# Difusión en redes: Contagio complejo

Aníbal Olivera Morales

- 1. Modelos de Umbral para Acción Colectiva
- 2. Contagios Complejos y la Debilidad de los Lazos largos
  - 1. De lazos débiles a mundos pequeños
  - 2. Efectos de lazos largos para contagio complejo
- 3. Difusión de Innovaciones con Preferencias Individuales: Elección Racional vs Influencia Social

+ Interludio: Transiciones de fase

Modelos de Umbral para Acción Colectiva

#### **Contagios Complejos**

Granovetter 1978:

"Toma de decisiones colectivas bajo interdependencia"

→ Los individuos enfrentan una decisión binaria (actual/no actuar), donde la utilidad depende de la elección del resto.

#### Threshold Models of Collective Behavior<sup>1</sup>

Mark Granovetter
State University of New York at Stony Brook

Models of collective behavior are developed for situations where actors have two alternatives and the costs and/or benefits of each depend on how many other actors choose which alternative. The key concept is that of "threshold": the number or proportion of others who must make one decision before a given actor does so; this is the point where net benefits begin to exceed net costs for that particular actor. Beginning with a frequency distribution of thresholds, the models allow calculation of the ultimate or "equilibrium" number making each decision. The stability of equilibrium results against various possible changes in threshold distributions is considered. Stress is placed on the importance of exact distributions for outcomes. Groups with similar average preferences may generate very different results; hence it is hazardous to infer individual dispositions from aggregate outcomes or to assume that behavior was directed by ultimately agreed-upon norms. Suggested applications are to riot behavior, innovation and rumor diffusion, strikes, voting, and migration. Issues of measurement, falsification, and verification are discussed.

#### BACKGROUND AND DESCRIPTION OF THE MODELS

Because sociological theory tends to explain behavior by institutionalized norms and values, the study of behavior inexplicable in this way occupies

#### **Granovetter 1978**

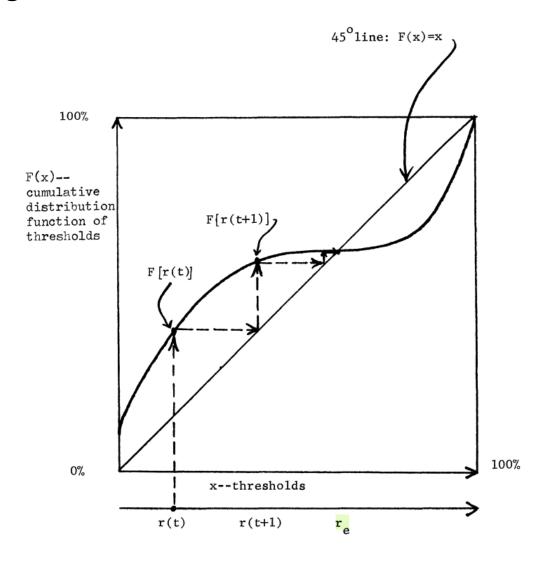
- fracción mínima de adoptantes que i • Preferencia individual:  $x_i \in [0,1]$ . necesita observar para actuar.
- Regla de decisión:

$$a_i(t+1) = egin{cases} 1 & ext{si } r(t) \geq x_i, \ 0 & ext{si } r(t) < x_i, \end{cases}$$

donde 
$$a_i(t) \in \{0,1\}$$
 y  $r(t) = rac{1}{N} \sum_{j=1}^N a_j(t)$ .  $igwedge$  proporción actual que ha actuado.

- Agregación: r(t+1) = F(r(t)).
  - Esto captura TODA la dinámica.
- Equilibrio:  $r^* an que F(r) = r$

#### **Granovetter 1978**



#### **Granovetter 1978 - Agregación ≠ preferencias individuales**

- Hay 100 trabajadores, cada uno con umbral  $x_i = \frac{i}{100}, \quad i = 0, 1, \dots, 99.$
- La dinámica de adopción está dada por

$$r(t+1)=F(r(t)),\quad r(0)=0$$

• Si iteramos:

$$r(0)=0, \quad r(1)=F(0)=0.01, \quad r(2)=F(0.01)=0.02, \ldots, r(100)=1.$$

$$r^*=1.$$

#### **Granovetter 1978 - Agregación ≠ preferencias individuales**

- Ahora quitamos el trabajador con  $x_i=0.01\,$  y lo reemplazamos por  $x_{i_2}=0.02\,$
- La dinámica de adopción está dada por

$$r(t+1) = F(r(t)), \quad r(0) = 0$$

• Si iteramos:

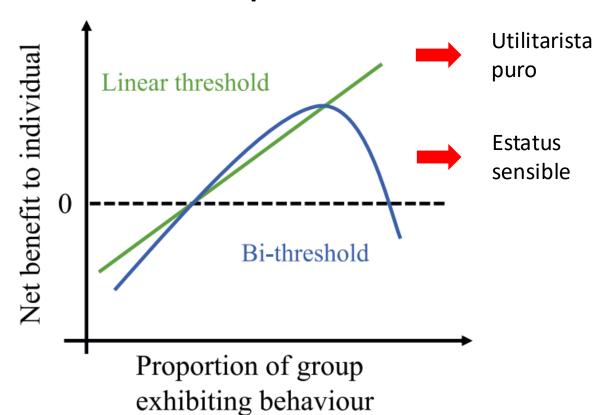
$$r(0)=0, \quad r(1)=F(0)=0.01, \quad r(2)=F(0.01)=0.01.$$
 $ightharpoonup r^*=0.01.$ 

• Preferencias individuales *iguales* producen outputs sociales opuestos:

$$\mathbb{E}\left[X_A
ight] = 0.495, \quad \mathbb{E}\left[X_B
ight] pprox 0.496,$$

#### **Granovetter 1978 - Detalles**

• El umbral **no** es preferencia, sino que es el **punto** en que la **utilidad neta** de actuar se vuelve **positiva**.



"El umbral es **una medida conductual contingente**, no una medida pura de preferencias normativas".

#### **Granovetter 1978 - Detalles**

• No captura influencias desiguales en la acción colectiva.

Los enlaces son **relacionales**, porque lo que importa no es solo cuántos participan, sino **quiénes** participan.

• El modelo asume mezcla completa: todos observan a todos por igual.

Esto omite que en la vida real los individuos no observan a toda la población, sino a sus **vecinos en una red social**.

#### **Granovetter 1978 – Thomas Valente 1996**

Network Threshold Models:

$$a_i(t+1) = egin{cases} 1 & ext{si } r(t) \geq x_i, \ 0 & ext{si } r(t) < x_i, \end{cases}$$

$$E_i \equiv rac{\sum_{j 
eq i} \mathbf{X}_{ij} a_j}{\sum_{j 
eq i} \mathbf{X}_{ij}}, \quad a_i = egin{cases} 1 & ext{if } au_i \le E_i \ 0 & ext{otherwise} \end{cases}$$



SOCIAL NETWORKS

Social Networks 18 (1996) 69-89

#### Social network thresholds in the diffusion of innovations

#### Thomas W. Valente

Population Communication Services, Center for Communication Programs, School of Hygiene and Public Health, 111 Market Place, Suite 310, The Johns Hopkins University, Baltimore, MD 21202, USA

#### Abstract

Threshold models have been postulated as one explanation for the success or failure of collective action and the diffusion of innovations. The present paper creates a social network threshold model of the diffusion of innovations based on the Ryan and Gross (1943) adopter categories: (1) early adopters; (2) early majority; (3) late majority; (4) laggards. This new model uses social networks as a basis for adopter categorization, instead of solely relying on the system-level analysis used previously. The present paper argues that these four adopter categories can be created either with respect to the entire social system, or with respect to an individual's personal network. This dual typology is used to analyze three diffusion datasets to show how external influence and opinion leadership channel the diffusion of innovations. Network thresholds can be used (1) to vary the definition of behavioral contagion, (2) to predict the pattern of diffusion of innovations, and (3) to identify opinion leaders and followers in order to understand the two-step flow hypothesis better.

#### 1. Introduction

Contagios Complejos y la Debilidad de los Lazos largos

## Complex Contagions and the Weakness of Long Ties<sup>1</sup>

Damon Centola

Harvard University

Michael Macy Cornell University

The strength of weak ties is that they tend to be long—they connect socially distant locations, allowing information to diffuse rapidly. The authors test whether this "strength of weak ties" generalizes from simple to complex contagions. Complex contagions require social affirmation from multiple sources. Examples include the spread of high-risk social movements, avant garde fashions, and unproven technologies. Results show that as adoption thresholds increase, long ties can impede diffusion. Complex contagions depend primarily on the width of the bridges across a network, not just their length. Wide bridges are a characteristic feature of many spatial networks, which may account in part for the widely observed tendency for social movements to diffuse spatially.

Most collective behaviors spread through social contact. From the emergence of social norms (Centola, Willer, and Macy 2005), to the adoption

#### **Contagios Complejos**

- La infección se comunica no por un solo contacto, sino que se necesitan varios contactos.
- Ejemplos de contagio complejo:
  - Credibilidad de leyenda urbana,
  - Adopción de nuevas tecnologías,
  - Disposición a participar de manifestaciones riesgosas,
  - Adopción de modas vanguardistas.

#### **Contagios Complejos - Mecanismos**

Complementariedad estratégica:

Las innovaciones son costosas para los primeros adoptadores, pero menos para aquellos que adoptan después.

Los costos y beneficios depende de una masa crítica que asegure que el esfuerzo adicional vale la pena.

Credibilidad:

Las innovaciones carecen de credibilidad sino hasta que son adoptados por varios vecinos.

Por ejemplo médicos que no usan una innovación sino hasta que varios de ellos lo usan.

#### **Contagios Complejos - Mecanismos**

• Legitimidad:

Validar una norma.

Simplemente saber de un movimiento social de alto riesgo no es suficiente para unirse. Ver a varios cercanos asistir sí. Modas de ropa, estilo de peinado, zonas del cuerpo a modificar.

Contagio emocional:

La mayoría de los modelos de acción colectiva comparten la suposición de que impulsos expresivos pueden ser amplificados en reuniones concentradas espacial y socialmente.

De lazos débiles a mundos pequeños

#### Lazos débiles – Contagio simple

- Se ha mostrado que la topología de una red tiene consecuencias importantes en el comportamiento colectivo.
- Los *lazos débiles* que conectan actores distantes aceleran dramáticamente la difusión de
  - Una enfermedad
  - Información laboral (Granovetter 1973)
  - Adopción de nuevas tecnologías
  - Coordinación de acción colectiva.
- Granovetter dijo: "Lo que sea que se deba difundir puede llegar a un mayor número de personas y atravesar una mayor distancia social cuando se transmite a través de lazos débiles en lugar de fuertes."

#### Lazos débiles

- Los autores están en contra de esta frase.
- "Mediante modelos formales, demostramos diferencias fundamentales en la dinámica de difusión entre contagios simples y complejos que resaltan el peligro de generalizar la teoría de los lazos débiles a cualquier cosa que se deba difundir"

### ¿En qué sentido hay fortaleza en los vínculos?

- Fortaleza en un sentido relacional: lazos fuertes conectan amigos cercanos, con interacciones frecuentes.
- Fortaleza en un sentido estructural: facilidad de propagar información conectado a nodos distantes.
- La clave de Granovetter es que los lazos que son *débiles* en el sentido relacional, son *fuertes* en el sentido estructural:

Proveen atajos a través de la topología social.

• Por contrario, los lazos *fuertes* en sentido relacional tienen debilidad estructural: Transitividad (redundancia en triada).

#### De contagio simple a contagio complejo

- El artículo muestra que, para contagio complejo, los lazos largos *pueden* ser débiles tanto en sentido relacional como **estructural**.
- Para el contagio simple, lo importa es la longitud del puente.
- Para el contagio complejo, lo que importa es el ancho del puente.
- El re-cableo de lazos largos hace muy *poco probable* la creación de puentes *anchos*, lo que es una debilidad estructural para el contagio complejo

Efectos de lazos largos para contagio complejo

en una red anillo

#### • Definamos:

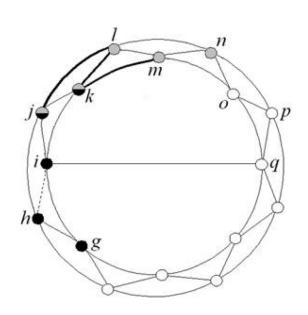
a: número de nodos mínimos para contagio.

z : número de vecinos.

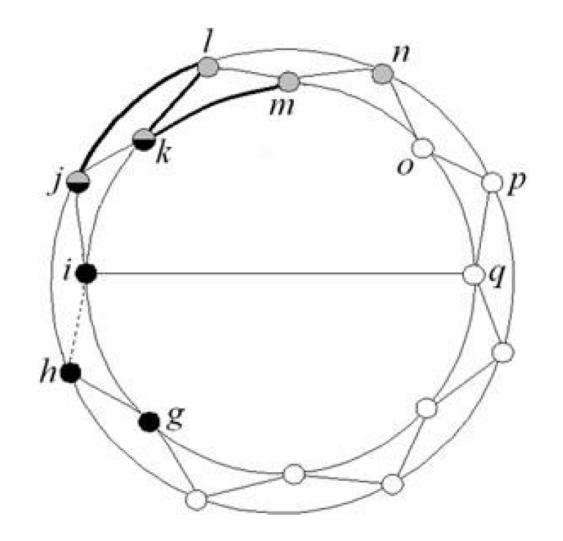
$$\tau = \frac{a}{z}$$
: threshold para el contagio.



- 1. Cada nodo tiene igual threshold  $\tau$ .
- 2. Cada nodo tiene igual influencia.
- 3. Todos los nodos tienen grados cercanos (casi iguales).

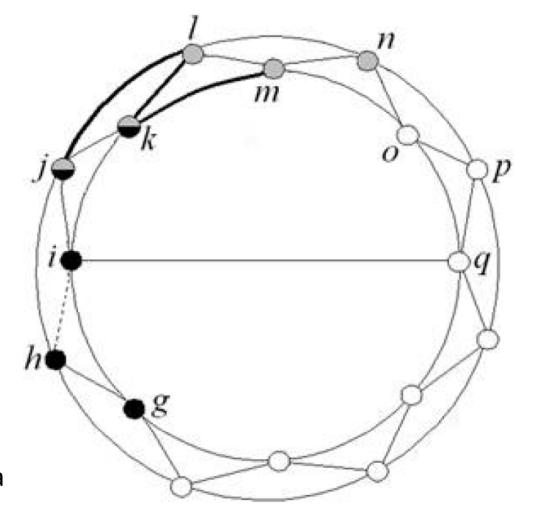


Los nodos j y k **contagian** al nodo l.



### Nodos negros = Nodos activados

En contagio **simple**, el cableo  $ih \rightarrow iq$  reduciría el tiempo de propagación.

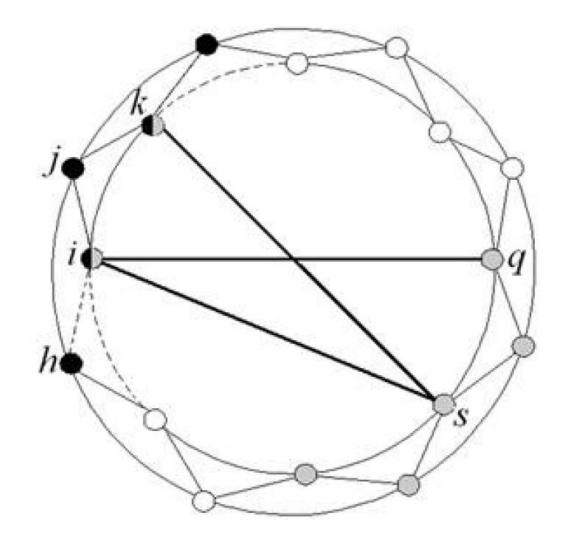


Nodos negros = Nodos activados

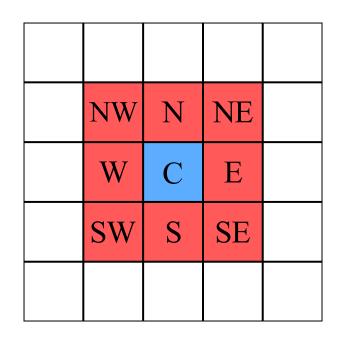
En contagio **complejo**, el cableo  $ih \rightarrow iq$  impide por completo la propagación.

Este re-cableo sí forma un puente entre los vecindarios *J* y *S*.

El efecto del re-cableo depende de si es más probable **formar** puentes a través del anillo que **romper** un puente, como en el caso anterior.



• Aumentamos la dimensionalidad del problema, por lo que usamos modelos computacionales.



Variamos threshold

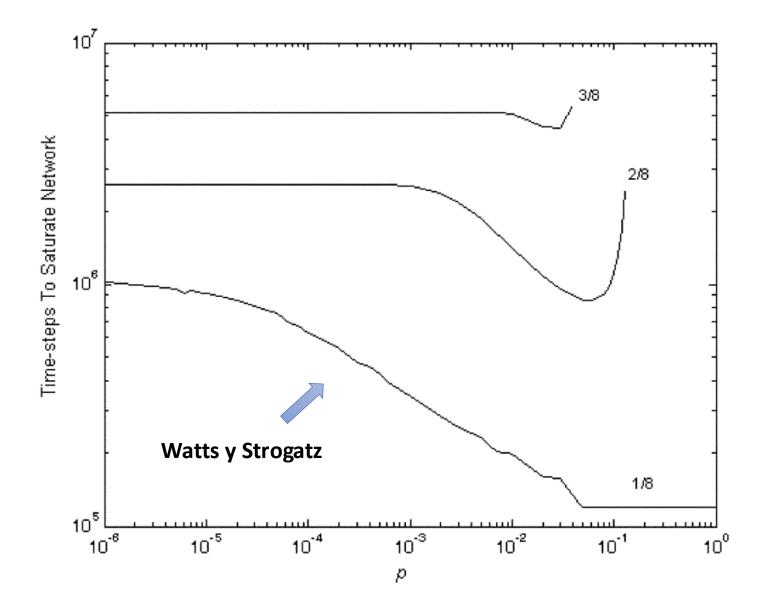
$$\tau = \frac{1}{8}, \frac{2}{8}, \frac{3}{8}.$$

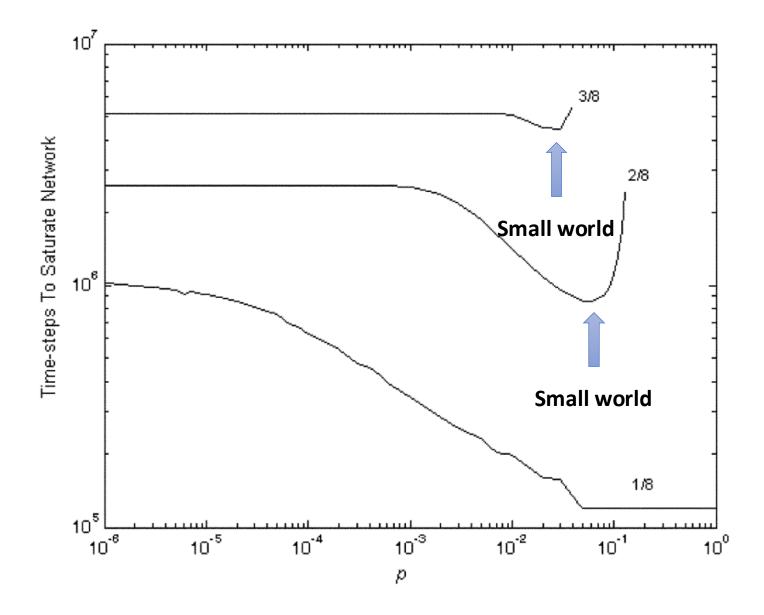
• Nos interesa ver la tasa de propagación (medido como las iteraciones requeridas para que el contagio llegue al 99%), en función del porcentaje p de lazos que son re-cableados.

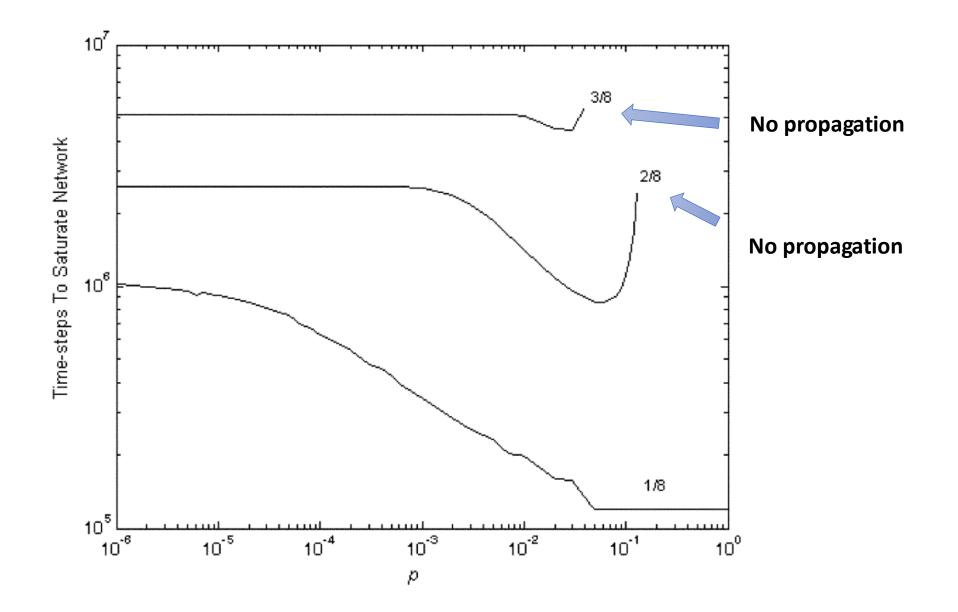
$$p=0$$
Red regular

z = 8

$$p=1$$
Red aleatoria







#### **Resultados principales:**

- 1. El contagio complejo **no** se beneficia de niveles bajos de aleatorización.
- 2. El efecto de la aleatorización p en los tiempos de saturación **no** es **monotónica** en los contagios complejos.
- Ejemplo de una transición de fase de primer orden en la propagación.
- 3. La difusión de información no es equivalente a la difusión de comportamientos.
  - Granovetter dijo: "Lo que sea que se deba difundir puede llegar a un mayor número de personas y atravesar una mayor distancia social cuando se transmite a través de lazos débiles en lugar de fuertes."

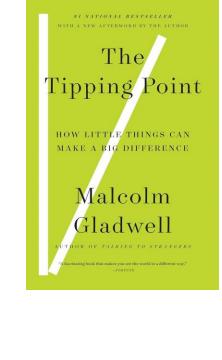
### Difusión de Innovaciones con Preferencias Individuales: Elección Racional vs Influencia Social

#### Difusión de Innovaciones

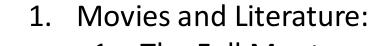
- Toma de decisiones bajo interdependencia.
- Ejemplo mínimo de **influencia diferencial** por personas cercanas. No cercano / cercano
- Preferencias individuales ≠ Comportamiento colectivo.
- Se hace cargo del **reduccionismo estructural**: setup realista.

### Motivation

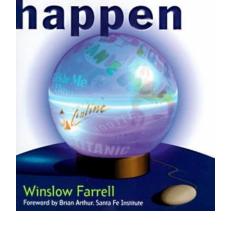
- 1. Hush Puppies rapid growth in 95'
- 2. East New York and Brownsville crime reduction
- 3. Abrupt growth of fax users
- 4. Book: Divine Secrets of the Ya-Ya Sisterhood



[1]



- 1. The Full Monty
- 2. Hootie & the Blowfish
- 2. Market Share:
  - 1. AT&T



Forecasting Predictability in a Chaotic Marketplace

these collective behaviors are studied as epidemics...

### Previous works

Previous works [4, 5] used *Percolation Theory* to construct a word-of-mouth diffusion model.

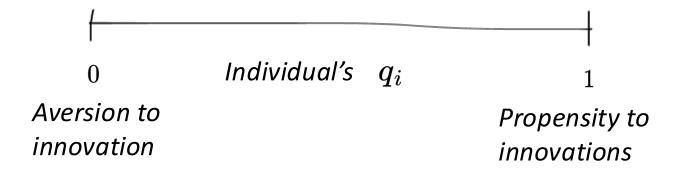
• An innovation has an *Intrinsic Utility Level* (IUL)  $\Gamma \in [0,1]$ 



### Previous works

Previous works [4, 5] used *Percolation Theory* to construct a word-of-mouth diffusion model.

- An innovation has an *Intrinsic Utility Level* (IUL)  $\Gamma \in [0,1]$
- Individuals are entirely characterized by
  - 1) their network position
  - 2) an attribute of *Minimum Utility Requirement* (MUR)  $q_t^i \in [0,1]$



Previous works [4, 5] used *Percolation Theory* to construct a word-of-mouth diffusion model.

- An innovation has an *Intrinsic Utility Level* (IUL)  $\Gamma \in [0,1]$
- Individuals are entirely characterized by
  - 1) their network position
  - 2) an attribute of *Minimum Utility Requirement* (MUR)  $q_t^i \in [0,1]$
- There is an influence driven by adopters

$$q_t^i = q_0^i \Bigg(rac{1}{\sum_{j 
eq i} \mathbf{X}_{ij} a_j}\Bigg)^lpha \;, \quad lpha > 0$$

• Finally adopt if  $q_t^i \leq \Gamma$  .

But the models use only synthetic data:

- Small World topology
- Random MUR from a distribution  $~X \sim U(0,1)$
- Global influence parameter  $\alpha$

$$q_t^i = q_0^i \Bigg(rac{1}{\sum_{j 
eq i} \mathbf{X}_{ij} a_j}\Bigg)^lpha$$

When it comes to peer influence, those dyads that are more similar to each other have greater influence [8].

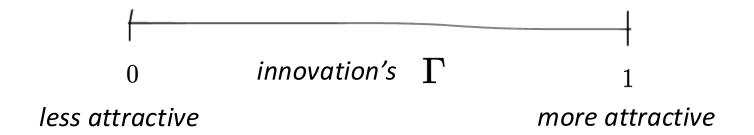
But the models use only synthetic data:

- Small World topology
- Random MUR from a distribution  $X \sim U(0,1)$
- Global influence parameter  $\alpha$

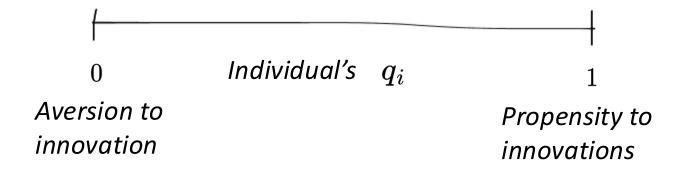
Some criticism has arisen due to the 'structural reductionism' of these works:

- "This literature often treats agents as cognition-free 'structural dopes,' operating like relay stations whose only purpose is to automatically respond to external stimuli." [6]
- Fewer attention to more realistic setups.

1. Let's assume an 'innovation' has an *Intrinsic Utility Level* (IUL)  $\Gamma \in [0,1]$  which characterize the attractiveness of that innovation.



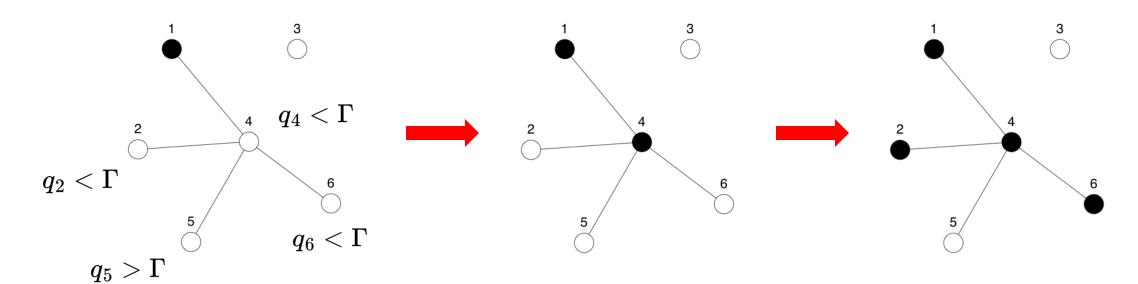
2. And individuals with *Minimum Utility Requirement* (MUR)  $q_i \in [0,1]$ .



Adoption can happen in two ways:

1) Rational Choice: *i* adopt if *Minimum Utility Requirement ≤ Intrinsic Utility Level* 

$$q_i \leq \Gamma \qquad \Rightarrow a_i = 1$$



### 2) Selective Social Influence:

$$E_i \equiv rac{\sum_{j 
eq i} \mathbf{X}_{ij} a_j}{\sum_{j 
eq i} \mathbf{X}_{ij}}, \quad a_i = egin{cases} 1 & ext{if } au_i \le E_i \ 0 & ext{otherwise} \end{cases}$$

### 2) Selective Social Influence:

$$ilde{E}_i \equiv rac{\sum_{j 
eq i} \mathbf{X}_{ij} ilde{a}_j}{\sum_{j 
eq i} \mathbf{X}_{ij}}, \quad a_i = egin{cases} 1 & ext{if } au_i \le ilde{E}_i \ 0 & ext{otherwise} \end{cases}$$

Here,  $\tilde{a}_i$  accounts for those infected individuals who are more influential:

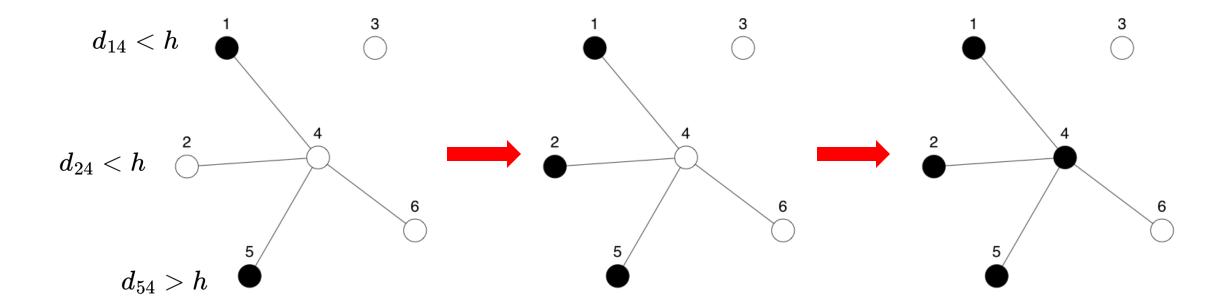
$$ilde{a}_i = 1 \quad \Leftrightarrow \quad a_i = 1 \land \overbrace{d_{ij} \leq h}$$

#### where:

- 1.  $d_{ij}$  is the social distance between individuals i and j ,
- 2. h is the Maximum Social Proximity (MSP), which measures the flexibility to be influenced by a person with different demographics.

### 2) Selective Social Influence:

Let's set  $au_4=0.5$  ; then, because the variability among the ties [6]...



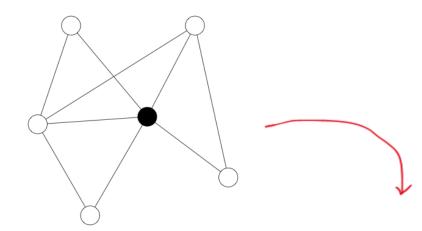
Setting up

To run the simulations in a plausible setup, we should have a network with:

- 1) a representative socio-demographic distribution, and
- 2) a plausible distribution of the population's willingness to adopt.

Following McPherson and Smith (2019) [7], you can impute a network structure on any other survey if that survey:

- 1) Is representative of the same population,
- 2) Has some basic demographic variables.



Homophilic strength in the m-th social dimension

### General Social Survey (GSS 2004):

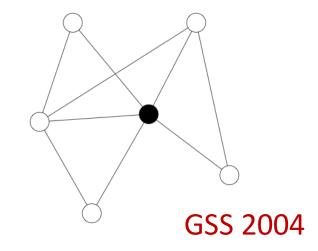
- 1. Representative of the US population
- 2. Demographics:

  Age Sex Years of Educ Race Religion
- 3. EGO data

### American Trends Panel (ATP 2014):

- 1. Representative of the US population
- 2. Demographics:

  Age Sex Years of Educ Race Religion
- 3. EGO data 'Openness to Innovations' items



	$ \operatorname{id}_1 $	$\mid \mathrm{id}_2 \mid$	$\mid \mathrm{id}_3 \mid$	•••
Age	36	27	41	• • •
Educ	12	15	16	•••
$q_i$	.54	.67	.32	
• • •				• • •

**ATP 2014** 

### General Social Survey (GSS 2004):

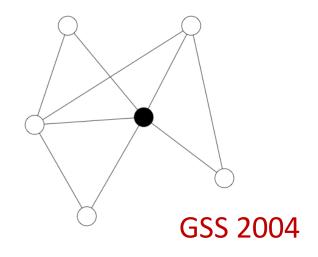
- 1. Representative of the US population
- 2. Demographics:

  Age Sex Years of Educ Race Religion
- 3. EGO data

### American Trends Panel (ATP 2014):

- 1. Representative of the US population
- 2. Demographics:

  Age Sex Years of Educ Race Religion
- 3. EGO data 'Openness to Innovations' items





Homophilic strength in the m-th social dimension

### General Social Survey (GSS 2004):

- 1. Representative of the US population
- 2. Demographics:

  Age Sex Years of Educ Race Religion

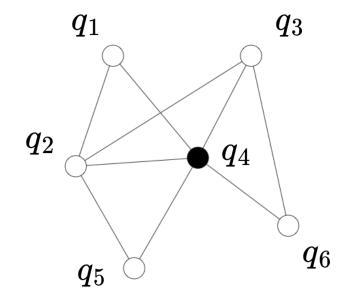


3. EGO data

### American Trends Panel (ATP 2014):

- 1. Representative of the US population
- 2. Demographics:

  Age Sex Years of Educ Race Religion
- 3. EGO data 'Openness to Innovations' items





**Table 1.** Case-Control Logistic Regression Predicting a Confiding Relationship based on General Social Survey Ego Network Data, 1985 and 2004.

Variable	Model I
Intercept	-14.456***
·	(0.048)
Different race	-1.819***
	(0.077)
Different religion	-1.362***
	(0.044)
Different sex	-0.317***
	(0.025)
Education difference	-0.049***
	(0.002)
Age difference	-0.173***
	(0.009)
Year	-0.179***
	(0.047)
N (respondents)	3,001
N (dyads)	1,139,161

We can use those values and ERGM to get networks with:

- 1. The right topology,
- 2. The right structural homophily (realistic socio-demographic attributes),
- The right distribution of your relevant variable (individual innovation propensity ).

 $q_i$ 

# Comprehensive Simulation Methodology

Using the ATP-networks (N=1000 nodes):

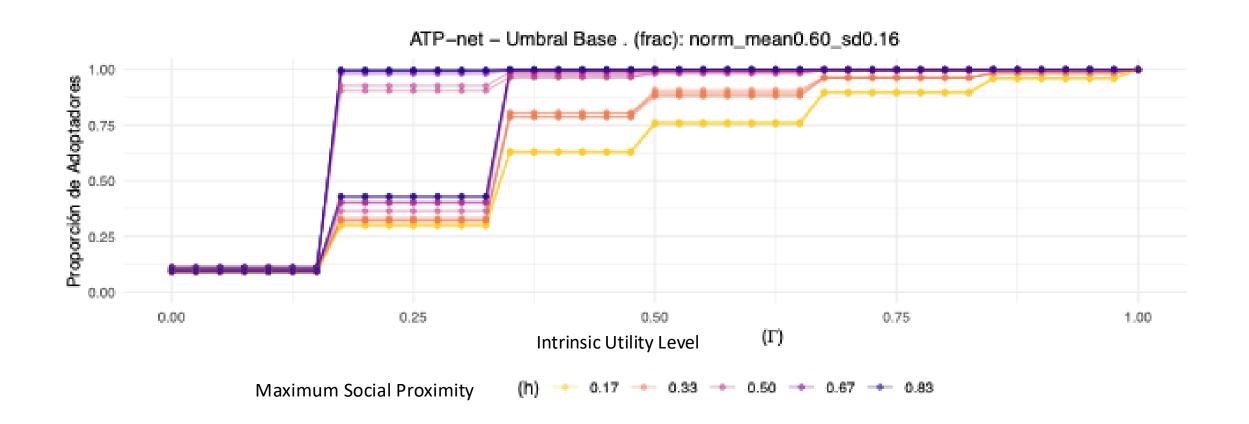
#### **Parameter Space Exploration:**

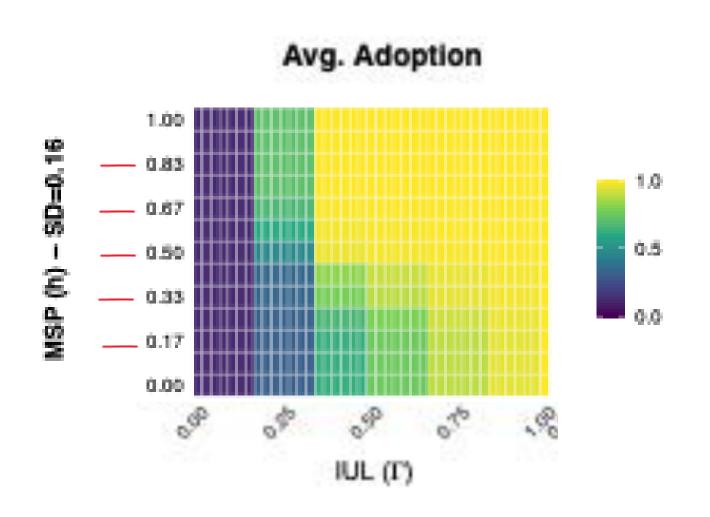
- **IUL** (Γ): 41 levels (0.0 to 1.0).
- **MSP** ( *h*): 13 levels (0.0 to 1.0).
- Thresholds (  $au_i$ ): Normally distributed  $au_i \sim \mathcal{N}(\mu_{ au}, \sigma_{ au})$ 
  - Means (  $\mu_{\tau}$ ): 4 levels (0.3, 0.4, 0.5, 0.6).
  - SD ( $\sigma_{\tau}$ ): 4 levels (0.08, 0.12, 0.16, 0.20).
- Seeding Strategies:

Random, Central, Marginal, Closeness, and Eigenvalue.

#### **Scale of Simulation:**

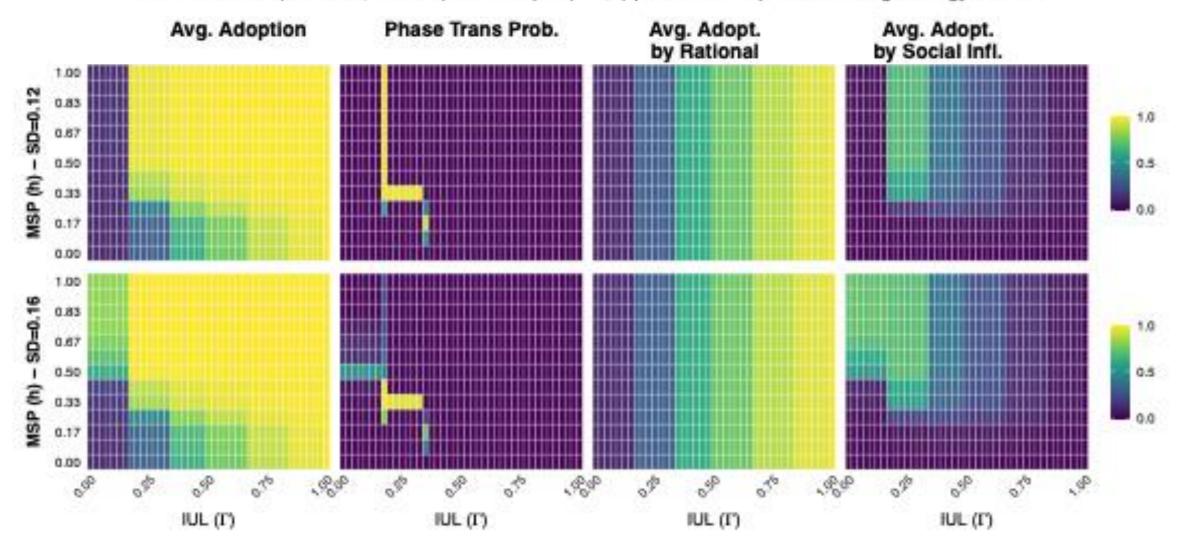
- 96 simulations for each combination of parameters
- Total simulations:  $41 imes 13 imes 4 imes 4 imes 5 imes 96 pprox 4 imes 10^6$





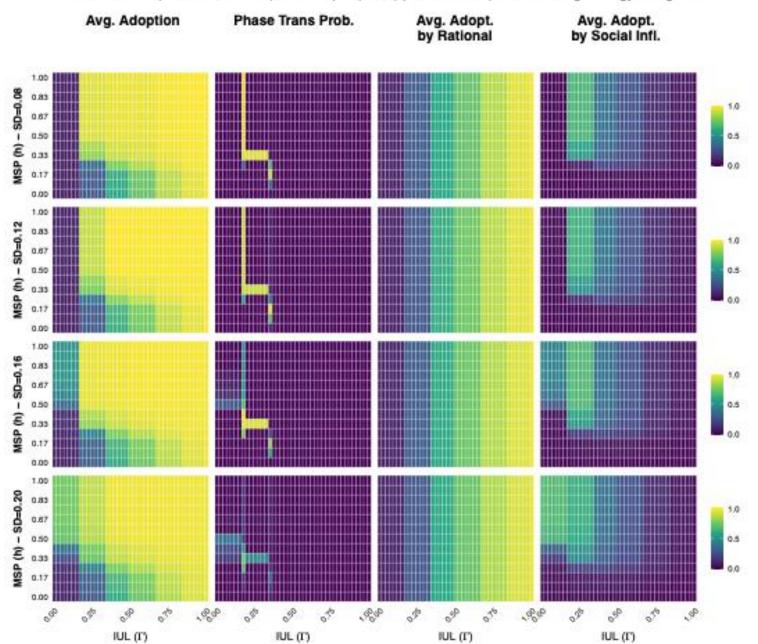
#### Consolidated Heatmaps for ATP-net - Mean threshold = 0.40

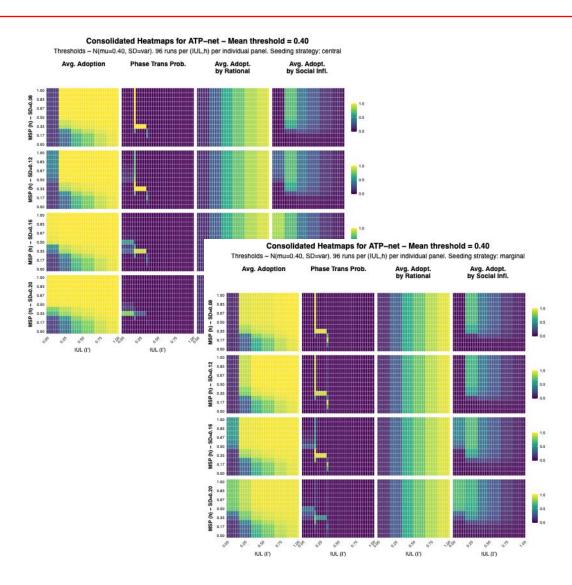
Thresholds ~ N(mu=0.40, SD=var). 96 runs per (IUL,h) per individual panel. Seeding strategy: random



#### Consolidated Heatmaps for ATP-net - Mean threshold = 0.40

Thresholds ~ N(mu=0.40, SD=var). 96 runs per (IUL,h) per individual panel. Seeding strategy: marginal





All the results are available on the GitHub repository!



# Conclusion

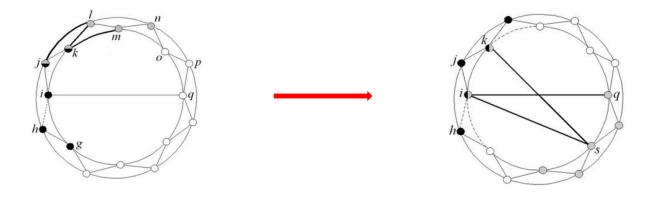
# Conclusion

#### So, what have we learned?

- Social influence plays a relevant role in adoption, even for configurations where the innovation is not particularly attractive.
- The model offers consistent results across several parameter configurations.
- The regions where social influence is on are separated abruptly by first-order phase transitions.

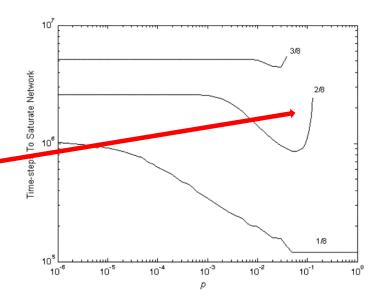
# Centola & Macy 2007 [3]

• While the rewiring is higher:



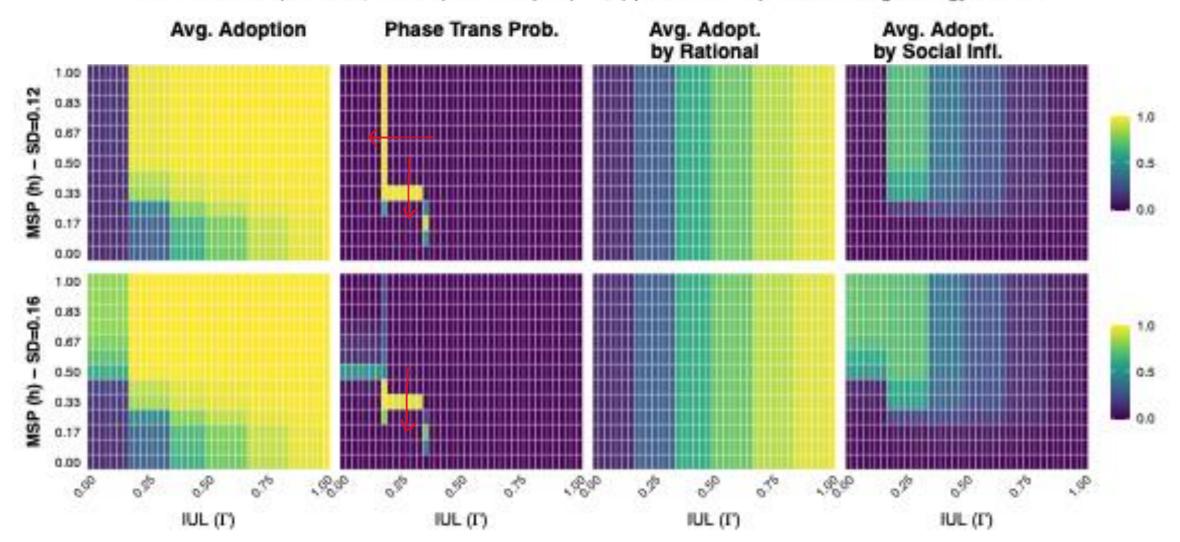
We get an abrupt barrier in diffusion:

Pure structural phase transition



#### Consolidated Heatmaps for ATP-net - Mean threshold = 0.40

Thresholds ~ N(mu=0.40, SD=var). 96 runs per (IUL,h) per individual panel. Seeding strategy: random



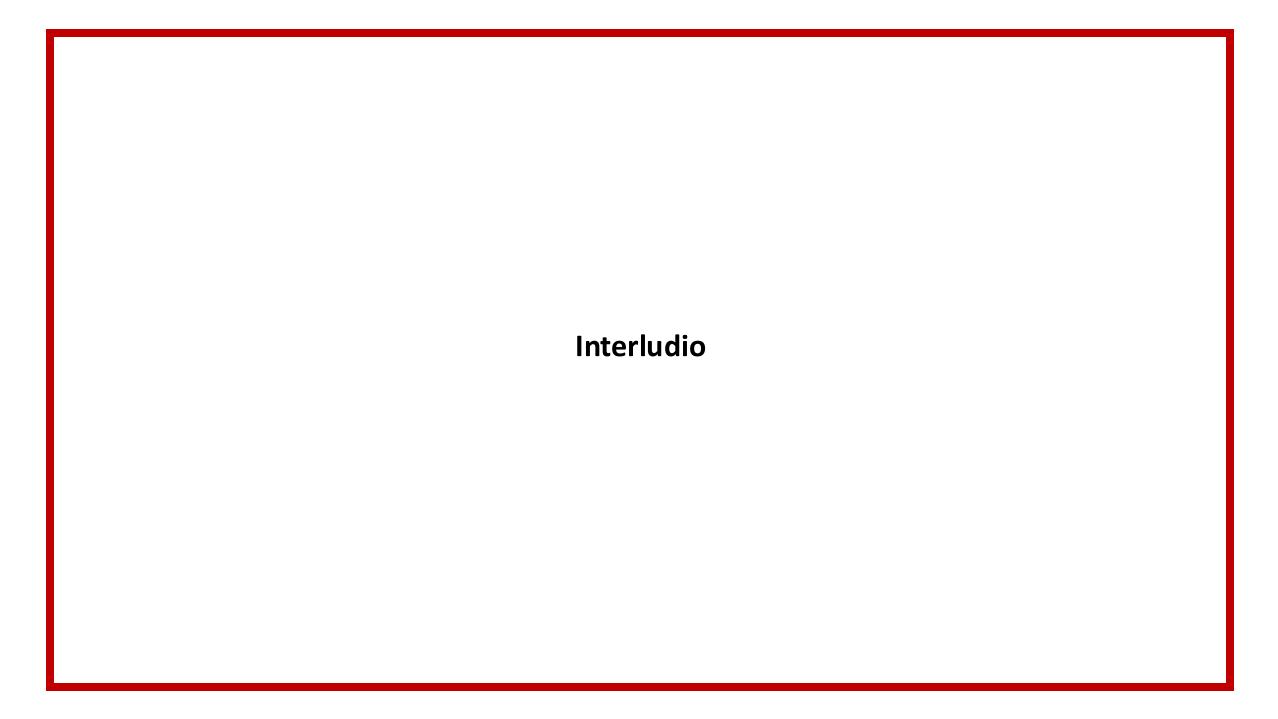
# Conclusion

#### So, what have we learned?

- Social influence plays a relevant role in adoption, even for configurations where the innovation is not particularly attractive.
- The model offers consistent results across several parameter configurations.
- The regimens where social influence is on are separated abruptly by firstorder phase transitions.
- There are non-structural barriers to word-of-mouth diffusion.

Gracias.

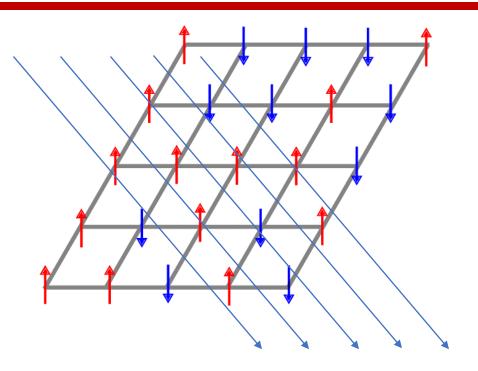
Contagio Complejo Aníbal Olivera



# **Modelo Ising**

$$\vec{B} = 0$$

Variamos T



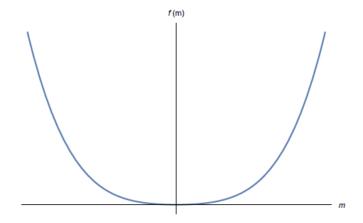


Figure 1: Free energy when  $T > T_c$ 

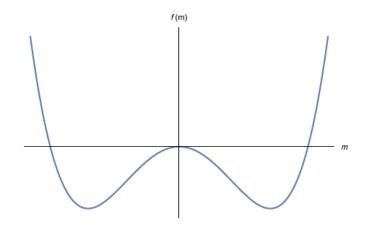
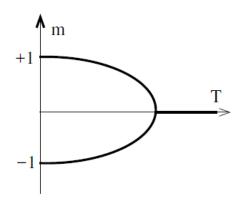


Figure 2: Free energy when  $T < T_c$ 



 $s_i = \pm 1$ 

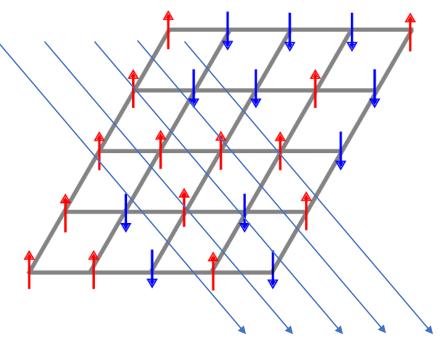
m

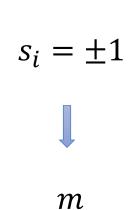
Figure 3:

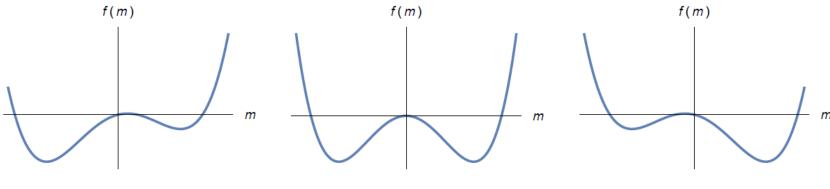
### **Modelo Ising**

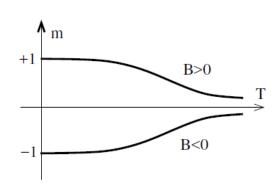
$$\vec{B} \neq 0$$

Variamos  $\vec{B}$ , manteniendo  $T < T_c$ 









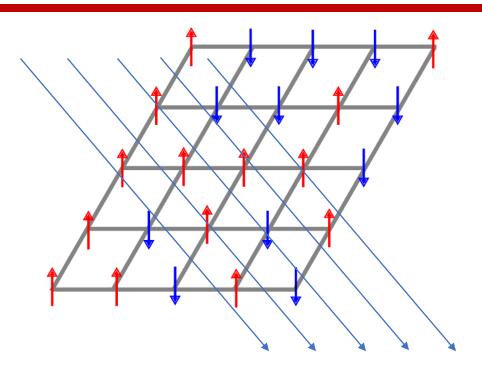
**Figure 8:**  $T < T_c$ : The free energy of the Ising model for B < 0 on the left, B = 0 in the middle, and B > 0 on the right.

Figure 7:

# **Modelo Ising**

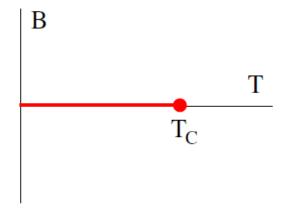
$$\vec{B} \neq 0$$

Variamos  $\overrightarrow{B}$ , manteniendo  $T < T_c$ 

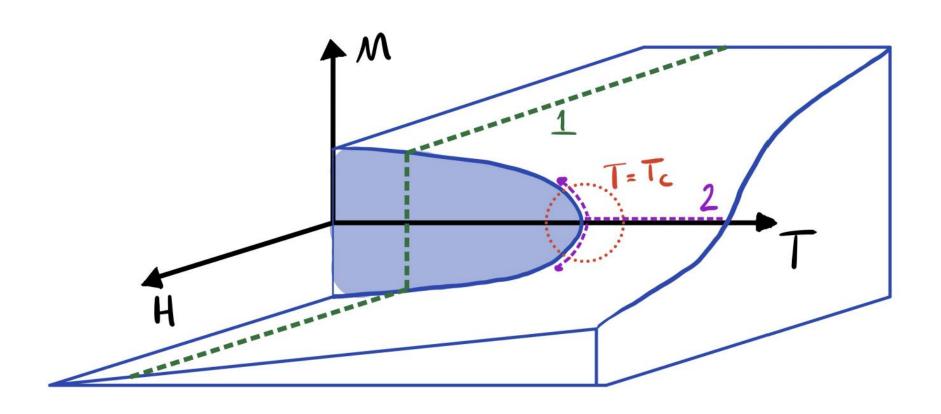


 $s_i = \pm 1$ 

m



# Ising Model

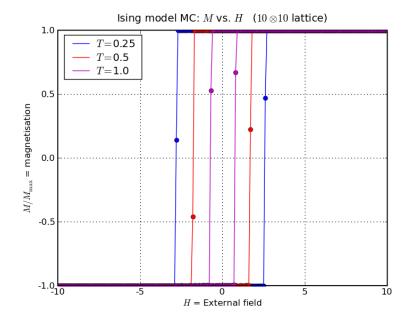


### **2nd Order Transitions**

# 

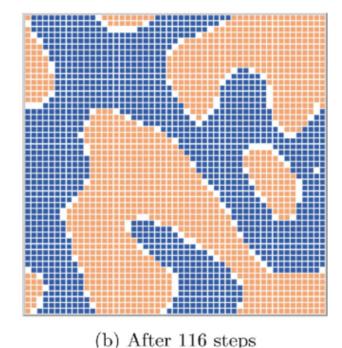
Ising Model

### **1st Order Transitions**



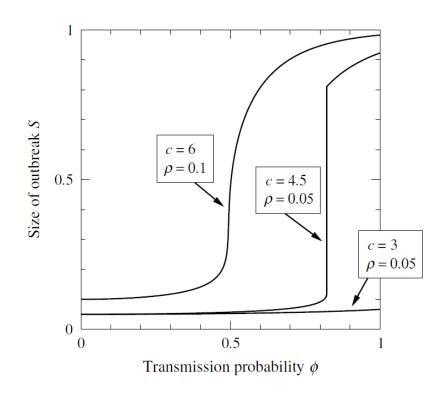
Ising Model

### **2nd Order Transitions**



Modelo de Schelling

### **1st Order Transitions**



Contagio Complejo

