Statistical tests for contagion in observational social network studies

Greg Ver Steeg and Aram Galstyan

Networks Workshop, SFI 2013

The New York Times
nytimes.com

July 25, 2007

Study Says Obesity Can Be Contagious

By GINA KOLATA

Obesity can spread from person to person, much like a virus, researchers are reporting today. When a person gains weight, close friends tend to gain weight, too.

2007

SCIENTIFIC AMERICAN™

Sign In / Register

Search ScientificAmerican.com

Subscribe | News & Features | Topics | Blogs | Videos & Podcasts | Education

Mind & Brain :: 60-Second Science :: December 5, 2008 :: 9 Comments :: Email :: Print

Happiness Is Contagious

A happy person within a social circle quickly influences those around him or her to be happy, extending to three degrees of separation. Adam Hinterthuer reports

2008

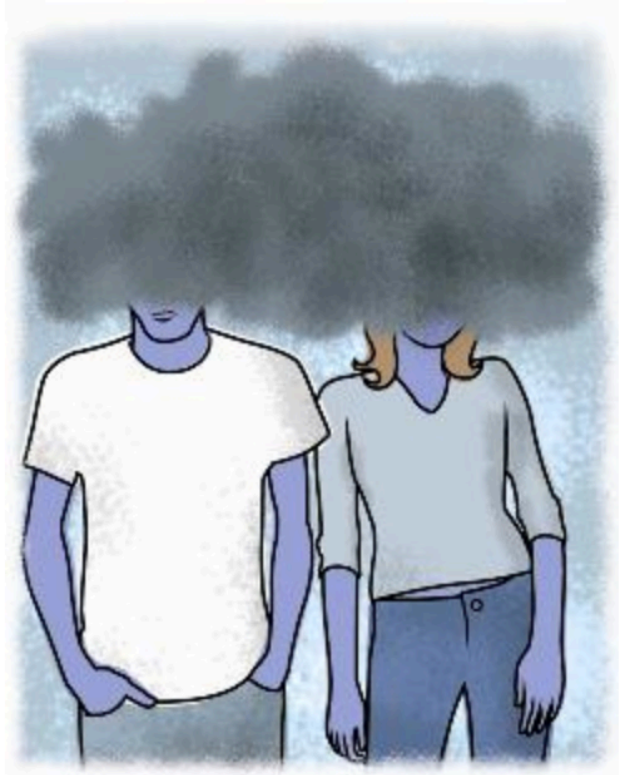
USA TODAY Home News Travel Money Sports Life Tech Wea

News » Health & Behavior ▪ Fitness & Nutrition ▪ Your Health: Kim Painter ▪ Medical Resources

Like happiness, loneliness is contagious

Updated 12/1/2009 5:14 AM | Comment | Recommend

E-mail | Print |



Enlarge

Illustration by Alejandro Gonzalez, USA TODAY

By Sharon Jayson, USA TODAY

Loneliness, like happiness, can be contagious, says research out today that shows how feeling lonely can make others lonely, too.

The study by John Cacioppo, a psychologist at the University of Chicago, builds on recent research showing that happiness is contagious and spreads through social networks.

LONELINESS: It's increasing and can harm your health
HAPPINESS: Staying positive in negative situations
CO-HOUSING: Seniors can be independent but not alone

Cacioppo worked with the two researchers who did the earlier happiness research: Nicholas Christakis, a professor of medical sociology at Harvard Medical School, and James Fowler, associate professor of political science at the University of California-San Diego.

Share

Add to Mixx
Facebook
Twitter
More

Subscribe

myYahoo
iGoogle
More

2009

The Telegraph

HOME NEWS WORLD SPORT FINANCE COMMENT BLOGS CULTURE TRAVEL LIFE FASH

Women | Motoring | Health | Property | Gardening | Food | History | Relationships | Expat

Politics | Business | Family | Sex | Life | Wonder Women Columnists

HOME » WOMEN » MOTHER TONGUE

Divorce 'is contagious'

Divorce can be contagious within groups of friends, according to a new study.

By Richard Savill
2:42PM BST 04 Jul 2010

Print this article

2010

Los Angeles Times

GPA can be contagious among high school students, study finds

February 13, 2013 | By Karen Kaplan

Researchers have some new advice for high school students who want to improve their grades: Become friends with academically oriented classmates.

It may sound obvious, but researchers went to considerable effort to prove it.

They surveyed all members of the junior class at Maine-Endwell High School in Endwell, N.Y., and asked students to rate each of their classmates as either a "best friend," a "friend," an "acquaintance" or someone they didn't know.

High school students' grades were strongly influenced by the grades... (Don Bartletti / Los Angeles...)

2013

Slate

BRIEFING NEWS & POLITICS ARTS LIFE BUSINESS & TECH SCIENCE PODCASTS & VIDEO BLOGS

HOME / SCIENCE : THE STATE OF THE UNIVERSE.

Disconnected?

We've heard that obesity and divorce can be passed from one person to another. Critics now wonder how the "social contagion" studies ever passed peer review.

By Dave Johns

Why the interest in this case?

- Marketing Facebook applications?
  A heuristic for contagion is fine.

- Making decisions about public health?
  Let's double-check that.

The New York Times

Health

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH

FITNESS & NUTRITION | HEALTH CARE

Study Says Obesity Can Be Contagious

By GINA KOLATA
Published: July 25, 2007

Obesity can spread from person to person, much like a virus, researchers are reporting today. When a person gains weight, close friends tend to gain weight, too.

Catching Obesity From Friends May Not Be So Easy

By GINA KOLATA
Published: August 8, 2011

Why are friends' behaviors correlated?

We mimic our friends:
Contagion/influence

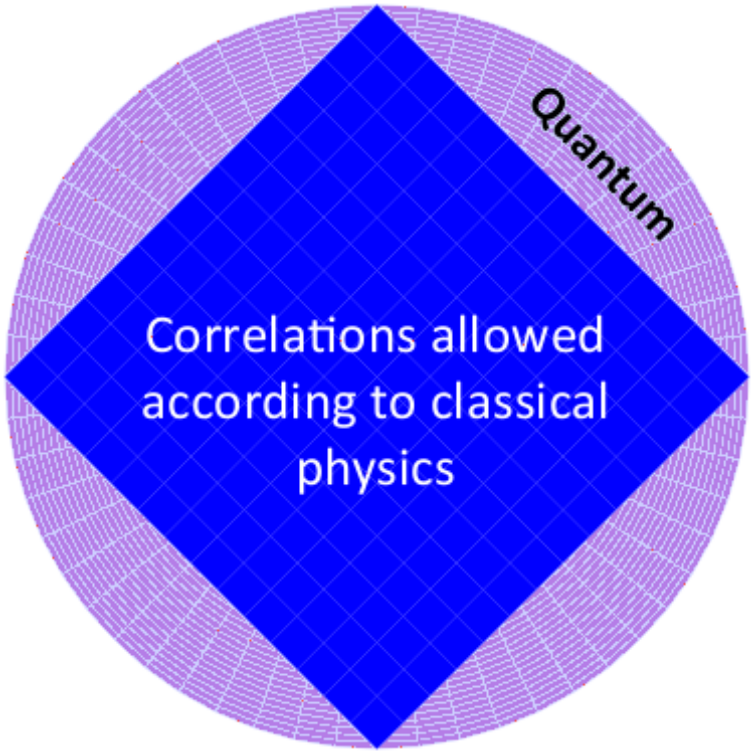
We befriend people who are already like us:
Selection/Homophily

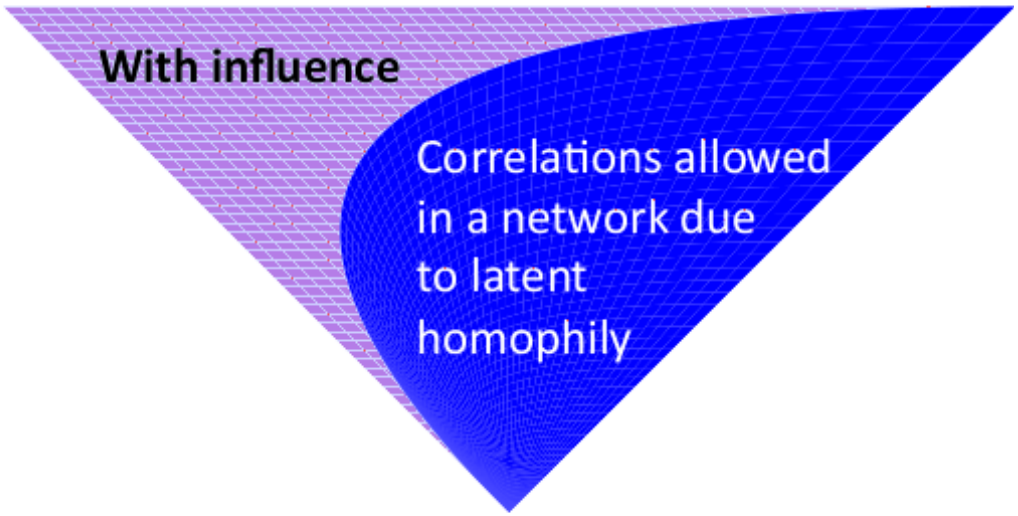
Difficulty: HIDDEN variables
(aka "unobserved covariates", "latent attributes")

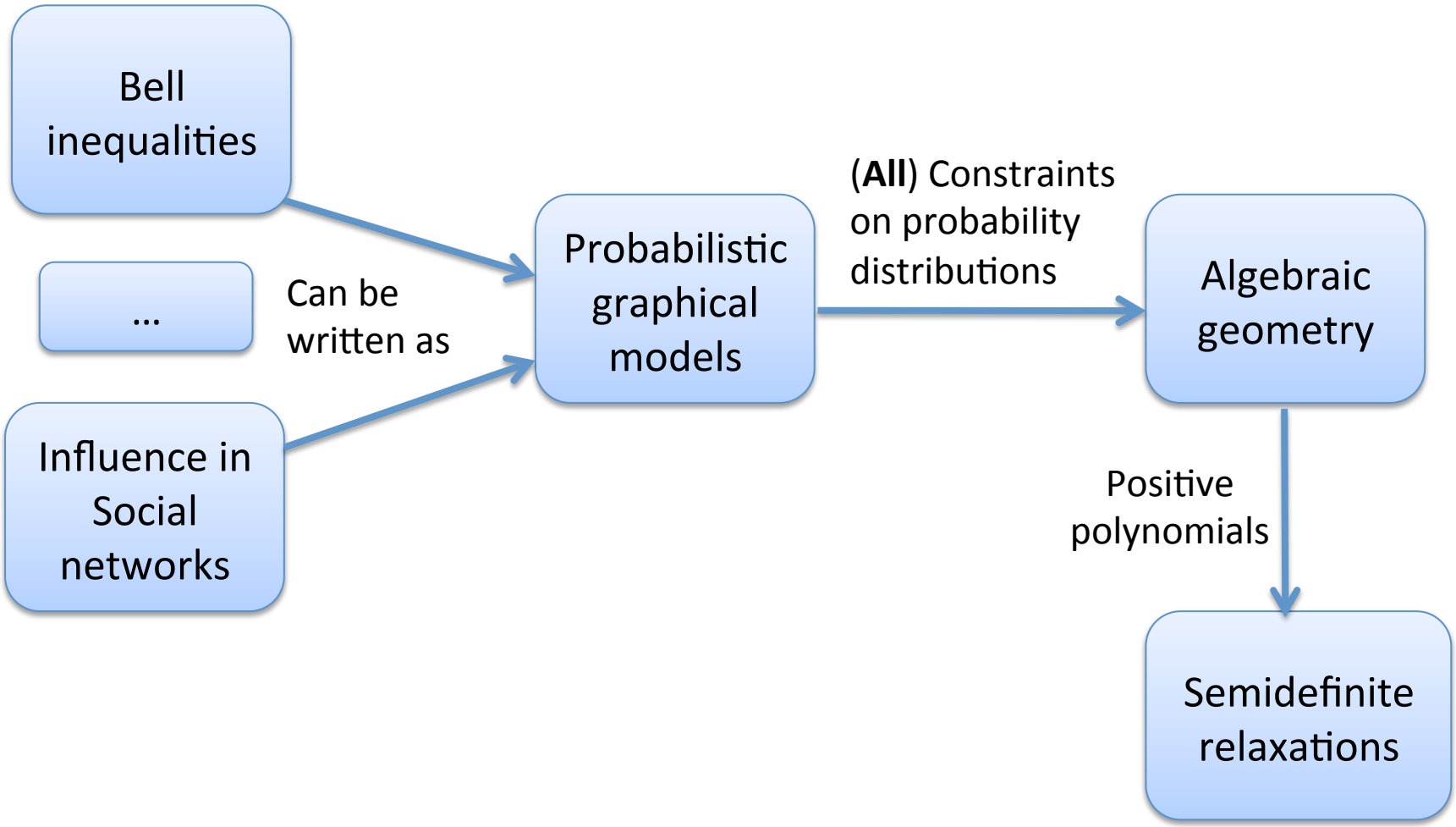
INDIVIDUALITY

ALWAYS REMEMBER THAT YOU ARE UNIQUE. JUST LIKE EVERYBODY ELSE.

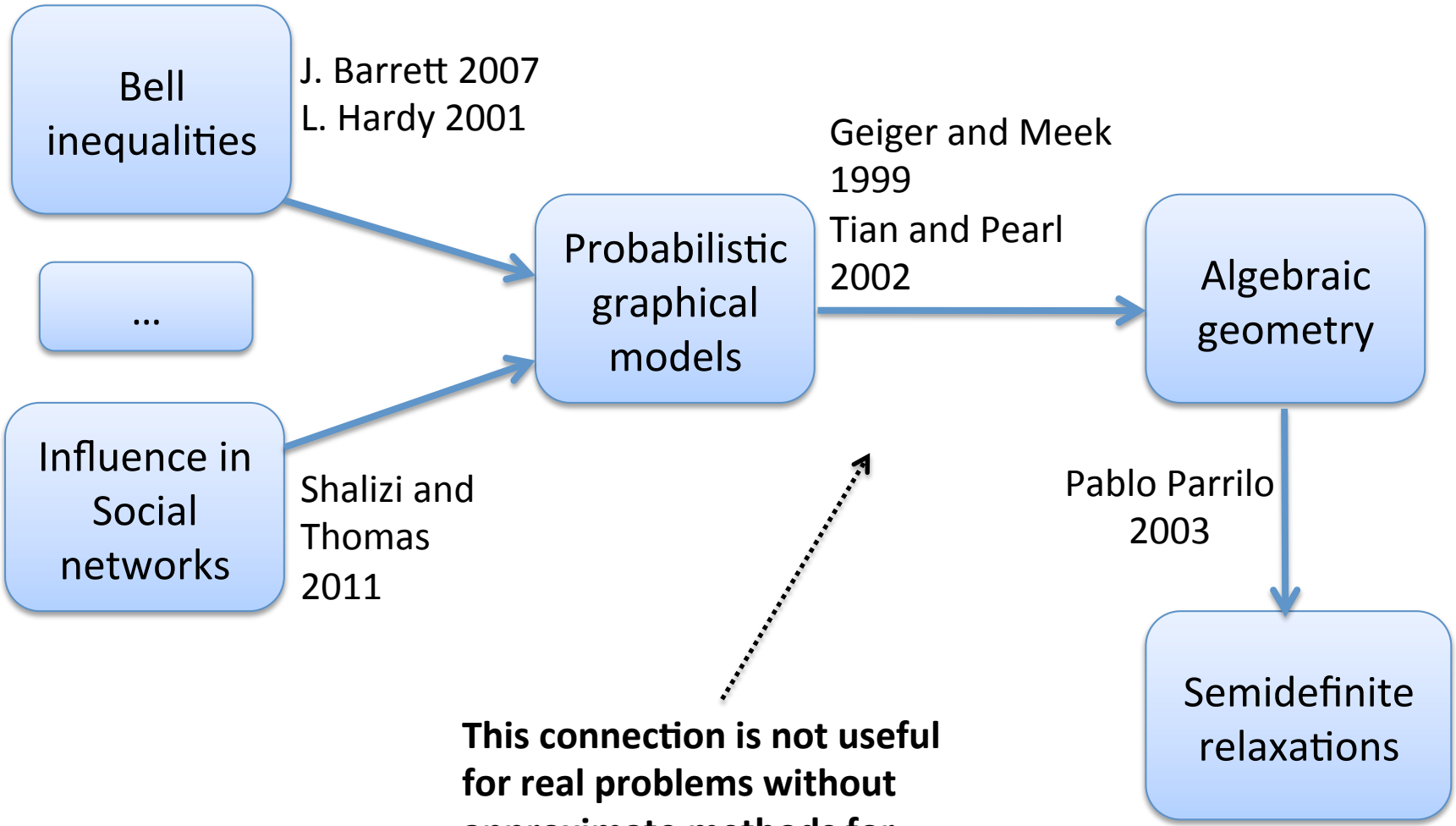
Ruling out hidden variables

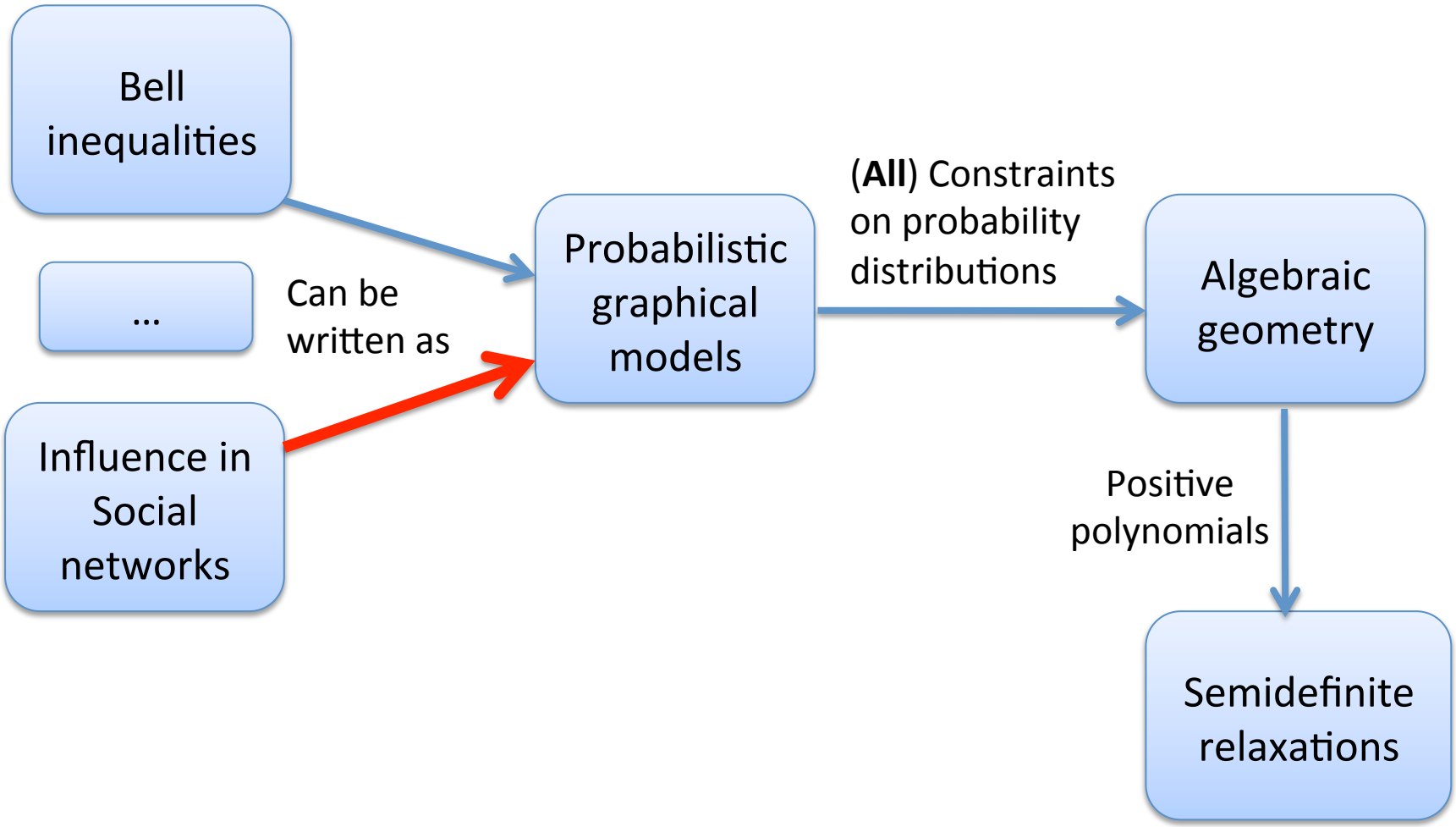




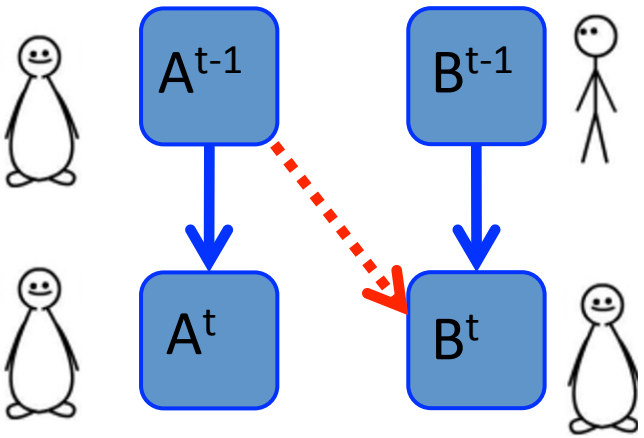

Bell inequalities
...
Influence in Social networks
Can be written as
Probabilistic graphical models
(All) Constraints on probability distributions
Algebraic geometry
Positive polynomials
Semidefinite relaxations

Context


Bell inequalities
J. Barrett 2007
L. Hardy 2001
...
Influence in Social networks
Shalizi and Thomas 2011
Probabilistic graphical models
Geiger and Meek 1999
Tian and Pearl 2002
Algebraic geometry
Pablo Parrilo 2003
Semidefinite relaxations
This connection is not useful for real problems without approximate methods for solving the AG problems


Bell inequalities
...
Influence in Social networks
Can be written as
Probabilistic graphical models
(All) Constraints on probability distributions
Algebraic geometry
Positive polynomials
Semidefinite relaxations

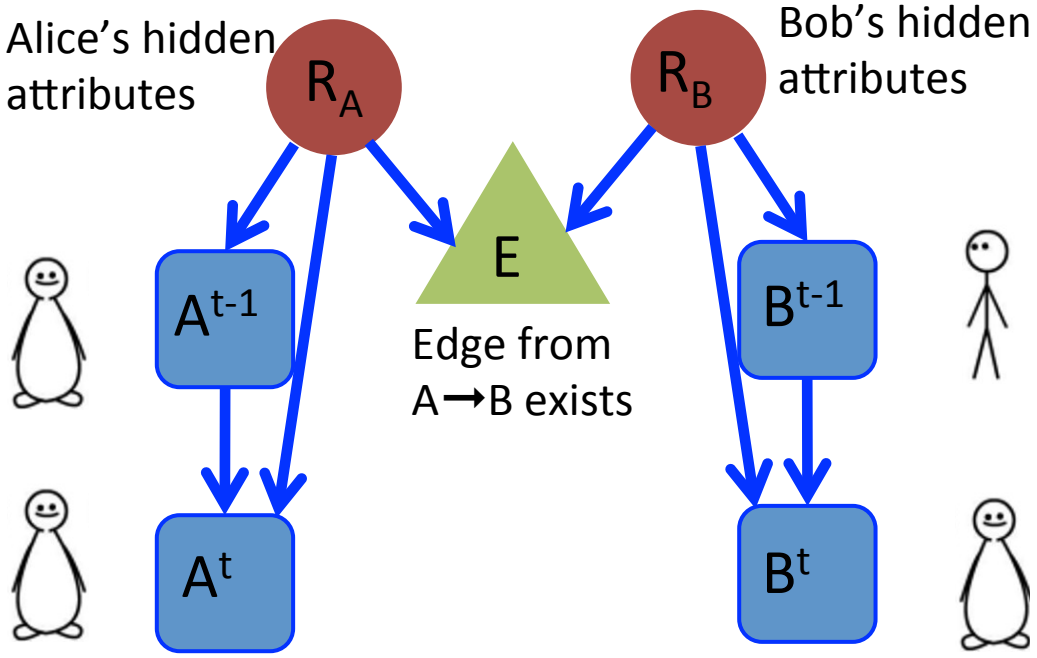
Influence or contagion explanation



Alice's state over time

Bob's state over time

Selection/Latent homophily explanation



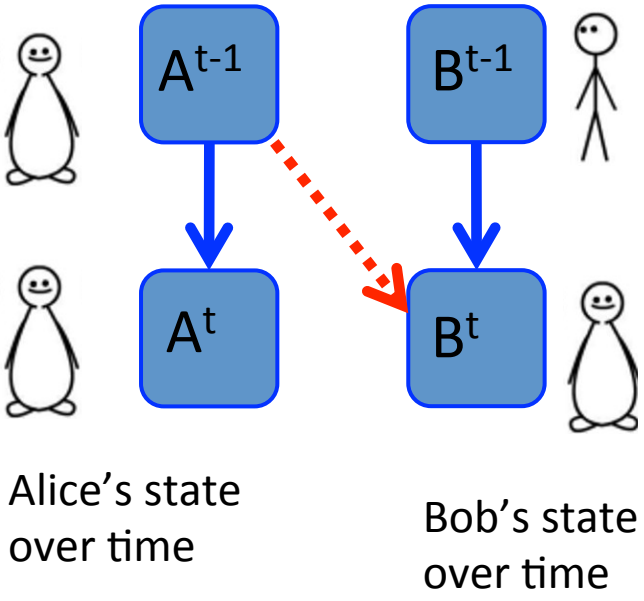
Alice's state over time

Bob's state over time

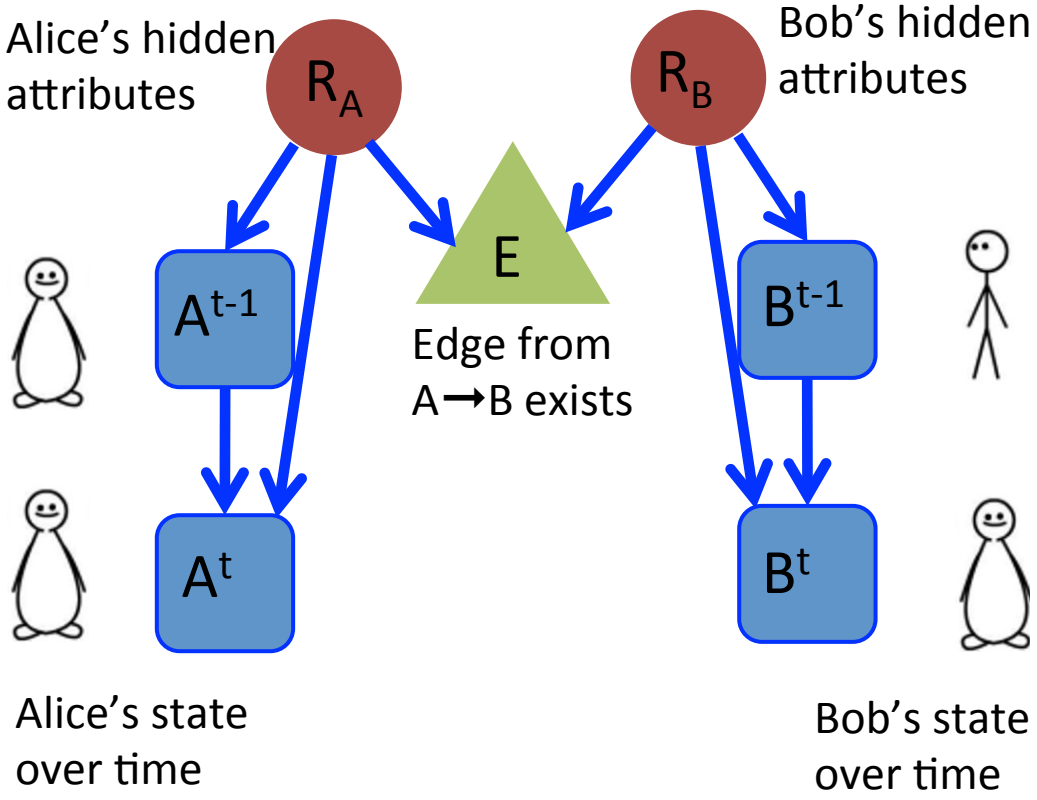
Is it easy to tell the difference?

NO: *Shalizi & Thomas, Soc. Methods & Research, 2011.

Influence or contagion explanation



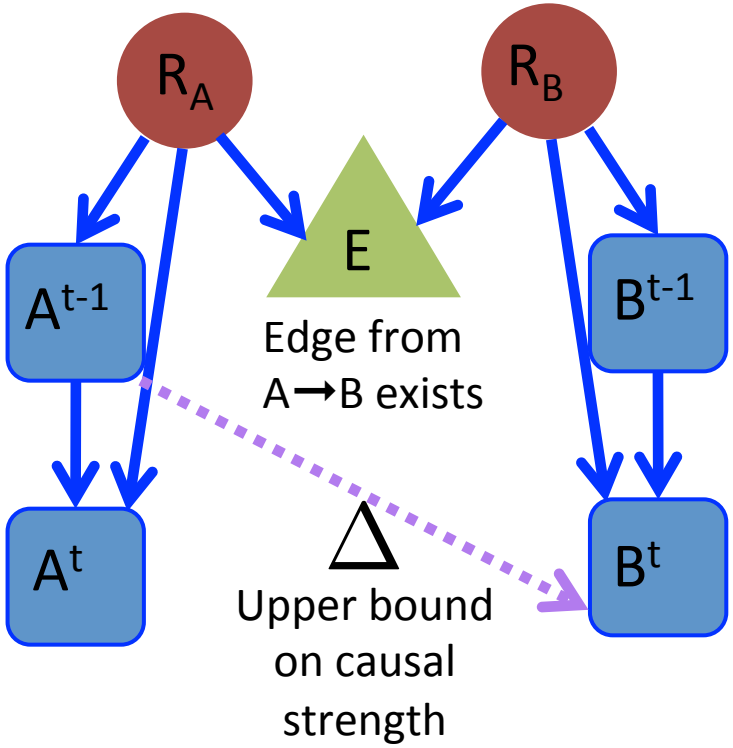
Selection/Latent homophily explanation



*Shalizi & Thomas, Soc. Methods & Research, 2011.

- Current tests for contagion (as in obesity study) can be fooled by latent homophily
- AND contagion can look like homophily

A Statistical Test



Find an observable, c(A,B), so that:

$$\forall P \in \mathcal{P}_\Delta \qquad \langle c(A,B) \rangle_P \leqq 0$$

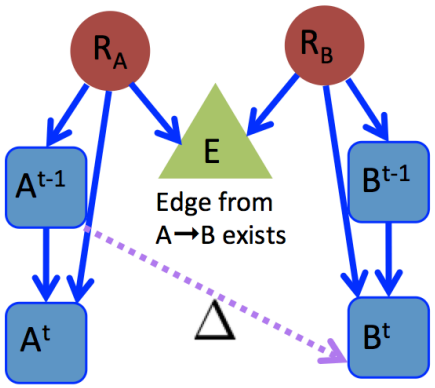
THEN, if an observed distribution $\hat{P}(A,B|E)$ violates this bound:

$$\langle c(A,B) \rangle_{\hat{P}} \neq 0$$

That implies causal strength > Δ is required to explain correlations. *

*With a caveat discussed in a few slides

How to Construct Tests



$$\forall P \in \mathcal{P}_\Delta \qquad \langle c(A,B)\rangle_P \leq 1$$

Equivalent to:

$$P \in \mathcal{P}_\Delta \iff \exists P(A_t|A_{t-1},R_A), P(R_A,R_B|E), P(B_t|B_{t-1},A_{t-1},R_B)$$

$$P(A_{1:T},B_{1:T}|E=1) = \sum_{R_A,R_B} P(R_A,R_B|E) \prod_{t=1}^{T} P(A_t|A_{t-1},R_A)P(B_t|B_{t-1},A_{t-1},R_B)$$

$$\delta_l \leq P(B_t=1|B_{t-1},A_{t-1}=1,R_B) - P(B_t=1|B_{t-1},A_{t-1}=0,R_B) \leq \delta_u \langle c(A,B)\rangle \leq 1$$

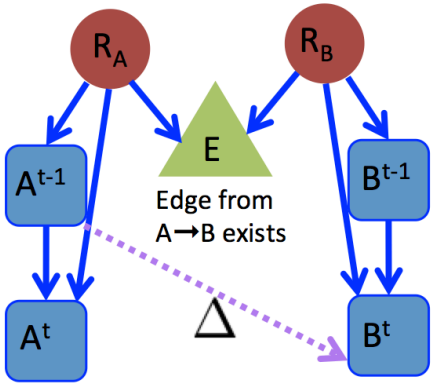
$f(\vec{x})$

$x_1, x_2, \ldots$

$g_i(\vec{x}) \geq 0$

Convex hull of an algebraic surface

How to Construct Tests



$$\forall P \in \mathcal{P}_\Delta \qquad \langle c(A,B) \rangle_P \leq 1$$

Equivalent to:

$$1 - \vec{c} \cdot \vec{f}(\vec{x}) \geq 0$$

Polynomial

$$\forall \vec{x}, g_1(\vec{x}) \geq 0, \ldots$$

(Linear) polynomial inequalities

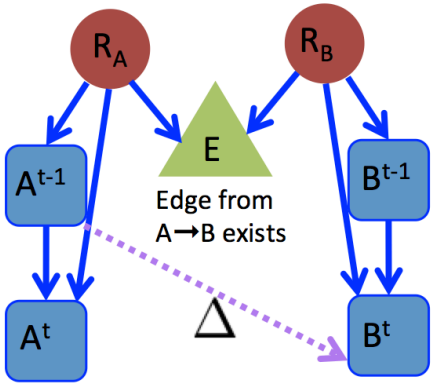
Handelman's representation

Theorem 0.1. *(Handelman's representation) Any polynomial, $\gamma - h(x)$, that is positive on a compact domain $\mathcal{K} = \{x : g_1(x) \geq 0, \ldots, g_s(x) \geq 0\}$, where the $g_i(x)$ are linear, can be written in this form,*

$$\gamma - h(x) = \sum_{k \in \mathbb{N}^s} \lambda_k \prod_{i=1}^{s} g_i(x)^{k_i},$$

using non-negative λ's, with k representing a vector of non-negative integers.

How to Construct Tests



$$\forall P \in \mathcal{P}_\Delta \qquad \langle c(A,B) \rangle_P \leq 1$$

Equivalent to:

$$1 - \vec{c} \cdot \vec{f}(\vec{x}) \geq 0$$

Polynomial

$$\forall \vec{x}, g_1(\vec{x}) \geq 0, \ldots$$

(Linear) polynomial inequalities

Search for a test, c, becomes a Linear Program

$$1 - \vec{c} \cdot \vec{f}(\vec{x}) = \sum_{k \in \mathbb{N}^s} \lambda_k \prod_{i=1}^{s} g_i(x)^{k_i}$$

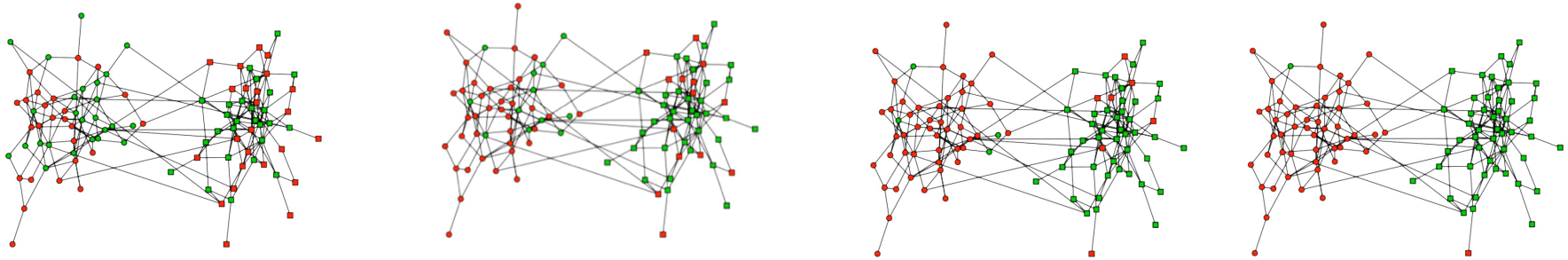
← Handelman representation

$$\vec{\lambda}_k \geq 0, \sum_i k_i \leq d_{max}$$

← Bounded degree relaxation

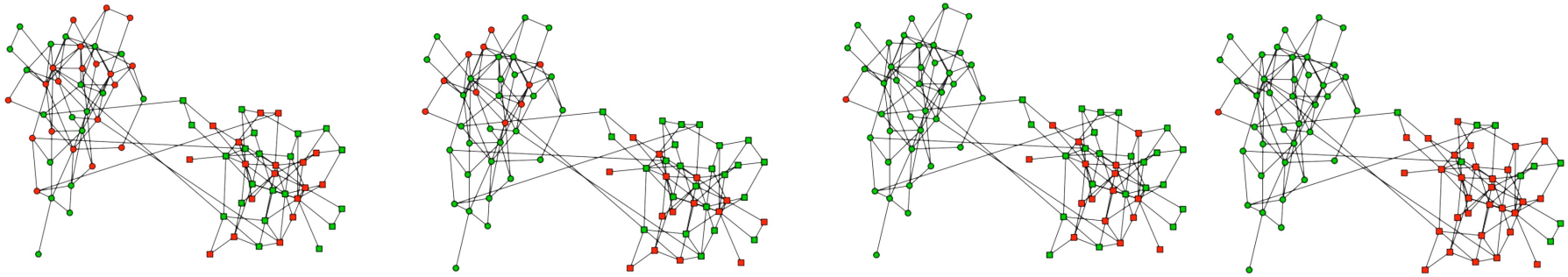
Spot the contagion!

t=1 t=2 t=3 t=4



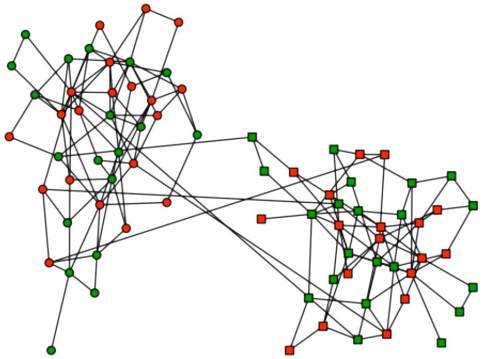
Model A

Model B

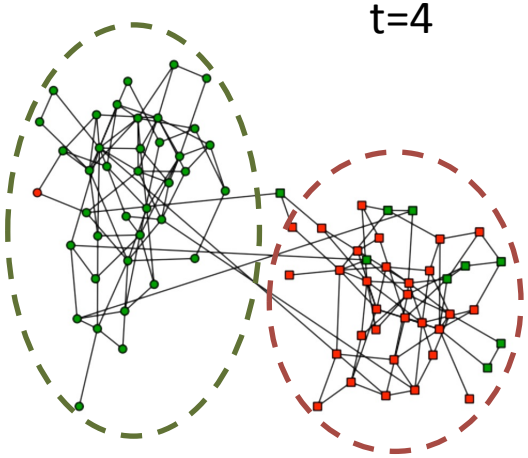


Two plausible explanations

t=1



t=4



Group 1:

- Prefer to befriend people in group 1 (homophily)
- Have a tendency to become green

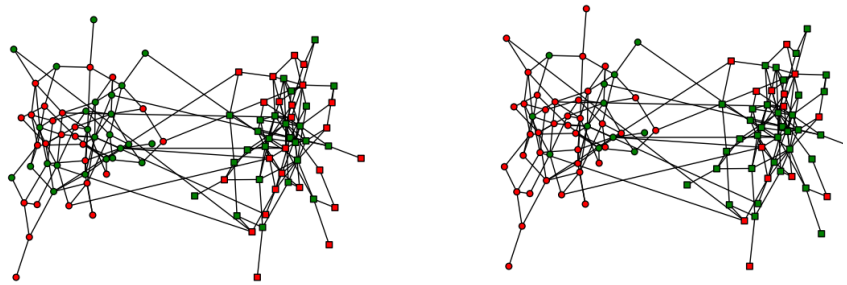
Group 2:

- Prefer to befriend people in group 2 (homophily)
- Have a tendency to become red

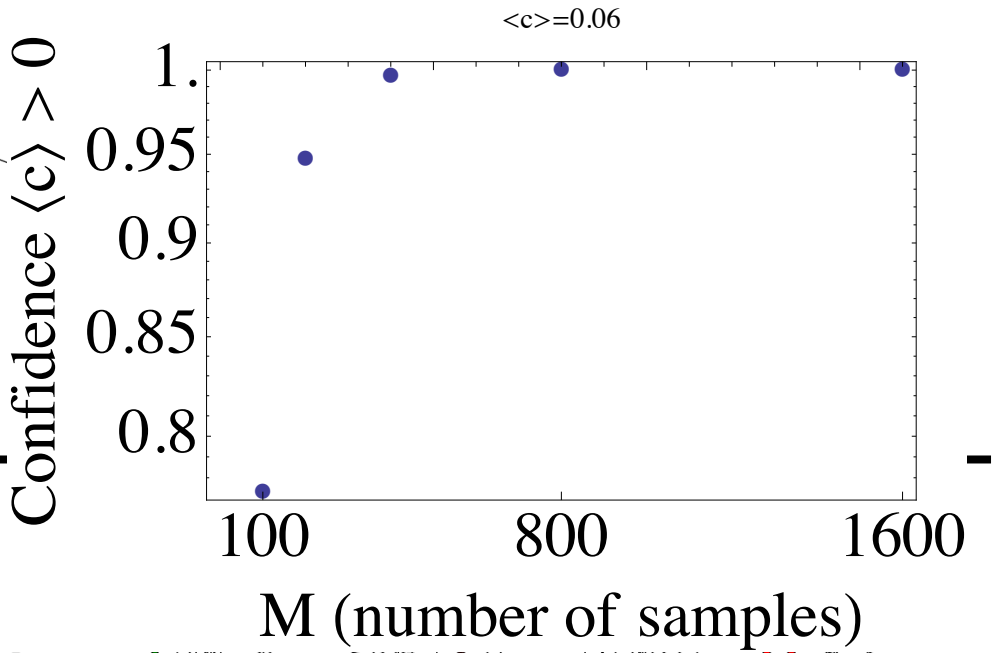
We construct a test, c, so that:

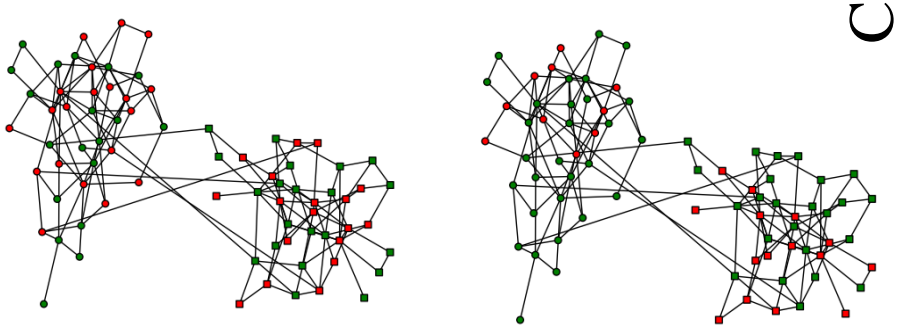
$$\text{No influence} \Rightarrow \langle c^{(1)}(A,B)\rangle_P = 0$$

$$\text{Estimate } \hat{P}(A_{1:4}, B_{1:4}|E) \Rightarrow \langle c^{(1)}(A,B)\rangle_{\hat{P}}$$

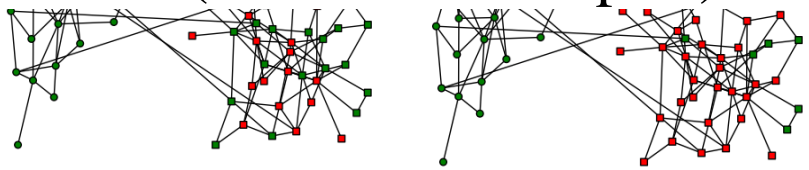


$$\langle c^{(1)}(A,B)\rangle_{\hat{P}} \approx 0.06$$





$$\langle c^{(1)}(A,B)\rangle_{\hat{P}} \approx 0$$



Framingham Heart Study Data

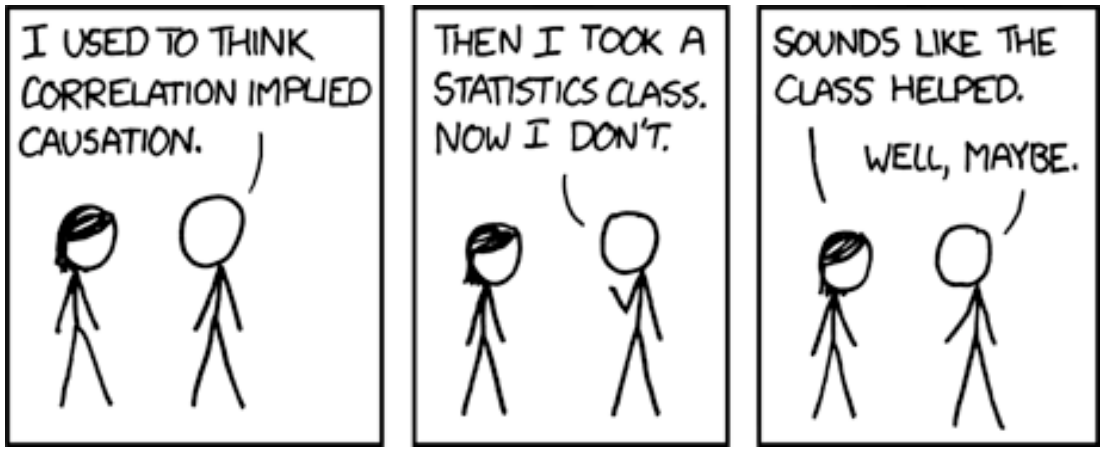
- 5124 subjects, decades long study
  - Extensive data, including: height, weight, marriage, social and familial ties, alcohol, smoking, cancer, heart disease, diabetes, psych exam…
- Edge: if B declared A as an unrelated friend
- Binary outcome: BMI increased more than median

$$\left\langle c^{(1)} \right\rangle_{\hat{P}_{obesity}} = 0.055 > 0$$

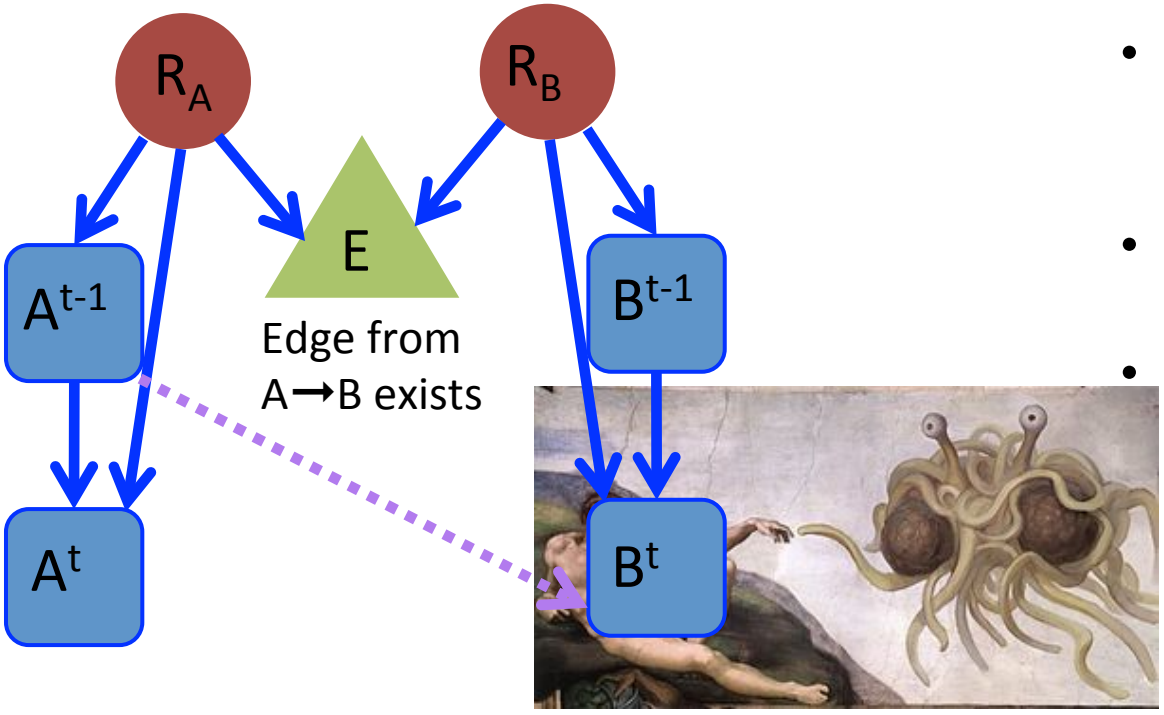
$$\text{with confidence } 99.8\ldots\% \quad \text{(Using 705 pairs)}$$

Homophily alone does not explain the spread of obesity

Causal Caveats



xkcd.com



- No model of this form can explain the spread of obesity, but…
- Adding this causal link is not the only alternative
- "external dynamics"

Conclusion

- Hidden Variables ➔ Nontrivial constraints on observations
- Allows results like: Spread of obesity NOT explained by homophily (regardless of hidden attributes)

- Future work: Other HV models, better confidence bounds, relax stationary assumption, *analyze all FHS data* (smoking, marriage, depression…)

Pre-print: bit.ly/UwUct2
Slides: bit.ly/T3LBis

Contact: gregv@isi.edu

Thank you. Questions?