable)? Under what structures do they occur?

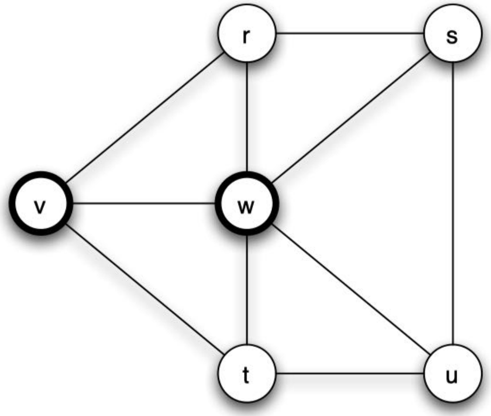
- **Assumptions for cascade analysis:**
    - Initially, everyone is using B (the old behavior/technology).
    - A small set S of initial adopters is "seeded" with A.
- The process: We then allow the network to evolve in discrete steps. In each round, all nodes with at least a q fraction of neighbors on A will switch to A (B → A). B-adopters with fewer than q fraction A-neighbors remain B.
- **Question:** Does this process result in A spreading to every node (a *complete cascade*) or does it stop at some point (a *partial cascade*)?
- **Answer:** It depends on three factors:
    i. The network structure (who is connected to whom),
    ii. The choice of initial adopters S (which nodes start with A),
    iii. The threshold q (based on relative payoff of A vs. B).

Example: Complete Cascade

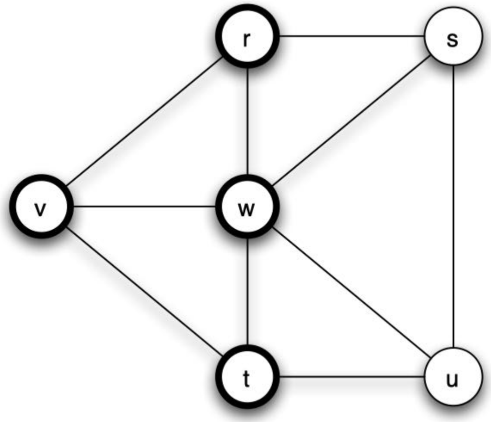
- Payoffs $a = 3, b = 2$. Then the threshold $q = \frac{b}{a+b} = \frac{2}{5} = 0.4$.
- Each node needs at least 40% of its neighbors on A to be willing to switch.
- Let the initial adopters be $S = u, v$ (two nodes that start with A).
- In the network shown (figure on the right), starting from S we observe that eventually **every node adopts A**. The innovation percolates through the entire network—this is a **complete cascade** at threshold q.



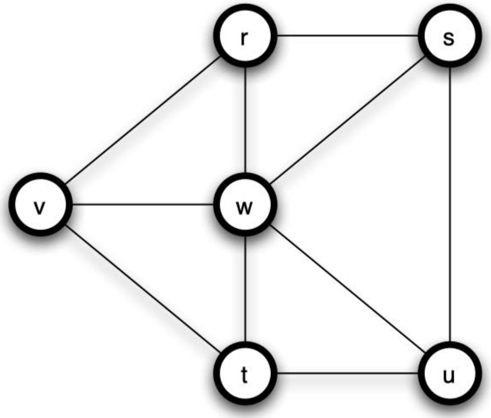
(a) *The underlying network*



(b) *Two nodes are the initial adopters*



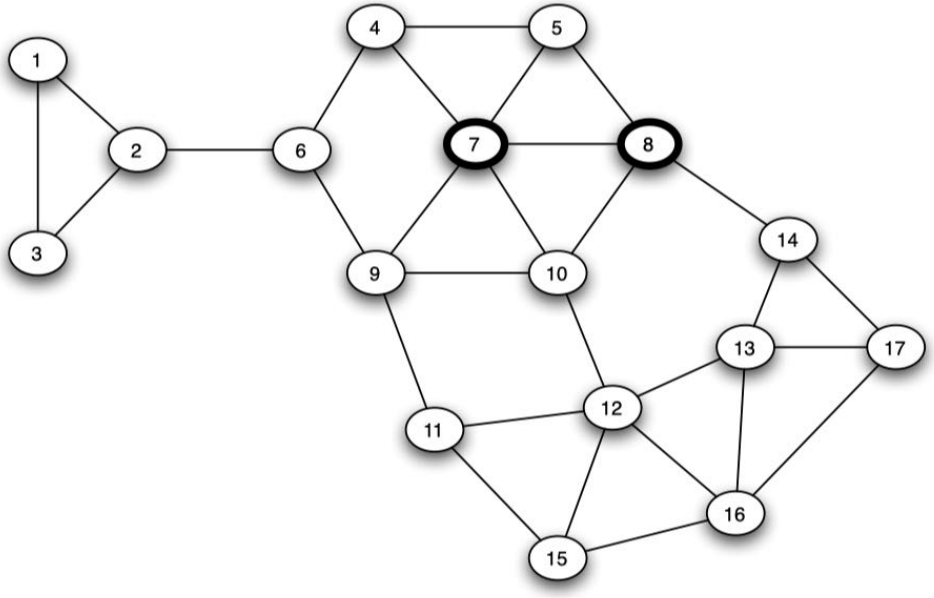
(c) *After one step, two more nodes have adopted*



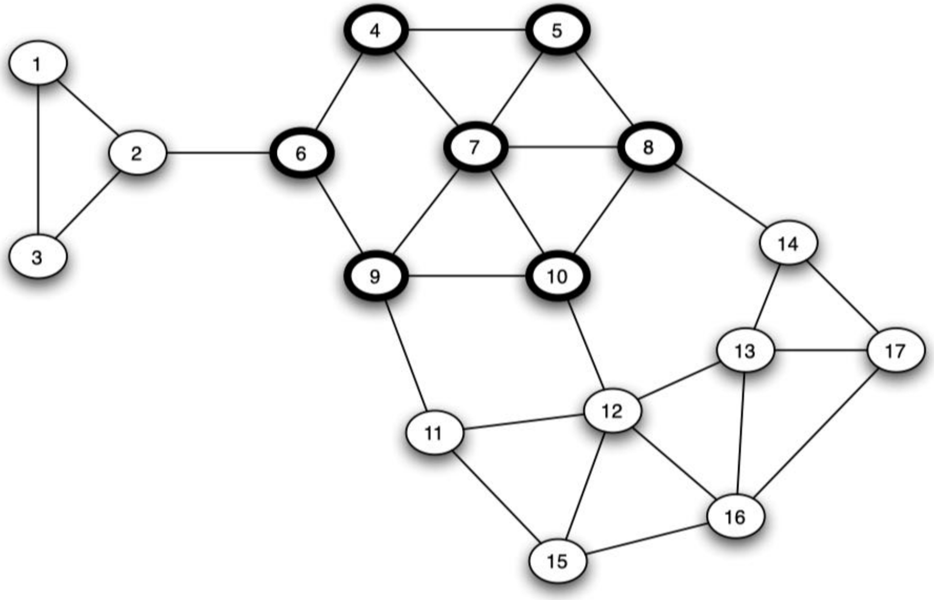
(d) *After a second step, everyone has adopted*

Example: Partial Cascade

- Same payoffs $a = 3, b = 2$ ($q = 2/5$) and initial seed $S = u, v$, but a slightly different network structure.
- In this case, A spreads to some additional nodes but **stops before reaching everyone**. The diffusion of A stops after a few rounds; some nodes remain using B indefinitely.
- This is a **partial cascade** – A gains some foothold but cannot overcome the resistant part of the network.



(a) *Two nodes are the initial adopters*



(b) *The process ends after three steps*

Recap: Cascade Dynamics

- We have a process where:
  - Start with a set S of initial adopters of A (all others initially B).
  - Iteratively, any node with at least a q fraction of neighbors on A will switch from B to A.
  - Nodes that switch to A stay with A thereafter (we assume no switching back to B in this model).
- **Complete cascade (at threshold q):** If eventually every node ends up adopting A (B → A everywhere). Then we say the initial set S *causes* a complete cascade at threshold q.
- **Partial cascade:** If the process converges to a mix of A and B (some nodes remain B), meaning A's spread reaches a limit and stops.
- Key insight: Whether a complete cascade happens depends on network structure relative to q. Highly interconnected clusters of B can resist if the cluster density is high (we formalize this next).
- This model captures phenomena like adoption reaching a "tipping point" and then taking over, versus stalling due to insufficient peer support.

Defining Clusters (as Obstacles)

- To analyze where cascades stop, we introduce the notion of a **cluster** in this context:
- A **cluster of density** p is a set of nodes in the network such that **every node in the set has at least a** p **fraction of its neighbors also in the set**.
  - In other words, each member's neighborhood overlaps significantly with the cluster itself (at least p of their neighbors are also in the cluster).
  - Example: A tightly knit community could form a cluster of high density (close to 1 if everyone is connected).
- Intuition: A cluster of high density means strong internal cohesion—members are connected mostly to each other, with relatively few outside neighbors.

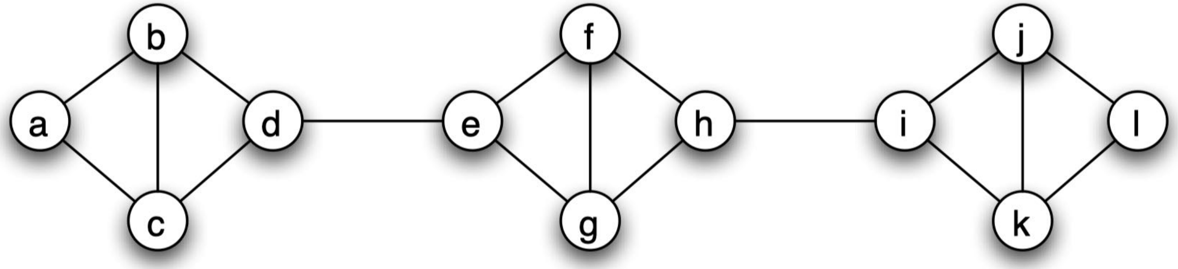
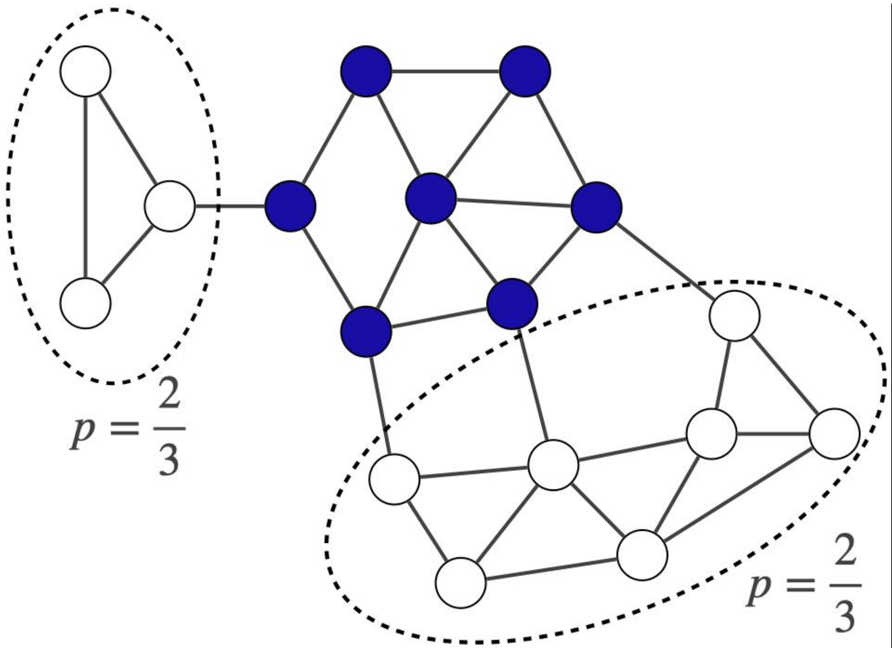


Figure 19.6: A collection of four-node clusters, each of density 2/3.

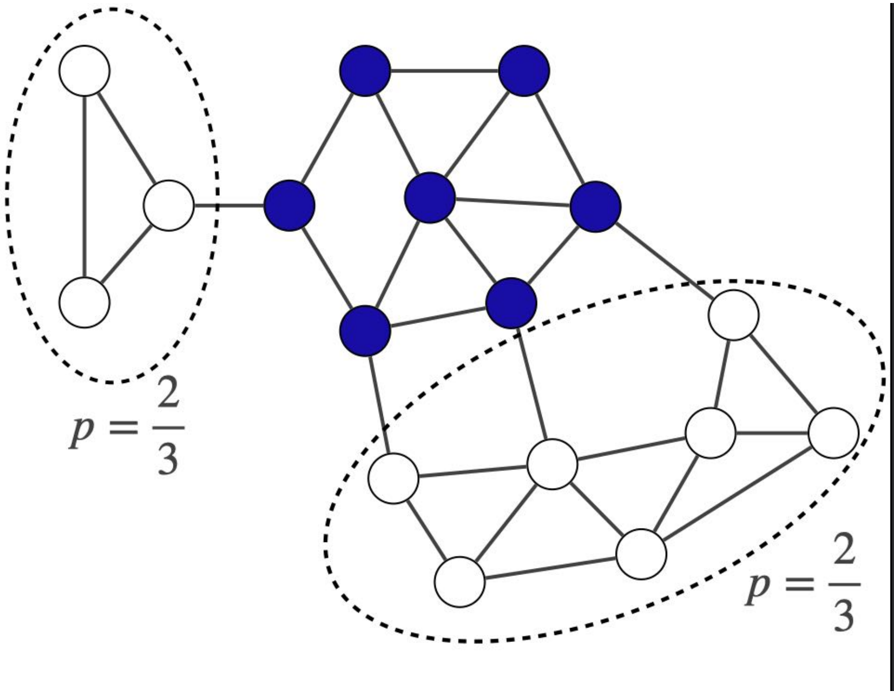
Stopping Cascades: Intuition

- **What prevents a cascade from spreading to the whole network?**
  - A cascade will fail to propagate if it encounters a subnetwork where newcomers (A-adopters) can't get a foothold.
  - **Homophily and community structure** can be barriers: an innovation introduced outside a tightly-knit community may have trouble penetrating that community.
- A tightly connected cluster of B-users can resist A if its internal links dominate external links.
  - Each member sees so many neighbors sticking with B that $p_A < q$ (fraction of neighbors on A stays below threshold).



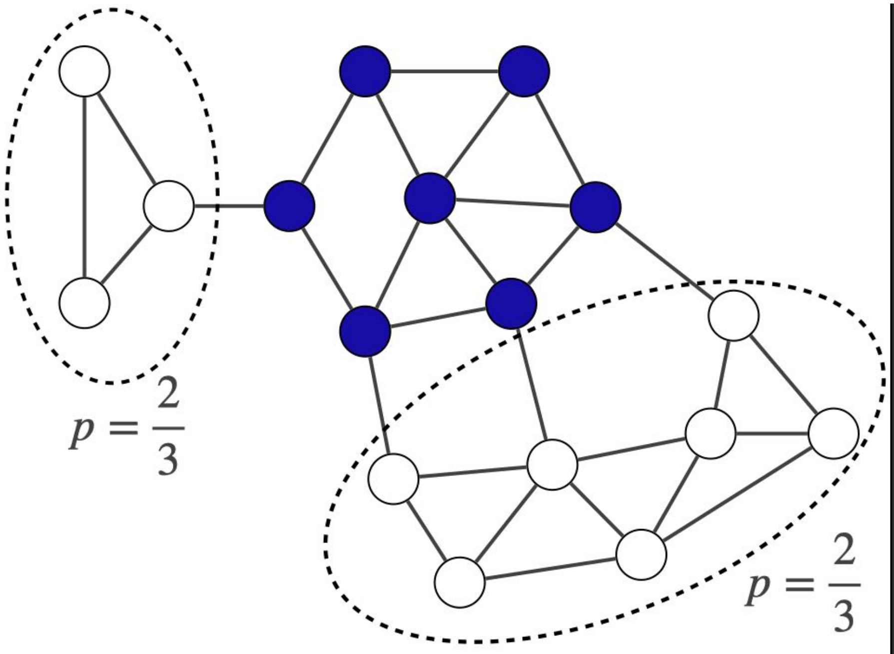
Stopping Cascades: Intuition (2)

- We quantify this with cluster density:
  - A **cluster of density** p (recall definition) provides "internal cohesion" for B: everyone in it sees at least p of neighbors inside the cluster (thus at most $(1-p)$ outside).
  - If $(1-p)$ (the max fraction of outside neighbors) is below the threshold q, then no one in the cluster will ever switch to A. Each member is always short of enough A-influence.



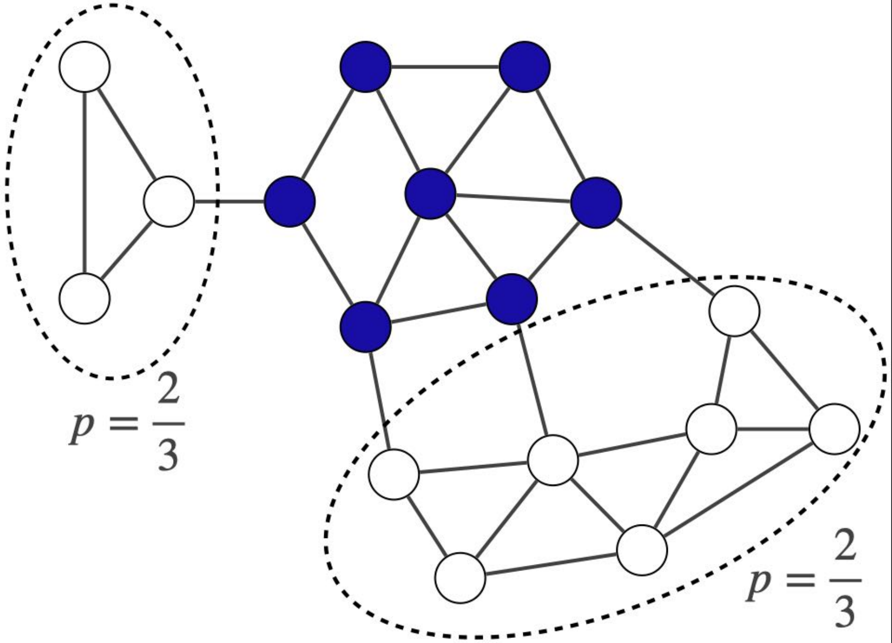
Internal Cohesion and Diffusion

- **Internal cohesion:** Nodes in a cluster support each other's current behavior. Each node has a significant fraction of friends within the same cluster (hence likely sharing the same behavior initially).
  - Note: Nodes in a cluster need not be identical or share an ideology; it's about network structure (who's connected to whom).
  - Extreme cases:
    - The entire network is a cluster of density $p = 1$ (trivial: everyone is interconnected).
    - The union of two clusters of density p is also a cluster of at least that density p. Clusters can exist at multiple scales overlapping.



Internal Cohesion and Diffusion (2)

- For diffusion: A cluster with high density can serve as an “immune system” against external contagion:
  - So long as each node in the cluster sees enough peers in B (the cluster’s behavior) to stay below threshold q for A, they will all resist switching.
  - More formally, if a cluster’s density p is greater than $(1 - q)$, then each member has fewer than q fraction neighbors outside. So even if all outside neighbors adopt A, it’s still not enough pressure.



Clusters and Cascades: Formal Relationship

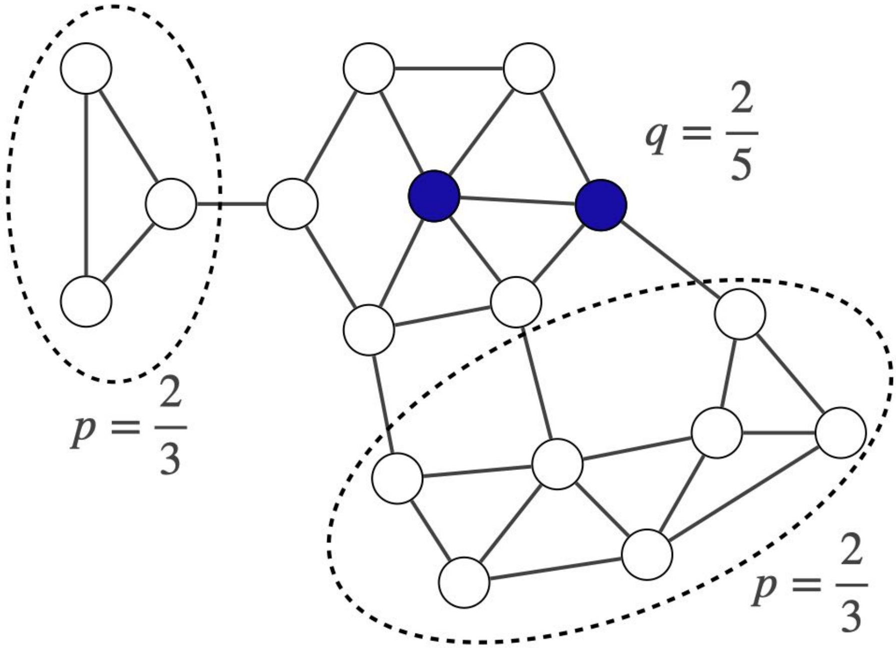
- **Key claim connecting clusters to cascades**:

  i. If the remaining network (after removing initial adopters S) contains a cluster of density $p > (1-q)$, then S **cannot trigger a complete cascade**.

     - Such a cluster has too strong internal support for B, halting the spread of A.

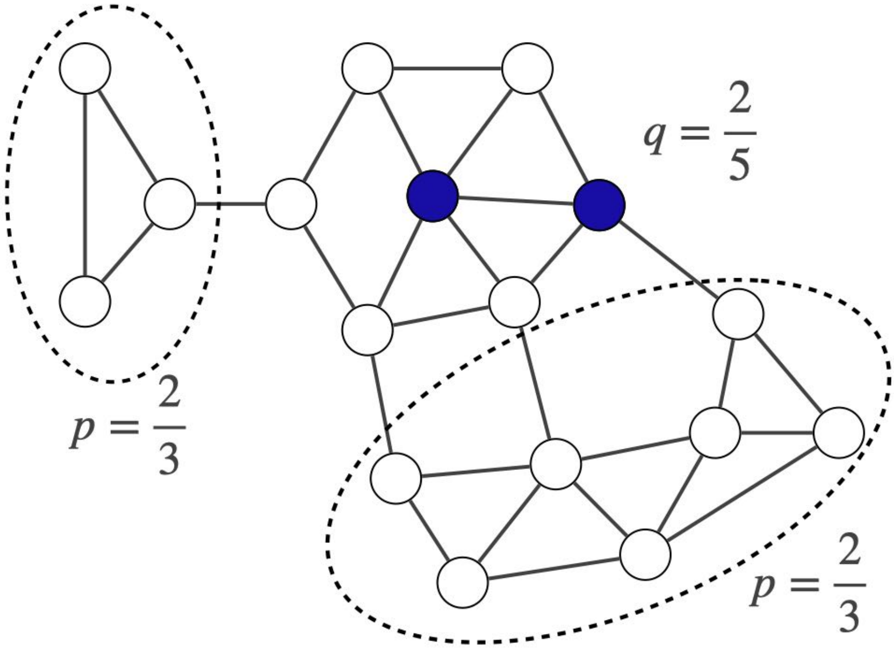
  ii. Conversely, if S fails to cause a complete cascade, then there must exist a cluster in the network (among those who remained B) with density $p > (1-q)$.

     - In other words, whenever a cascade stops short, it's precisely because it ran into a cluster dense enough to block it.



Clusters and Cascades: Formal Relationship (2)

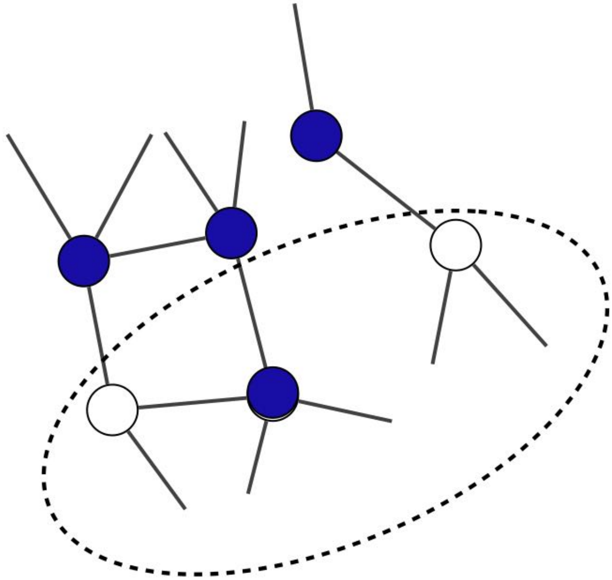
- Therefore, **clusters are the natural obstacles to cascades**:
  - A cluster of density greater than $(1 - q)$ is sometimes called a **blocking cluster** (it blocks A's spread).
  - This gives a structural criterion to evaluate if a given initial seed set S will fully take over or not.



Why Clusters Imply No Cascade

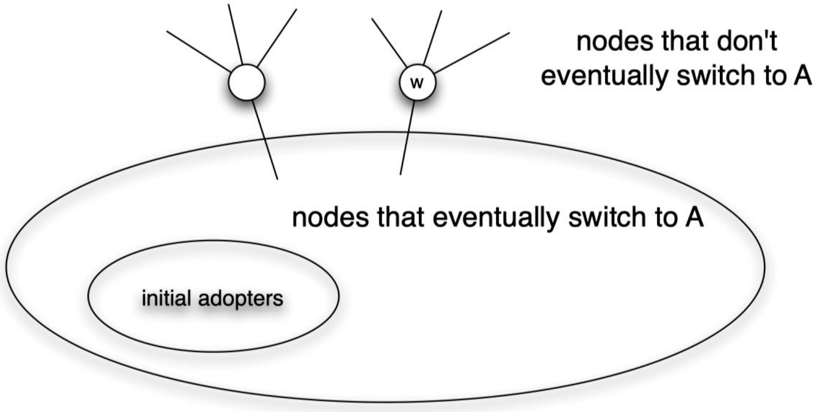
Proof Sketch (by contradiction):

- Assume there is a cluster C of density $p > (1-q)$ in the network (outside the initial adopters) but suppose, for contradiction, that A eventually invades C (meaning someone in C adopts A).
- Let v be the first node in cluster C to switch to A at time t. By definition of "first," at the previous time step $t-1$, no other node in C had adopted A yet.
- At time $t-1$, v's neighbors outside C were the only possible A neighbors (since no one inside C had A yet).
- Because cluster density $p > (1-q)$, the fraction of v's neighbors outside the cluster is less than $1-p < q$. So at $t-1$, the fraction of v's neighbors who had A was < q.
- Thus v should not have switched at time t (threshold rule wasn't met). This is a contradiction. Hence no one in C ever



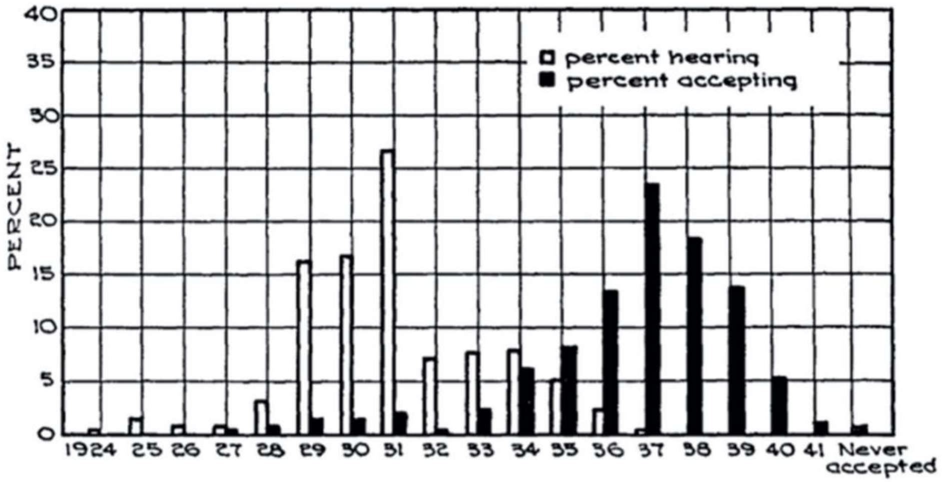
Clusters Imply No Cascade (Alternate View)

- Let S = initial adopters of A. Suppose the process stops and not everyone switches (so we end with some R remaining on B).
- By definition of termination, each node in R didn't switch because it didn't meet the threshold:
    - For each node $w \in R$, fraction of friends that adopted A $< q$.
    - Equivalently, fraction of friends still in R (sticking with B) is > $(1 - q)$.
- Thus, *every* node in R has more than $(1 - q)$ of its neighbors in R. That means R is a cluster of density $p > (1 - q)$.
- So either the cascade succeeds entirely, or the



Viral Marketing Implications

- **Lesson:** Tightly-knit communities can impede the spread of an innovation. Even if the innovation spreads in the rest of the network, a dense community (cluster) might hold onto the old behavior (coexistence of A and B).
    - Example: different regions favoring different technologies due to local network effects (e.g., SMS usage persisting in one region vs. instant messaging in another).
- If a company can improve the intrinsic quality of product A (increase a relative to b), the threshold $q = b/(a+b)$ drops. A lower q means even dense clusters might yield (they need an even smaller fraction of A neighbors to switch).
    - E.g., making the new technology significantly better (or compatible) can overcome resistance.
- If quality can't be easily changed, another strategy is **targeted seeding**: identify and **convince a few key individuals in the resistant cluster (still using B) to adopt A**.
    - By seeding inside the cluster, you break its cohesion from within.
    - These "key people" should be chosen based on network position – ideally well-connected in that cluster (to maximize impact). This connects to influence maximization problems.

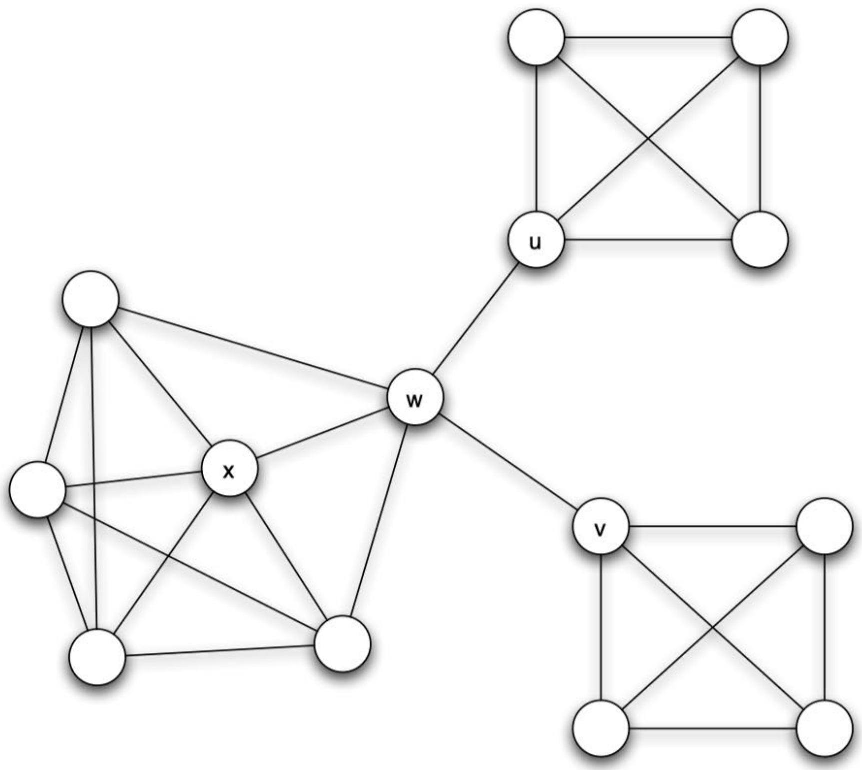


- **Historical example – Ryan & Gross (1943):** Studied adoption of hybrid seed corn in Iowa. Farmers learned about the innovation (many heard about hybrid seeds), yet adoption lagged until they saw enough peers try it successfully.
  - Demonstrates the gap between awareness and adoption – similar to individuals on the edge of a cluster who hear about A but don't switch because not enough of their close peers have switched.
- This is analogous to nodes on the boundary of a cluster: they get exposed to the new idea (from outside neighbors) but *still* don't adopt because their internal ties dominate and reinforce the old behavior.
  - In the seed corn study, even after almost all farmers knew about hybrid corn, some waited to see a larger fraction of their neighbors adopt before they did (a threshold behavior).
- **Takeaway:** Viral marketing must often overcome this reinforcement problem: simply getting the word out (information) isn't enough if people are waiting for social proof from their network.

Reading: Ryan, B. & Gross, N. (1943). "The diffusion of hybrid seed corn in two Iowa

The Role of Weak Ties

- Threshold models shed light on Granovetter's **"strength of weak ties"** theory:
  - **Weak ties** (acquaintanceships bridging different communities) are great for spreading simple information widely (news, gossip) because they connect distant parts of the network.
  - But for behaviors that require reinforcement (our threshold q models that need multiple confirmations), weak ties alone might not be enough to trigger adoption.
- Nodes connected by weak ties often get the *first news* of innovations from outside their circle, but one weak tie might not provide enough peer pressure to adopt if q is high. For example, if $q = 1/2$, a single friend (especially an outsider) isn't sufficient.



- **Strong ties** (close friends/family) often cluster together, providing redundant signals. Thus, strong ties can be more influential in causing actual behavior change, even if the idea originally came via a weak tie.
- Reference: Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology, 113*(3), 702–734. Their work showed that high clustering can facilitate diffusion of behaviors requiring social reinforcement.

Simple vs. Complex Contagion

- **Simple contagion:** One contact might be enough to transmit the "infection." Analogy: diseases or viral memes. In network models:
  - *Independent Cascade model:* each infected node gets a chance to infect each neighbor (one-to-one chance-based influence). This does not require multiple confirmations – just one successful transmission event.
  - *Simple threshold case:* effectively q is low enough (or probabilistic contagion) that a single activation can trigger you.
  - Weak ties are very powerful here: one connection can spread it onward.
- **Complex contagion:** Adoption requires multiple reinforcing signals. Our threshold model is a prime example:
  - If $q > 0.5$, you need more than one neighbor to convince you (a "critical mass" in your neighborhood).
  - Each additional friend adopting increases the likelihood you do – there's synergy or reinforcement.
- In complex contagion, *each new adopting neighbor has greater influence than the previous ones*. The first friend might not sway you, but the third or fourth might tip you over.
- **Key distinction:** Simple contagions thrive on network reach (many long-range weak ties = good). Complex contagions thrive on clustered reinforcement (multiple friends in the same circle adopting).

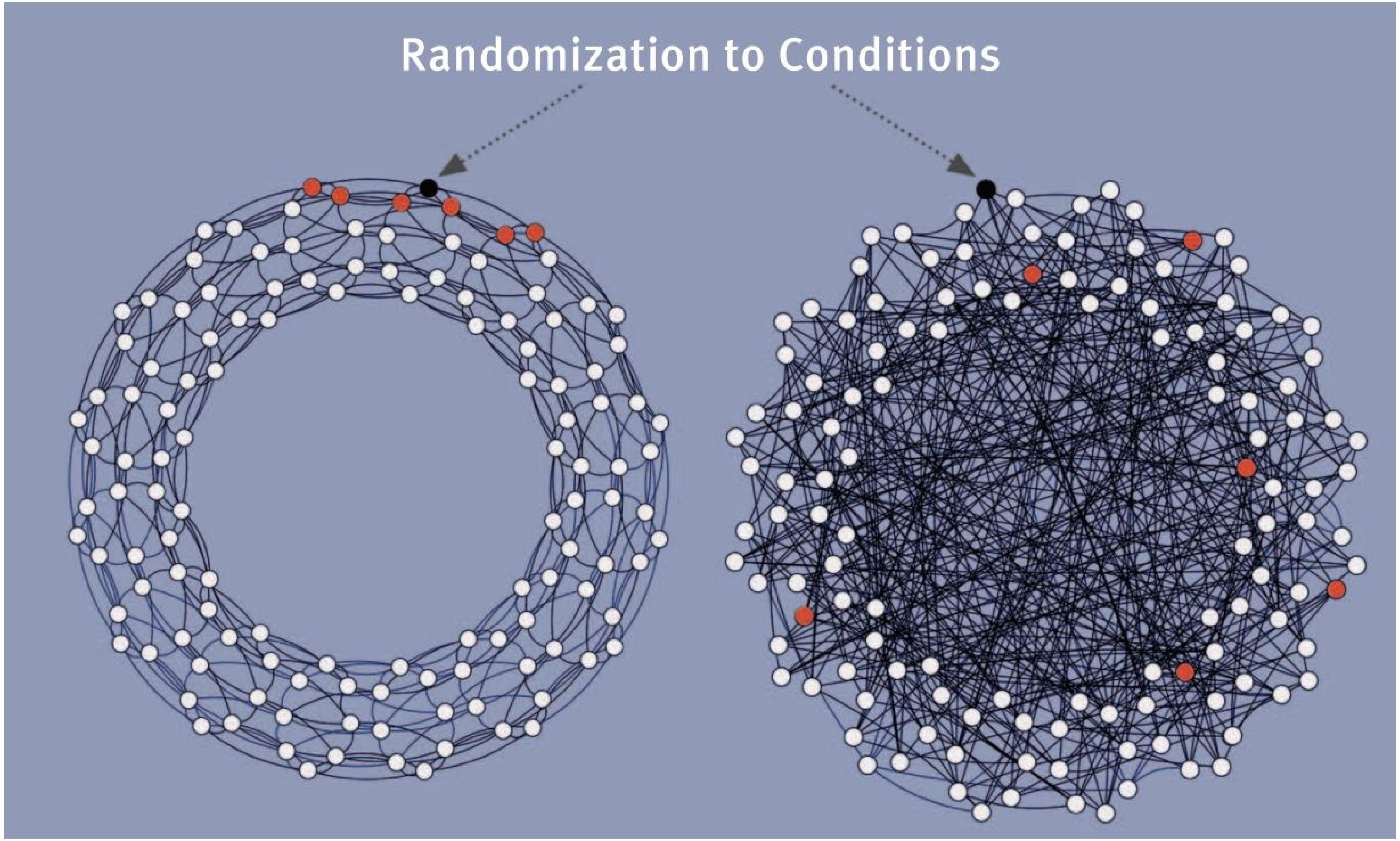
Health Behavior and Artificial Communities (Centola 2010 Experiment)

- Sociologist **Damon Centola (2010)** ran a landmark experiment on health behavior diffusion in online communities:
  - Recruited 1,528 participants interested in health, and created an online "health community."
  - Participants were randomly placed into one of two network conditions:
    a. A **clustered lattice** network (highly clustered, lots of local ties, longer path lengths).
    b. A **random network** (low clustering, many long-range ties, shorter path lengths akin to "small-world").
  - Each participant had a few "health buddies" (neighbors) in their network condition and could see their activities.
- Participants were asked to consider adopting a health behavior (registering for a health forum site).
  - When a participant's buddy adopted (registered), they'd get a notification (signal).
  - If multiple buddies adopted, they received multiple signals (one per adopting neighbor).

Health Behavior and Artificial Communities (Centola 2010 Experiment) (2)

- Thus, Centola's setup directly tests simple vs. complex contagion:
    - In the random network, individuals are likely to have *single* contacts from various parts of the network (weak ties).
    - In the clustered network, individuals might receive *redundant* signals from a tight-knit cluster of buddies.

Reading: Centola, D. (2010). "The Spread of Behavior in an Online Social Network Experiment." Science 329(5996):1194-1197.


Randomization to Conditions

Results of Centola's Experiment

Findings: Network structure had a significant effect on behavior diffusion.

- Adoption reached a significantly **higher fraction of people in clustered networks** than in random networks. Example: in one condition ~54% adoption in clustered vs ~38% in random.
- **Speed:** The behavior diffused faster in clustered networks as well. On average, diffusion in clustered nets was more than four times faster than in random nets.
- *This is opposite to what we'd expect for a simple contagion.* It supports the complex contagion idea: redundant signals spur adoption.
- At the individual level, **redundant signals greatly increased adoption probability**. If you had two or three friends adopting, you were much more likely to follow than if you had just one.
    - They measured a *hazard ratio* g per additional signal, indicating how much each extra adopting neighbor multiplied the baseline probability of adoption.

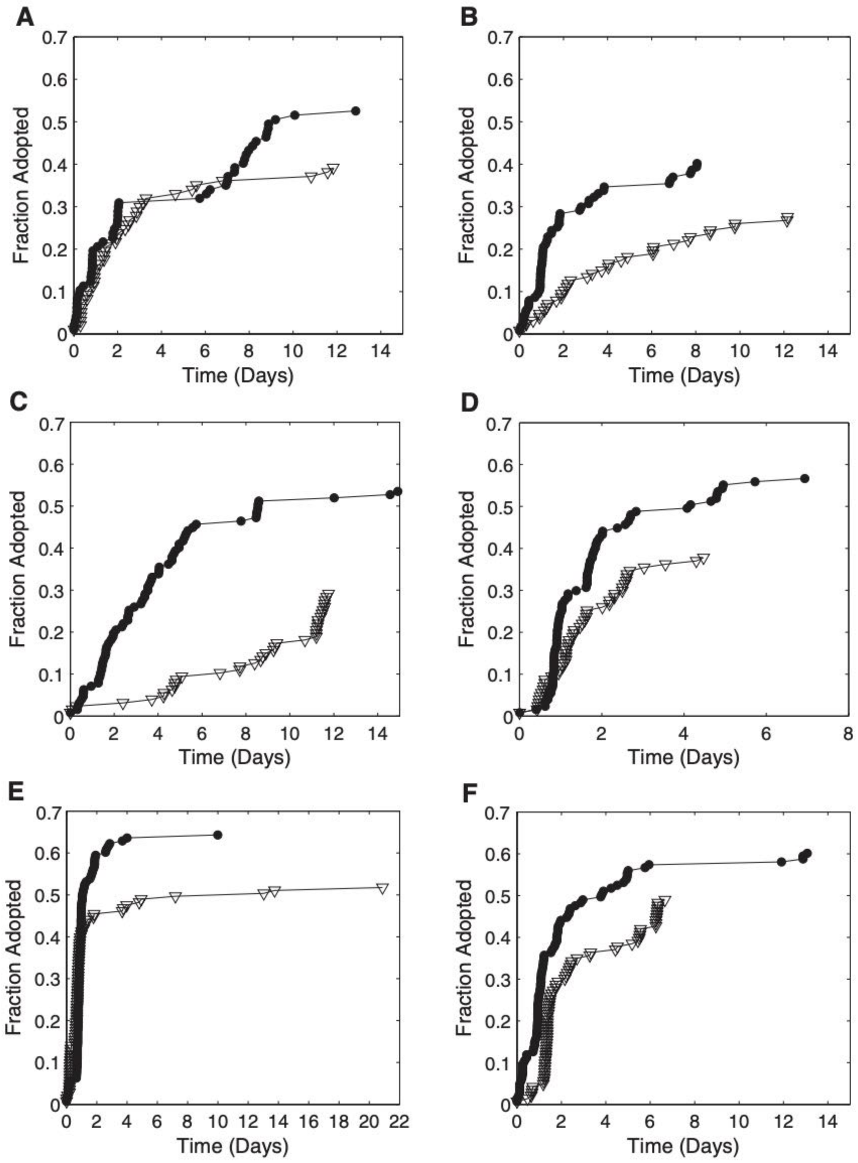
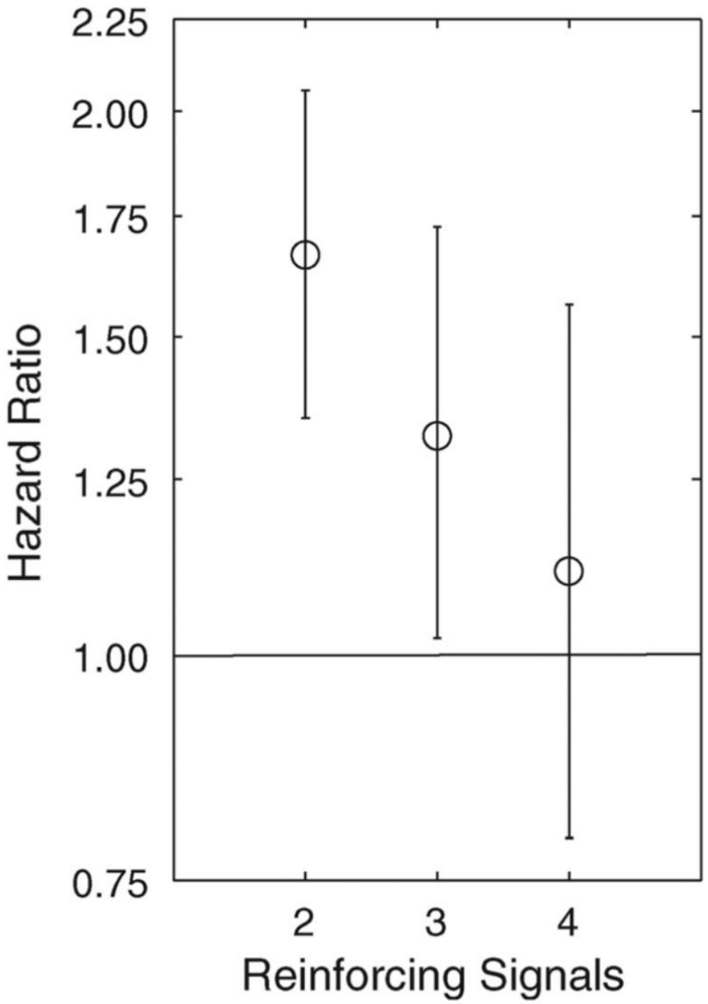
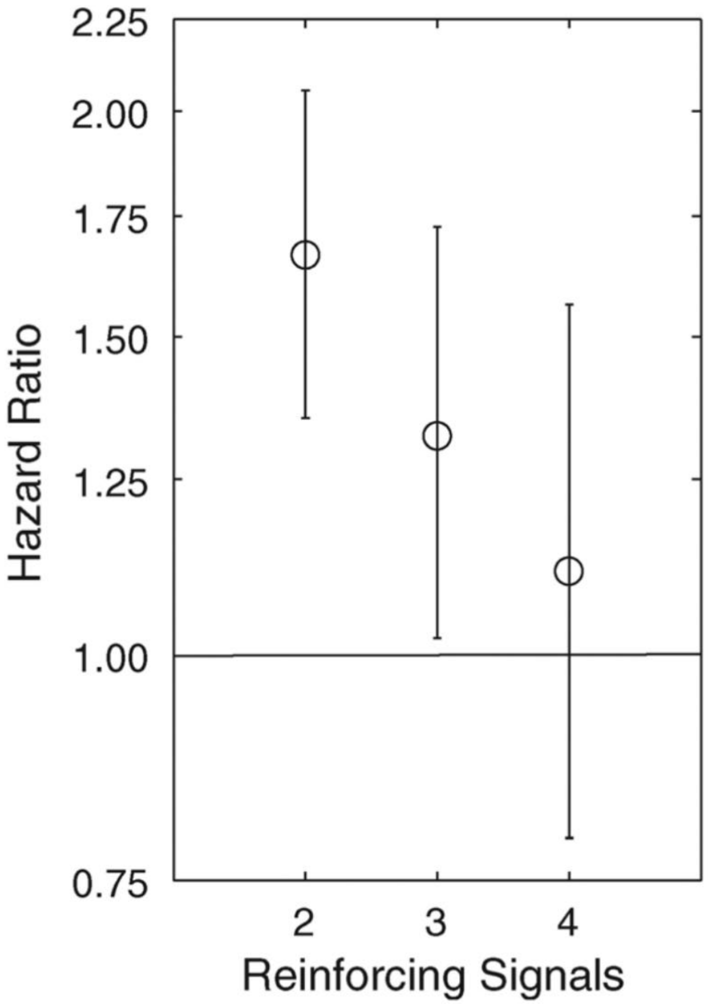


Fig. 2. Time series showing the adoption of a health behavior spreading through clustered-lattice (solid black circles) and random (open triangles) social networks. Six independent trials of the study are shown, including (**A**) $N = 98$, $Z = 6$, (**B** to **D**) $N = 128$, $Z = 6$, and (**E** and **F**) $N = 144$, $Z = 8$. The success of diffusion was measured by the fraction of the total network that adopted the behavior. The speed of the diffusion process was evaluated by comparing the time required for the behavior to spread to the greatest fraction reached by both conditions in each trial.

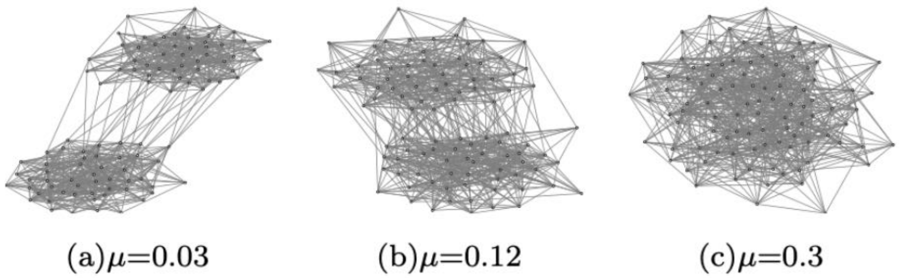
- The **hazard ratio** g (from Centola's paper) quantifies how additional social signals affect adoption:
  - g > 1 indicates each extra neighbor adopting multiplies your chance of adopting. For example, if $g = 3$, two neighbors adopting makes you 3 times as likely as with one neighbor.
  - In the experiment, g was significantly > 1, confirming **social reinforcement**: each extra adopting friend sharply raised the probability of adoption.
- **Conclusions**:
  - Individual adoption is improved by having reinforcing signals from multiple social ties.
  - System-level: clustered networks (which provide those reinforcing ties) achieved larger and faster diffusion compared to random networks.
  - For simple contagions (like info or disease), redundant ties are unnecessary. But for complex contagions (behavior change, social movements), redundancy can make diffusion *more efficient*.

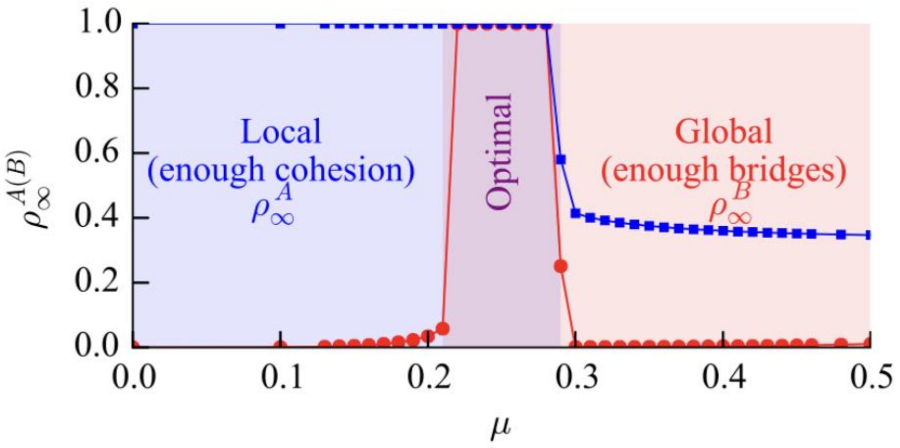


- **Implication for interventions:** To promote a new health behavior, it might be more effective to seed in a clustered community (ensuring people see multiple neighbors adopt) than to rely on a few influencers spreading the word broadly.
    - E.g., target tightly-knit residential communities for interventions like exercise programs or healthy eating, rather than hoping a mass media campaign alone will do it.



Finding the Optimal Clustering for Spreading





- There is a **trade-off** between intra-community spreading and inter-community spreading:
  - **High modularity / strong communities (low inter-connectivity, small μ):** Great for *within* community spread (local cascades easily among similar nodes) but poor for jumping to other communities.
  - **Low modularity / weak communities (lots of bridging ties, large μ):** Easier for something to jump between communities (global reach), but without local reinforcement it might not fully take off in each community.

- Research by Nematzadeh et al. (2014) suggests an **optimal intermediate level of modularity** for diffusion:
  - At this optimal point, communities are clustered enough that social reinforcement works within them, yet enough inter-community links exist that once one community "ignites," others can catch the flame.
  - Global diffusion requires minimal seeds at this sweet spot of modularity.
- In the figure:
  - μ could represent fraction of edges that are between communities.
  - ρ_∞^A (and ρ_∞^B) represent final fractions adopting A (or B) in communities A and B.
  - Notice how intermediate μ yields both communities ending with high ρ_∞^A.
- **Takeaway:** Neither a completely fragmented network nor a completely random network is best for complex contagions. Some community structure helps, but too much creates blocking clusters. There's an optimal balance.

Reading: Nematzadeh et al., "Optimal Network Modularity for Information Diffusion" (Phys. Rev. Lett. 113, 088701, 2014).

Extensions of the Basic Cascade Model

Heterogeneous Thresholds

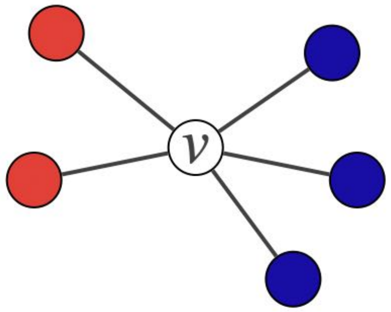
- So far, we assumed everyone has the same payoff values a, b, hence the same threshold $q = b/(a+b)$. In reality, individuals value behaviors differently:
  - Some people might require fewer neighbors to adopt before they do (lower q – more *easily influenced* or find A very appealing).
  - Others need a lot of friends doing it before they switch (high q – *stubborn* or find A only marginally better than B).
- We can extend the model: each person v has their own payoff a_v for A and b_v for B, giving them a personal threshold $q_v = \frac{b_v}{a_v+b_v}$.

Heterogeneous Thresholds (2)

- Now the population is heterogeneous:
  - "Early adopters" might have low q_v (they switch with very little peer support or even unilaterally if $a_v > b_v$ a lot).
  - "Laggards" might have q_v near 1 (they need almost everyone around them to adopt first).
- This is closer to reality and connects to diffusion of innovations theory (innovators vs. early majority, etc., have different thresholds in effect).

v \ w	A	B
A	a_v, a_w	$0, 0$
B	$0, 0$	b_v, b_w

Heterogeneous Thresholds (Formalism)



- For a given node v:
  - Let p = fraction of neighbors adopting A (as before),
  - d = number of neighbors,
  - Payoffs: a_v, b_v (node's subjective benefit for coordinating on A vs B with a neighbor).
- v chooses A if
  $p \cdot d \cdot a_v \geq (1-p) \cdot d \cdot b_v,$
  $p \geq \frac{b_v}{a_v+b_v} = q_v.$
- Thus each node has its own threshold q_v. The diffusion dynamic now depends on the distribution of thresholds across the network.

Blocking clusters with heterogeneous q

- Watts and Dodds (2007) argued we should consider not just “influentials” with low q, but also the network of easily influenced people.
- A cluster generalization: A set C is a blocking cluster if every node v in C has more than $(1 - q_v)$ fraction of neighbors in C.
  - That is, each node is well supported given *its own* threshold.
- This is trickier since one very stubborn person (q_v near 1) can raise the needed density for the cluster.
- Takeaway: Heterogeneity can either help or hinder cascades. A few low-threshold nodes (social “spark plugs”) can get things going, but a few high-threshold individuals can hold up a region unless others circumvent them.
- Large cascades often depend more on a critical mass of easily influenced individuals than a few super-influencers

Reading: Watts, D. & Dodds, P. (2007). “Influentials, Networks, and Public Opinion Formation.” J. of Consumer Research 34(4): 441–458.

Knowledge, Thresholds, and Collective Action

Integrating Network Effects at Different Levels

- Not all diffusion processes are purely local. Especially in collective actions (e.g., protests, revolutions), people consider:
  - **Population-level perception:** "How many people overall do I think will join?"
  - **Local network cues:** "How many of my close contacts do I know are joining?"
- Topics:
  - **Collective action & pluralistic ignorance:** Sometimes everyone is waiting for everyone else.
  - **A model of knowledge effects:** People have thresholds for participation (need to expect a certain count).
  - **Common knowledge and institutions:** Society often establishes common knowledge (everyone knows that everyone knows...) to facilitate coordination (e.g., public announcements, media).
- These ideas extend our cascade model by including more global signals or knowledge beyond immediate neighbors, blending game theory and network theory.

Collective Action and Pluralistic Ignorance

- **Scenario:** Organizing a revolt under a repressive regime. Success depends on enough people participating (critical mass). But you can only safely communicate with a few trusted contacts.
- **Collective action problem:** You benefit from participating only if enough others also participate (otherwise it fails and you're at risk). This is like a threshold: "I'll join if I expect at least K others will join."
- Each person might have a personal threshold k (number of people they need to believe will protest for them to protest).

UNIVERSITÀ DI TORINO

Collective Action and Pluralistic Ignorance (2)

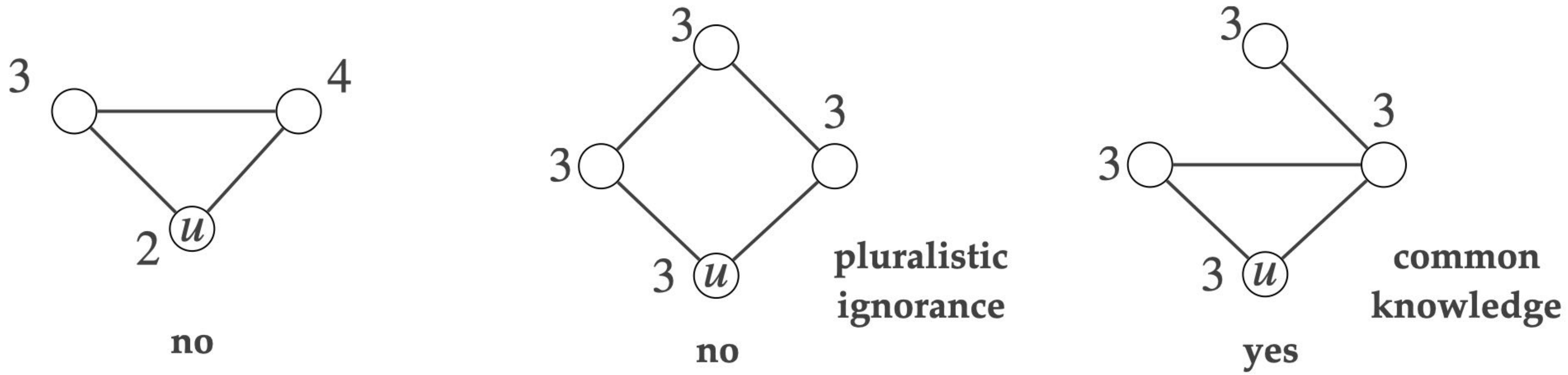
- **Pluralistic ignorance:** A situation where people vastly misjudge the popularity of an opinion or intention in the broader population.
    - E.g., Everyone might be unhappy with the regime (privately willing to protest if others would), but each thinks others are content or too afraid, so no one acts.
    - Essentially, lack of common knowledge of shared grievances.
- In networks, you only talk to a few friends. You might not realize how many others share your sentiment.
- This can freeze action: everyone is waiting for a signal that others will step up.

A Model: Knowledge and Collective Action

- Model each individual with threshold k_i: "I will protest if I am sure at least k_i people (in total) will protest".
- Assumption: You know your neighbors' thresholds (friends you talk to), but not the whole population's.
  - You have limited local info. How do you infer if k_i will be met globally?
- Think of a small network:
  - You and your friends exchange commitments or hesitations. If your friend has threshold 5 and you know 5 including them are ready, you can reassure them.
  - Scenarios to consider:
    - If two friends each need 10 people but they only see each other, none will move.
    - If one friend is an "instigator" needing only 2, they might trigger others if they communicate their intent.

A Model: Knowledge and Collective Action (2)

- This model highlights that network structure (who knows whose willingness) can determine if a cascade of participation happens. It's not just the distribution of thresholds but who observes whom.



- *Question:* In the above configuration, if u needs to know 3 people will join and only knows two neighbors, can u safely join?
    - This requires reasoning about what u thinks those neighbors will do, and what those neighbors think.

Common Knowledge and Social Institutions

- Many social institutions exist to create **common knowledge** — where *everyone knows that everyone knows* some information.
- Examples:
  - A widely-publicized speech or a front-page news story: not only do people get the message, they also know millions of others got it too.
  - This changes potential actions: e.g., a public call to protest in a newspaper means any one reader knows others saw it. If they show up, they expect others might too.
  - The famous Apple "1984" Super Bowl ad: not only advertising the Mac, but ensuring 90 million people see it simultaneously, creating buzz partly because *everyone knows everyone saw that iconic ad* (a coordination around the product's hype).

Common Knowledge and Social Institutions (2)

- **Common knowledge enables coordination:** It's a way to overcome pluralistic ignorance. If I know others got the same message, I'm more willing to act on it together.
- Thus:
  - **Freedom of press and assembly:** Crucial for enabling people to coordinate (you can gauge the true level of dissent/support in society).
  - **Marketing strategies:** Big launch events or ads during popular broadcasts aim for common knowledge. (Apple's 1984 ad during the Super Bowl, mentioned above, made a splash because it was a shared cultural moment).
- In network terms, common knowledge shortcuts the network – it's like making the whole group a clique where everyone "heard" each other get the info.

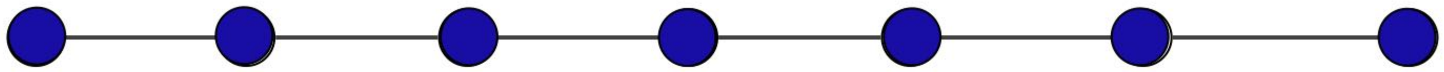
The Cascade Capacity

Infinite Networks

- Thus far we talked about finite networks. For a broader theory perspective, consider diffusion in infinite networks:
  - Helps us reason about long-term or large-scale limits of cascades.
- **Cascade capacity:** For a given infinite network structure, the cascade capacity is the **maximum threshold q such that a small finite set of initial adopters can lead to a complete cascade (adopt A) throughout the network**.
  - Essentially, how resilient is the network to new innovations? If the threshold is below this capacity, even a tiny seed can take over.
  - If the threshold is above this, no finite seed can cause full adoption.

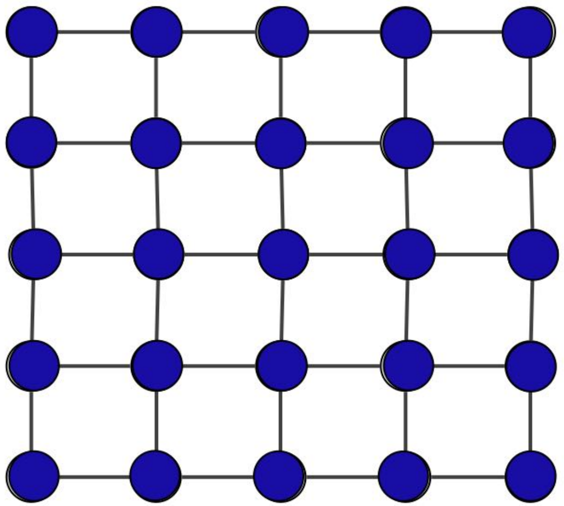
The Cascade Capacity (2)

$$q \leq \frac{1}{2}$$

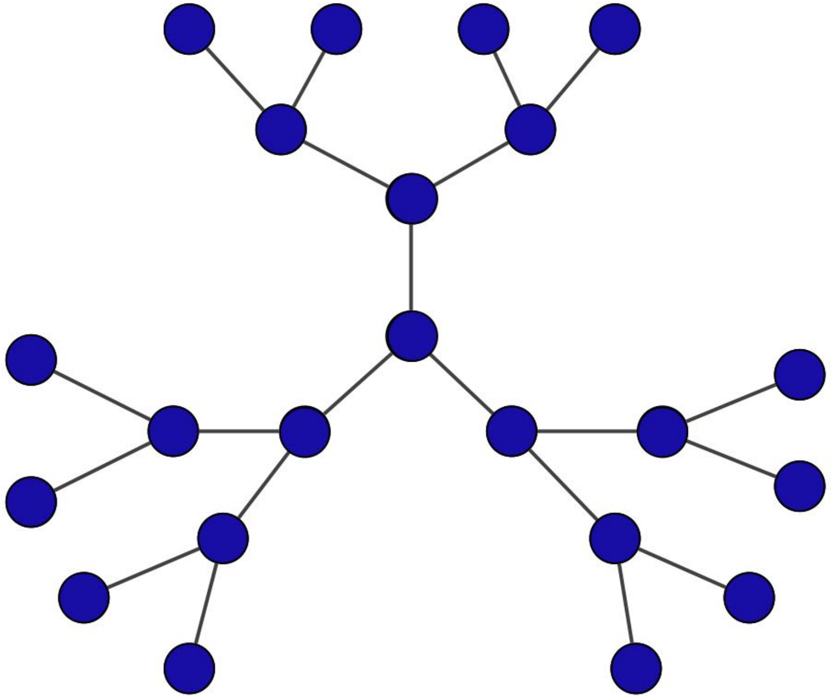


- We usually assume each node has finitely many neighbors (so we don’t get weird infinities locally).
- Example: **Infinite path (1-D lattice):** Each node has 2 neighbors (except initial ones).
  - Intuition: What q allows a finite seed to spread infinitely far? If $q < 0.5$, a single initial adopter can eventually infect the line (two neighbors eventually, etc.). If $q > 0.5$, two neighbors might not be enough to flip an intermediate node.
  - This can be formalized by seeing how a cluster (or a “block”) can stop the spread.

Infinite grid $q \leq \frac{1}{4}$



Infinite tree $q \leq \frac{1}{3}$



How Large Can Cascade Capacity Be?

- **Claim:** There is no network (no matter how cleverly structured) with cascade capacity more than $1/2$.
  - In other words, you cannot have a network where a cascade can happen if everyone needs a majority of their neighbors to convince them. Majority-needed ($q > 0.5$) cascades always fizzle out unless you initially flip infinitely many people.
- **Proof idea:** Suppose for contradiction a network G has cascade capacity > 1/2, meaning say $q = 0.6$ works – a finite seed can cause a complete cascade.
  - Then consider the "interface" between A adopters and B holdouts as the cascade progresses. We will argue this interface area must shrink each step if $q > 0.5$, meaning the cascade eventually stops.

- For $q > 0.5$, when a B node switches to A, it means it had *more* neighbors in A than in B (since $q > 0.5$ requires > half neighbors in A).
    - When that happens, think of edges between A and B regions (the interface). That B node flipping to A *removes some B-A edges and adds more A-A edges*, effectively reducing the boundary.
    - More formally, each newly flipped node has more connections to the A side than B side, so the number of B–A edges drops.
- If in each step the interface (the cut between A region and B region) strictly decreases, the process can only have a finite number of steps (it can't go on forever because you can't keep losing edges beyond zero).
    - Thus the cascade stops after at most as many steps as the initial interface size.
- This contradicts the assumption that a full cascade occurs for $q > 0.5$ (since full cascade on an infinite network would need infinite steps). Hence no network can have cascade capacity > 0.5.

Observations on Cascade Capacity

- If $q > 0.5$, **no finite seed can cause a complete cascade in an infinite network**.
  - This also suggests in very large finite networks, if more than half of each person's neighbors need to adopt, you won't see global cascades from small seeds (you'd need a very large initial push).
- This threshold $q = 0.5$ corresponds intuitively to requiring a majority of friends to convince someone. If everyone needs a majority, you can't start from a minority seed without external help.
- By contrast, if $q \leq 0.5$, it's possible to have cascades (depending on structure). Many real innovations probably have q well below 0.5 (e.g., if the innovation is slightly better or there are early adopters who inherently prefer it).
- This result is actually quite general: **no network topology** (no clever graph structure) can overcome the > 50 barrier for cascade from infinitesimal seeds. It's a fundamental limit.
- It tells firms or activists: if your product or idea isn't attractive enough that people will go for it even with less than half their friends on board, then you *must* convert more than a small fraction directly (or improve its appeal).

Another Extension: Compatibility (Bilingual Option)

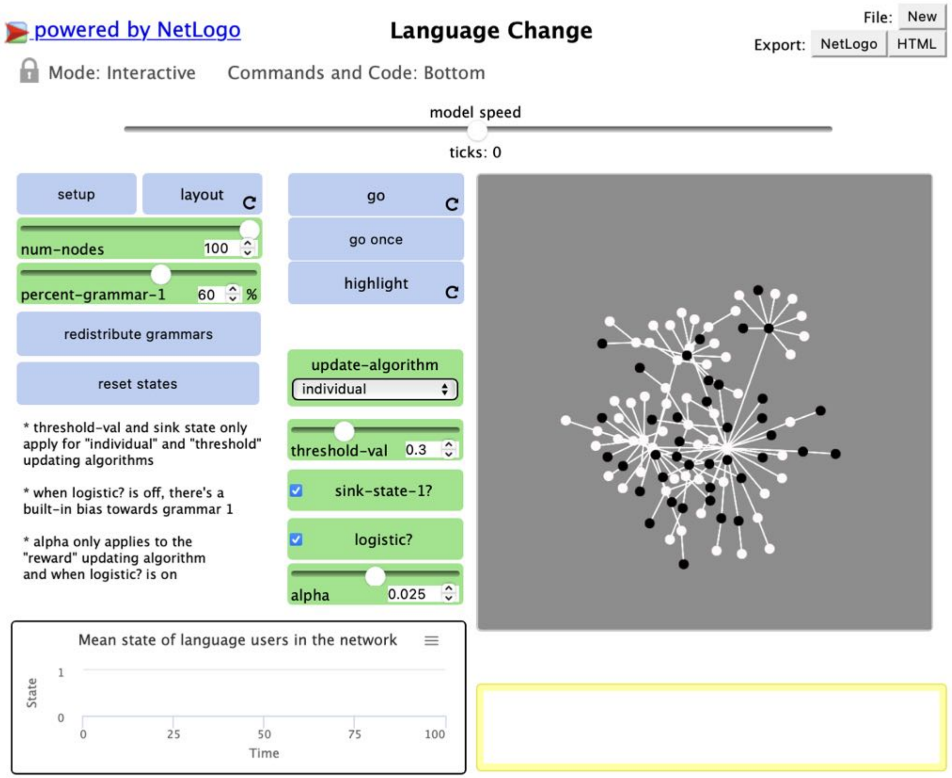
- Realistically, people can sometimes adopt a new behavior **without completely dropping the old one**. Think of technologies that can coexist:
  - E.g., one can use **both** WhatsApp and traditional SMS, or both a new streaming service and cable TV.
  - In language or convention terms, people can be **bilingual** (use A and B).
- Extension: nodes can choose A, B, **or both (AB)**. This is a form of *partial adoption* or compatibility.
  - E.g., a new product is backward-compatible with the old one, so you can interact with either.
- This is called the **bilingual option** in the literature.
- Surprisingly, even this simple addition can lead to complex effects:
  - Sometimes offering a compatible option can accelerate the adoption of A (since people don't have to risk losing B, they can adopt A gradually).
  - Firms may use compatibility strategically to enter a market dominated by an incumbent, then gradually phase out the old tech.
- We won't go deep into equations, but conceptually:
  - AB users get payoff with both A and B neighbors (some reduced payoff possibly).
  - They can serve as "bridges" in the network, facilitating a later full shift to A if A proves better.

Language Change Model

- A classic example of compatibility is bilingualism in language adoption:
  - Imagine two languages A and B. People can speak either or both.
  - Bilingual individuals (AB) can communicate with both A-monolingual and B-monolingual individuals.
- Typically findings:
  - If there's even a small benefit to coordinating on a single language, bilinguals can act as a catalyst: they adopt the new language A while still interacting with B speakers, gradually tipping B speakers to add A as well, and eventually everyone might switch to A.
  - If A has higher payoff ($a > b$), even a few initial bilinguals can start an increase of A usage.
  - But if bilingualism has a cost or if q remains high, people might stick to B unless many go bilingual.
- **Strategic insight:** Compatibility (bilingual option) can lower the effective threshold for adoption because people don't have to abandon B immediately. It smooths the transition.

Language Change Model (NetLogo Simulation)

NetLogo's Language Change allows playing with parameters like payoff advantages and initial bilingual fraction.



Reading Material

[NS2] Chapter 19 of "Networks, Crowds, and Markets" by D. Easley & J. Kleinberg – *Cascading Behavior in Networks*. Sections 19.1–19.5 cover diffusion games, cascades, clusters, and extensions, with examples and proofs.

Q&A