Behavioral Communities and the Atomic Structure of Networks

Matthew O. Jackson and Evan C. Storms*

October 2017

Abstract

We develop a theory of 'behavioral communities' and the 'atomic structure' of networks. We define atoms to be groups of agents whose behaviors always match each other in a set of coordination games played on the network. This provides a microfoundation for a method of detecting communities in social and economic networks. We provide theoretical results characterizing such behavior-based communities and atomic structures and discussing their properties in large random networks. We also provide an algorithm for identifying behavioral communities. We discuss applications including: a method of estimating underlying preferences by observing behavioral conventions in data, and optimally seeding diffusion processes when there are peer interactions and homophily. We illustrate the techniques with applications to high school friendship networks and rural village networks.

JEL Classification Codes: D85, D13, L14, O12, Z13

Keywords: Social Networks, Networks, Cohesion, Community Detection, Communities, Games on Networks, Coordination, Complementarities, Peer Effects, Peer Influence, Diffusion, Contagion, Atoms

*Department of Economics, Stanford University, Stanford, California 94305-6072 USA, Jackson is also an external faculty member at the Santa Fe Institute, and a fellow of CIFAR. Emails: and jacksonm@stanford.edu and estorms@stanford.edu. We gratefully acknowledge support under NSF grant SES-1629446 and ARO MURI Award No. W911NF-12-1-0509. We thank Kostas Bimpikis, Ben Golub, Omer Tamuz, and participants from RINSE and the Social and Economic Networks conference at the University of Chicago for helpful comments and suggestions.

1 Introduction

Networks shape who influences whom in a society. Strong divides driven by homophily and other forces result in communities of people who are more connected with each other than with outsiders (e.g., see McPherson, Smith-Lovin, and Cook (2001)). These divides enable different communities to maintain different behaviors, norms, and cultures. Differences in behaviors range from relatively innocuous ones – such as whether you shakes hands or kiss a friend on the cheek, and even how many times and which cheek you start with, when meeting them – to more consequential ones such as whether one cheats on taxes, how much a child studies in school, which profession one pursues, and so forth (e.g., see Jackson (2018, forthcoming)). Understanding how these "behavioral communities" depend on network structure is the subject of this paper.

Our approach is to define communities according to which people influence each other's behaviors in the network. In particular, we examine settings in which there are complementarities in peoples' behaviors - for instance as in the adoption of a new technology, or the maintenance of some norm of behavior or culture.

Our approach is based on drawing a cut in the network if there is at least one situation (Nash equilibrium) in which people on one side of the cut adopt a behavior and on the other side of the network do not. The intersections of all these cuts provide the basic building blocks or "atoms" of behavior that are possible in a given network. This is what we use to define a community. This means that we consider two people in the same community if and only if they behave the same way in every one of some set of behavioral conventions. This is a minimalist notion of community, but given that these form the basic atoms of possible behaviors, our communities provide the building blocks that can aggregate into larger norms or culture in different combinations. Most obviously, our analysis shows how network structure determines which norms of behavior and cultural conventions are feasible. By understanding those basic building blocks - the 'atoms' or 'communities' in our analysis - we learn about the possible norms and behavior contagions in a society and how they depend on the behavior and network in question.

We examine canonical settings in which people choose between two behaviors: adopting a technology or not, buying a product or not, joining a movement or not, smoking or not, engaging in criminal activities or not, going to college or not, etc. We consider two different ways in which they might be influenced by their peers. One case is in which their decision to undertake a behavior depends on the *fraction* of their friends who are undertaking the behavior, and the other is in which their decision depends on the *absolute number* of friends who are undertaking a behavior. Each of these is relevant in different applications.

The atoms that emerge from these definitions are fascinatingly complex objects that uncover relationships in the network structure that are not otherwise obvious. Some atoms emerge from closely knit groups whose high internal density and low external density tie their behaviors inexorably together. However, other atoms emerge because of parallels in nodes' positions, despite the fact that those nodes are not connected to each other and do

not even have friends in common – nonetheless the nodes' positions lead them to take similar behaviors in all situations. Our main theorems then provide results on how the communities that result under these relate to block structures in large random graphs, and how they differ across the fractional versus absolute setting. Fractional-threshold-based atoms tend to subsets of blocks (for certain behaviors), while absolute-threshold atoms are supersets of the blocks. As a by-product of our theoretical results, we provide a new theorem on k-cores of large random graphs and also introduce a new technique for proving such theorems.

In addition to these theorems, we also explore several other aspects of our approach:

We explore definitions of 'robust' atomic structures. This considers conventions of behavior that work for a whole range of thresholds. This is important in many applications, since there may be heterogeneity in a population's behavioral thresholds, and there may also be measurement error in the network. These definitions are then robust to removal or addition of links or nodes, as well as variation in the thresholds that people are using. We show how 'robustness' coarsens the atomic structure, both theoretically and in application to high school friendship networks.

We also provide an algorithm for identifying the atomic structure of a network. As the number of potential atomic structures grows exponentially in the number of nodes, this is computationally demanding in some networks. Nonetheless, an algorithm that builds conventions by expanding from small seed sets works in polynomial time and finds the atoms in many cases. For instance, we show that it computes the correct atomic structure with a probability going to one in a class of stochastic block models.

Our approach also suggests a method of recovering the behavioral threshold (preferences of the agents) from the observation of a network and an associated behavioral norm. We also show how one can distinguish whether behavior is driven by a fractional threshold or an absolute threshold. We illustrate this by showing that smoking behaviors in a US high school is better explained by a fractional than an absolute threshold.

Finally, our approach can also be useful in devising a new algorithm for optimally seeding the diffusion of a new technology or behavior. By identifying the behavioral communities in a network, one knows how many nodes need to be seeded and in which communities, and can devise good approximation algorithms for optimal seeding, and ones that significantly outperform random seedings.[1]

Although there is an enormous literature that examines how to partition the nodes of a network into a collection of "communities" (see Fortunato (2010)), much of that literature is motivated entirely by the position of nodes in a graph.[2] This traces back to Lorrain

[1]Random seedings can do well in simple contagion processes in which there is no threshold for behavioral choices (Akbarpour, Malladi, and Saberi, 2017). However, if behavior involves a threshold, then our approach significantly outperforms random seedings.

[2]There is a literature that builds models of homophily which result in communities, such as Currarini, Jackson, and Pin (2009, 2010); Chen, Liu, Sun, and Wang (2010); Currarini and Mengel (2012); Kets and Sandroni (2016). That is a very different perspective, since it is preferences over partners that determines the network, while here is it networks that are determining behavior.

and White's (1971) notion of structural equivalence among nodes in a network, which identified nodes that were interchangeable in terms of their positions. While their notion of equivalence was an important first step in understanding the communities in a network, it was so restrictive (nodes need to have identical sets of friends in a network) that only very particular networks produce interesting equivalence relations. Their definition was loosened and extended in various directions (e.g., see White and Reitz (1983); Borgatti and Everett (1992)) including block models (Holland, Laskey, and Leinhardt (1983)), and ultimately to the extensive community detection literature.

Our approach is not only different in terms of what we identify, but also in terms of the basic perspective that drives our definition. We derive community structure from the patterns of behavior that a network can support - so our approach is driven by what a network does rather than, and only indirectly, by the structural patterns that it exhibits. Of course which patterns of behaviors a network supports depends on its structure, but this change in focus is fundamental and does a couple of things. First, it provides a reasoning behind what a community is – a microfoundation. Second, it often leads us to find more basic "atoms" – or different communities than standard community detection algorithms. In particular, our community partitions are often refinements of those produced by other algorithms because larger blocks identified by others tend to have subtle substructures along which behaviors can diverge. Also, for some behaviors the atoms go across blocks as the thresholds for contagion are low enough that blocks are not the right objects for identifying behavior. Which community partition is correct depends, of course, on the application, but the atoms that our approach identifies are important in attempting to influence or predict behavior.

Obviously, our results relates to a games-on-networks analysis (e.g., see Jackson and Zenou (2014)). A first contribution in that direction was Morris (2000),[3] whose results addressed a question of when it is that a network allows for two different actions to be played in equilibrium. Our analysis provides new definitions, perspectives, and results concerning the behavioral conventions that a network can sustain.

Our results also provides new insights into the estimation of preferences in settings with complementarities (Bramoullé, Djebbari, and Fortin (2009)), as well as how to best seed a network when there are complementarities in diffusion (e.g., see Granovetter (1978); Centola, Eguíluz, and Macy (2007); Aral, Muchnik, and Sundararajan (2013)).

[3]There is an earlier literature on the majority game or voting game, e.g., Clifford and Sudbury (1973); Holley and Liggett (1975) that is a precursor to Morris's analysis for a specific threshold.

2 A Model

2.1 networks

A finite set $N = \{1, \ldots, n\}$ of people or nodes, with generic indices i, j, are connected in a network g.

A *network* is a simple graph, (N, g), consisting of a finite set of nodes or vertices N together with a list of the undirected edges that are present g. We let $ij \in g$ indicate that the undirected link (or edge or tie) between nodes i and j is present in the graph.

We consider undirected networks (mutual friendships) and so g is taken to be symmetric. We adopt the convention that $ii \notin g$ so that agents are not friends with themselves.

Agent i's neighbors in g is the set $N_i(g) \subset N \setminus \{i\}$ $N_i(g) \equiv \{j | ij \in g\}$.

Isolated nodes are not of much interest in our setting, so we ignore them. They are easily incorporated by allowing them to each be their own convention. In what follows we presume that a network is such that each node has at least one neighbor.

2.2 Conventions

A behavior is characterized a threshold $q \in (0, 1)$, and agents adopt that behavior if at least q of neighbors do.

One interpretation is that a person plays a coordination game with each one of his or her neighbors:

		Neighbor's Choice :	
		Adopt Behavior	*Not Adopt*
Own Choice :	*Adopt Behavior*	x	y
	Not Adopt	w	z

where $x > w$ and $z > y$. The agent makes one choice but cares about how that behavior matches in each of his or her interactions. This coordination game has a corresponding threshold that if at least q neighbors adopt the behavior, then the agent's best response is to adopt too:

$$q = \frac{z - y}{x - w + z - y}.$$

Exactly at this threshold an agent is indifferent, but otherwise has a unique best response.

Generically, the threshold would never be hit exactly. However, some rational thresholds, such as $1/2$, are prominent in the literature,[4] and so we allow for rational thresholds.

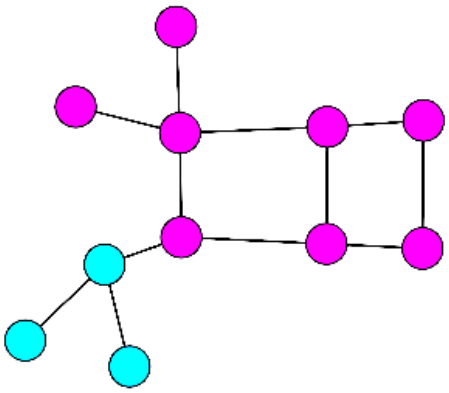
Unless otherwise noted, we break ties so that an agent adopts the behavior if exactly q neighbors do. Similar results are induced if one presumes that agents do not adopt the behavior exactly at the threshold by then looking at a threshold of $q - \varepsilon$ for some small

[4]The 'voter game' or 'majority game' is well-studied (e.g., see the discussion in Jackson and Zenou (2014)).

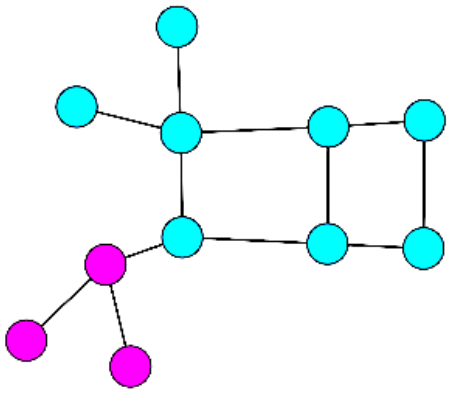
$\varepsilon < \frac{1}{n-1}$. If one reverses the tie-breaking then our use of open and closed intervals below reverses.

A *convention* associated with some threshold q on a network g is a group $S \subset N$, for which all members of S have a fraction of at least q of their neighbors in S, and all agents not in S have strictly less than q of their neighbors in S.

A couple of conventions are pictured in Figure 1 for a q of .4.



(a) The pink nodes form a convention



(b) In this example, the complement is also a convention for the same q

Figure 1: **Conventions:** A couple of conventions when $q = .4$.

Although the two conventions in Figure 1 are complementary sets, it is not always true that a convention's complement is also a convention. An example of a convention whose complement is not a convention is given in Figure 2.

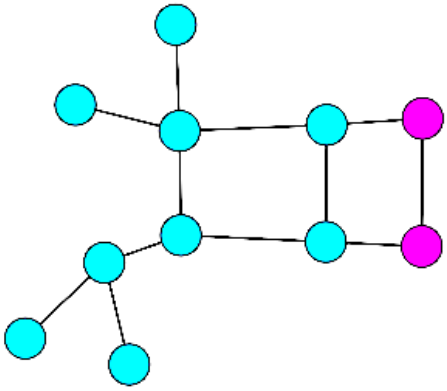


Figure 2: **Conventions:** A convention for $q = .4$ for which its complement is not a convention.

Although the complement of a convention is not always a convention for the same q, it is necessarily a convention for the complementary $1 - q$ (up to tie-breaking), as we note in the following observation.

OBSERVATION **1** *Consider q such that q cannot be expressed as a fraction with d_i in the denominator for any $i \in N$ (e.g., if q irrational). Then S is a convention for q if and only if $N \setminus S$ is a convention for $1 - q$.*

2.3 Cohesive Sets and Conventions

Following Morris (2000), we define a group $S \subset N$ to be q-cohesive if each of its members have a fraction at least q of their neighbors in the group.

A group $S \subset N$ is q-closed if every individual outside of S (in $N \setminus S$) has a fraction of his or her friends in the group that is less than q.

OBSERVATION **2** *A group S forms a convention with a threshold q if and only if it is q-cohesive and q-closed.*

2.4 Absolute Thresholds

The above definitions are relative to some fraction of neighbors taking an action. This applies naturally to coordination problems.

For some other games of strategic complements, it can be natural to adopt a behavior if at least t neighbors do, for some $t \in \{0, 1, 2, \ldots, n-1\}$.

For instance, one might want to learn to play bridge if at least three other friends do.

Clearly, if the network is regular of degree d, then S is a convention for q in a relative (fraction setting) if and only if S is a convention for $t = qd$ in the absolute setting. However, when one examines networks that are not regular, then there are q's that generate conventions that are not conventions for any t, (and vice versa).

Thus, the absolute thresholds are also of interest.

In what follows, there are equivalent definitions just switching q and t everywhere.

In the threshold setting, $t = 1$ generates the component structure as its atoms, while $t \geq 2$ subdivides the component structure.

2.5 Community Structures as Partitions Generated by Conventions

We now define the central concept of the paper: how community structures are defined from conventions.

Given a network g, let $C(q, g)$ denote the σ-algebra generated by all conventions that are robust relative to q.

The atoms of $C(q, g)$ (minimal nonempty sets in $C(q, g)$) exist by finiteness. Note that they form a partition that generates $C(q, g)$.

Thus, the atoms of $C(q, g)$ form a community structure.

We use the terms 'communities' and 'atoms' interchangeably in what follows to describe the atoms of $C(q, g)$.

To see how conventions define communities, let us consider all of the other conventions associated with the network from Figure 1:

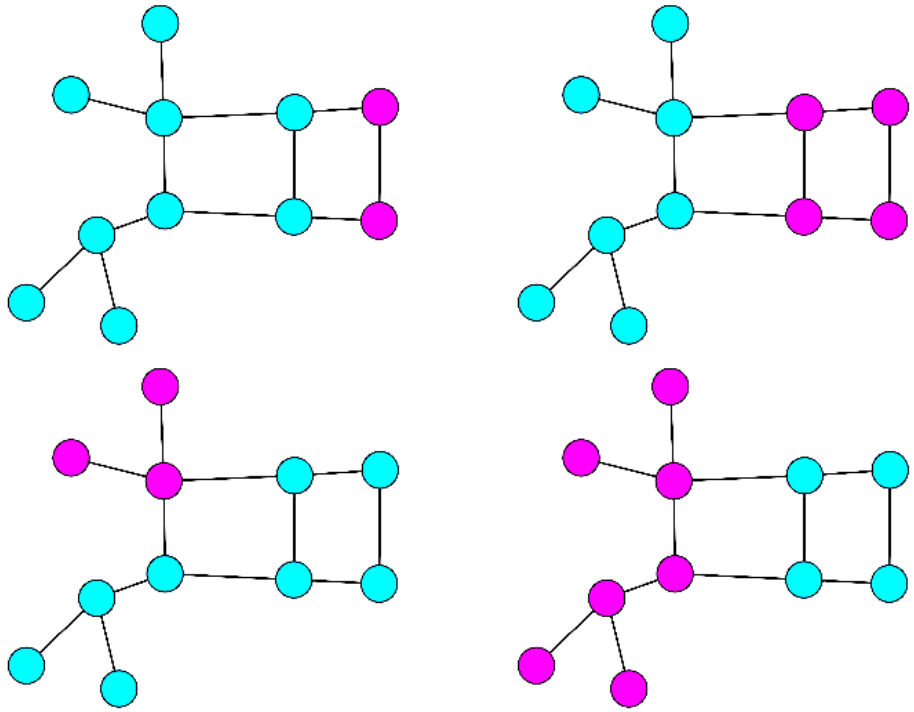


Figure 3: **More Conventions:** The other (non-degenerate) conventions when $q = .4$ (which are also conventions for $q \in (1/3, 2/3]$) .

The partition induced by all of these conventions is pictured in Figure 4.

Nodes inside an atom always behave the same as all other nodes in the same atom in all conventions for q, and if they are in different atoms then there is some convention under which behave differently. Thus, there is a behavioral sense in which these nodes are tied together and define a community. These atoms are also subsets of any convention, and conventions are necessarily unions of atoms, and thus the atoms are the basic building blocks of coordinated behaviors in a society.

2.6 Dependence of Community Structure on the Threshold q

The set of conventions can vary with the threshold, q, as different q's impose different restrictions in terms of cohesiveness and closure.

In Figure 5, we display the atoms for the social network in an American high school from The National Longitudinal Study of Adolescent to Adult Health (Add Health) for a few different choices of q.

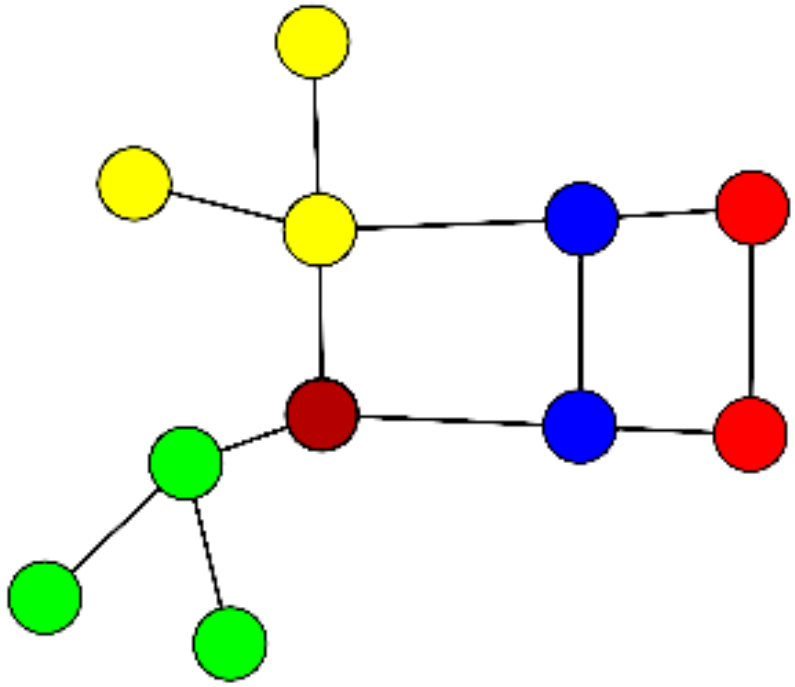
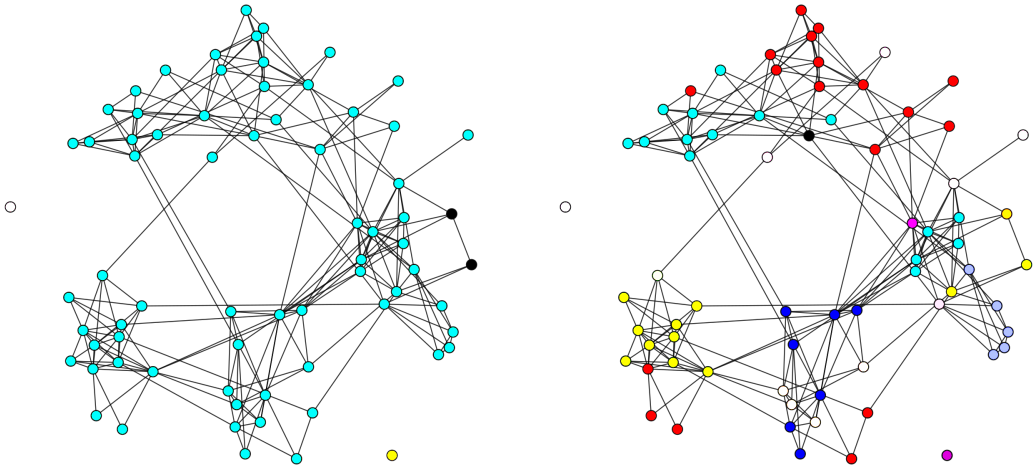
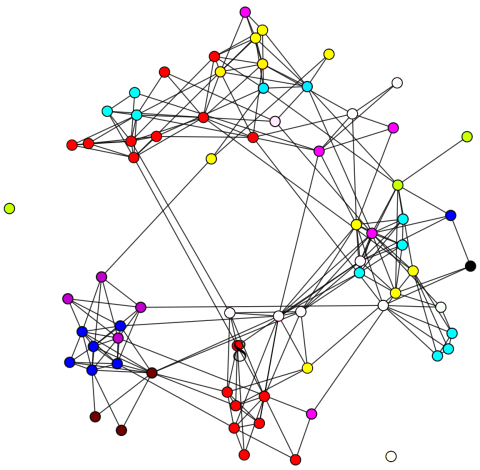


Figure 4: **The Behavioral Communities:** The behavioral communities for $q = .4$.



(a) q=1/4

(b) q=1/3



It is important to note that community structures are not necessarily nested as q varies. As q is increased the cohesiveness requirement for a convention gets *harder* to satisfy, while the closure requirement gets *easier* to satisfy. Thus, there is no monotonicity of conventions as q changes. This means community structures may change non-monotonically in q.

2.7 Community Structures from Robust Conventions

Given the dependence of the community structure on the specific level of q, it is useful to define conventions that are 'robust' in the sense that they remain conventions for some set of qs.

There are at least four reasons for considering communities that are generated by conventions that are robust to some range of q's:

- One may wish to identify robust communities that remain intact for a variety of behaviors.
- Individuals may be heterogeneous in their preferences and so behave according to a range of qs at the same time.
- The network that is observed may have measurement error in it, so that there may be missing links and/or nodes (or contain extras), and so one would like to have a convention that is robust to changes in the fractions of neighbors that are undertaking a given action.
- A network may evolve over time, and so the current network might only be an approximation of what might be in place at some other time.

Considering conventions that work for a range of q's addresses all of these issues.

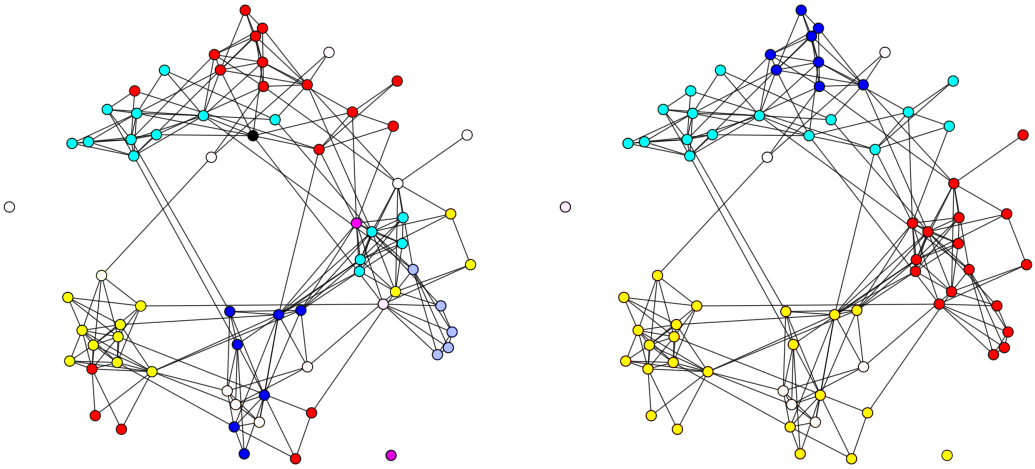
We say that a convention is a *robust convention* relative to some set $Q \subset [0,1]$ if it is a convention for all $q \in Q$.

As an illustration, the conventions in Figure 1 are both robust conventions for $(1/3, 2/3]$, but not for any additional q's.

Given a network g, let $C(Q, g)$ denote the σ-algebra generated by all conventions that are robust relative to Q.

Robustness matters, as it is more stringent to require that a convention hold for a range of q's rather than just a single q. That leads to fewer conventions and a coarser convention structure.

One reason that robustness can be very useful is easily seen in the following analysis of the behavioral communities in our high school social network. Even a small amount of robustness can have a large impact when dealing with rational q's, as illustrated in Figure 6.



(a) $q = 1/3$ (b) $q = 1/3 \pm \varepsilon$

Figure 6: **:** Even a small amount of 'robustness' ($\varepsilon < \frac{1}{n}$) coarsens the atomic partition of the high school social network at a threshold of 1/3.

3 Some Facts about Behavioral Communities

To begin to illustrate how behavioral conventions work, we first note some simple facts about how community structures change as we change q and Q.

3.1 Varying q and Q

The following lemma collects some obvious but important observations that help characterize robust community structures. A key observation is that only the extreme points of Q are needed to check whether some convention is robust with respect to Q. This follows since cohesivness is most demanding at the $\sup(Q)$, while closure is most demanding at the $\inf(Q)$, and so requiring something be a convention for all of Q then has to work at the extremeties.

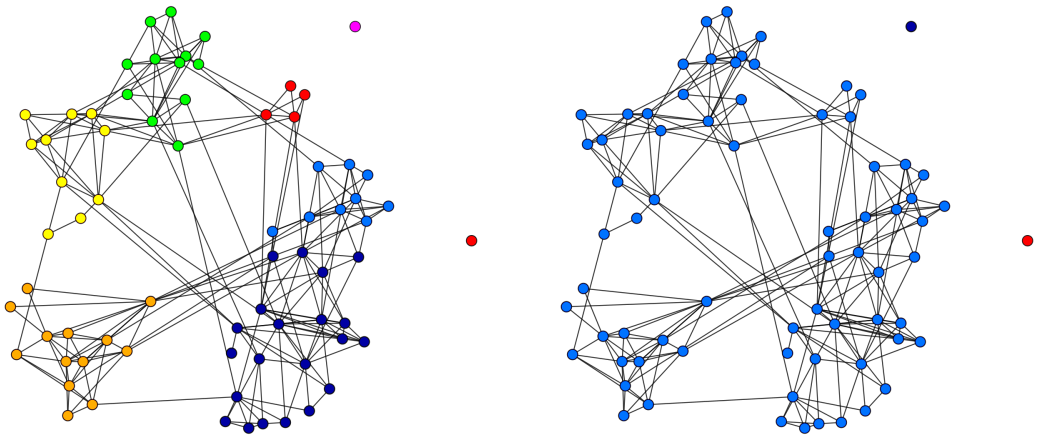
Lemma **1**

- *If Q contains $\inf(Q)$, then S is a robust convention with respect to Q if and only if it is $\inf(Q)$-closed and $\sup(Q)$-cohesive.*
- *If $\inf(Q) \notin Q$ then S is a robust convention with respect to Q if and only if it is $\inf(Q) + \varepsilon$-closed for all $\varepsilon > 0$ and $\sup(Q)$-cohesive.*
- *If S is a robust convention relative to Q, then it is a convention for all $q \in (\inf(Q), \sup(Q)]$.*[5]

[5] If $\inf(Q) \notin Q$ and $\inf(Q)$ is rational, then S may fail to be a convention at exactly $\inf(Q)$.

- *If S is a robust convention relative to Q and $Q' \subset Q$, then S is a robust convention for Q'.*
- *If $Q' \subset Q \subset [0,1]$, then $C(Q,g)$ is coarser than $C(Q',g)$ for any g.*

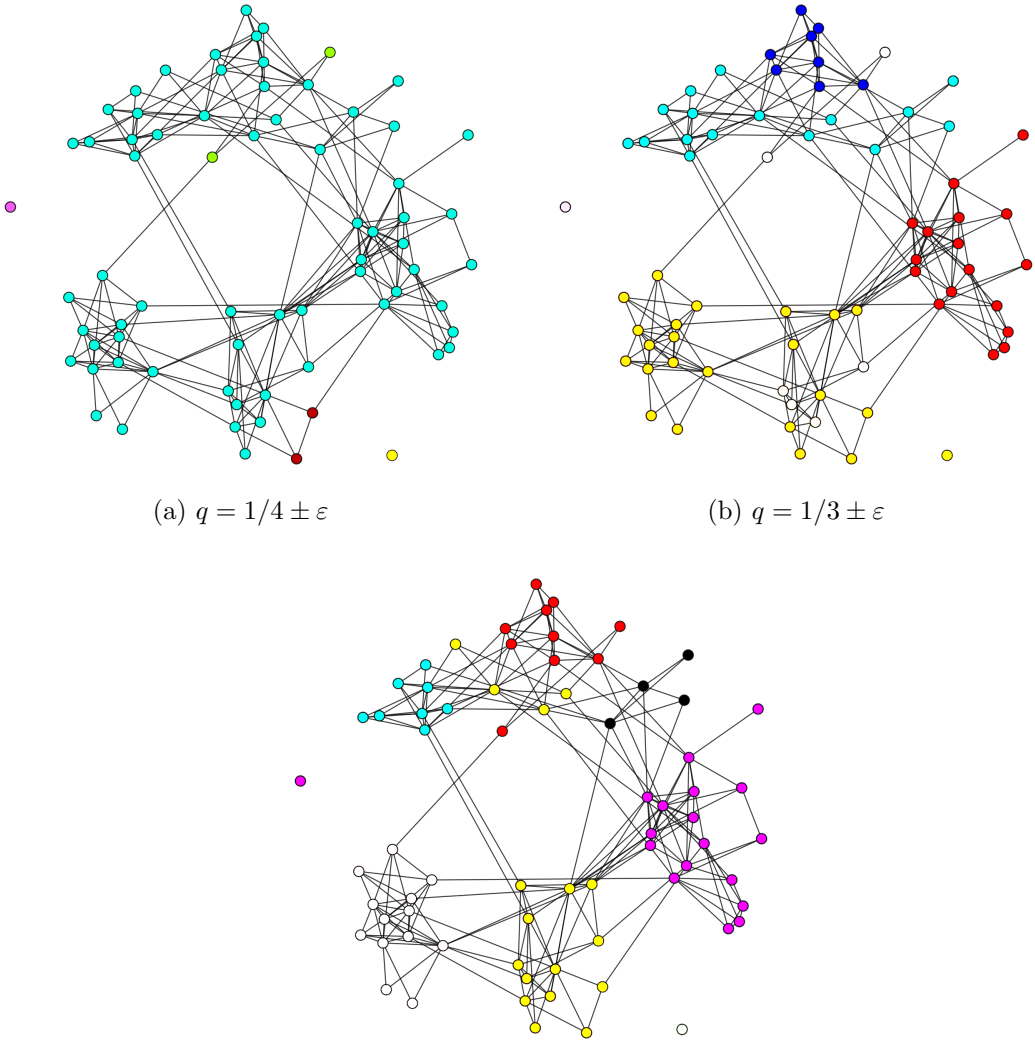
Figure 7 illustrates how the atomic structure coarsens as we expand the robust-Q range.



(a) $q = 1/2 \pm 1/10$ (b) $q = 1/2 \pm 1/5$

Figure 7: **:** How the robust atomic structure for the high school social network coarsens as Q is widened.

We can also compare how the communities change as we vary the q at which our robust community structures are centered, as seen in Figure 8. With these non nested changes in Q, the atomic structures are no longer nested.



(a) $q = 1/4 \pm \varepsilon$

(b) $q = 1/3 \pm \varepsilon$

(c) $q = 1/2 \pm \varepsilon$

Figure 8: **:** The atomic partitions for the high school social network for various q's plus or minus ε.

3.2 Conventions for q and $1 - q$

The conventions for q and $1 - q$ coincide (provided these are irrational). The following is a corollary to Observation 1.

OBSERVATION **3** *If q is irrational (or such that no agent could ever be indifferent so are rational but with a denominator different from all multiples of agents' degrees), then $C(q, g) = C(1 - q, g)$.*

This observation is useful as it means that when we are interested in conventions that are robust to a set of thresholds that lies to one side of $q = 1/2$, then understanding what

happens for small conventions implies that we know the conventions for the case of large conventions.

Interestingly, however, even though the behavioral communities in Figure 4 are the same for $Q = (1/3, 1/2]$ and $Q = (1/2, 2/3]$, those are not the behavioral communities for $Q = (1/3, 2/3]$. That is because, the conventions for $Q = (1/3, 1/2]$ and $Q = (1/2, 2/3]$ are complements of each other, but only some are robust conventions for all of $Q = (1/3, 2/3]$: essentially only the conventions for which both it and its complement are conventions for both of the sub intervals. This is pictured in the Figures 9 and 10.

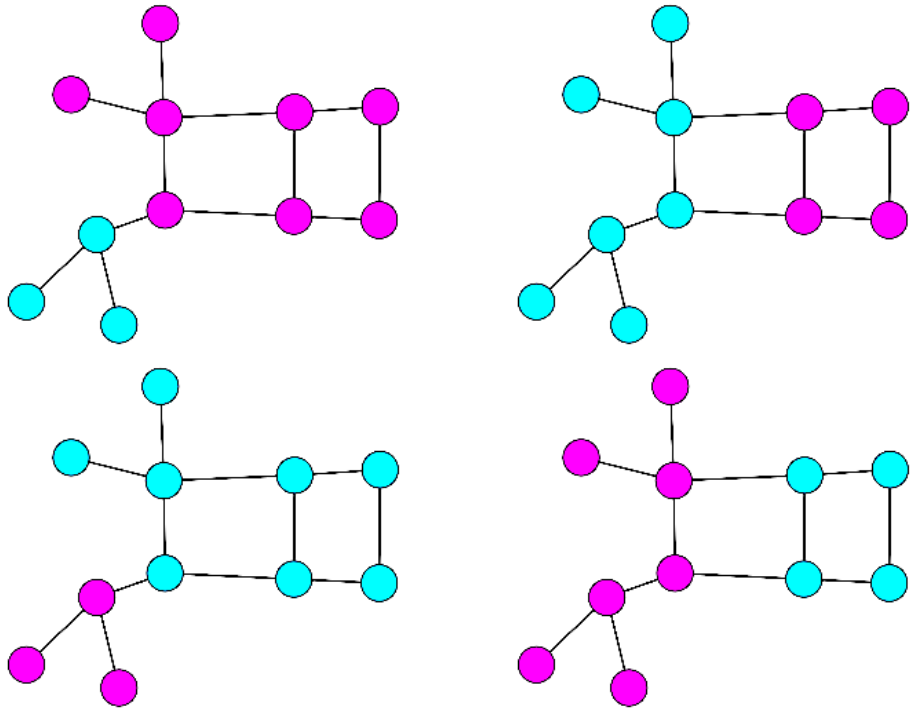


Figure 9: **Robust Conventions:** The only conventions that are robust when $Q'' = (1/3, 2/3]$.

A robust convention for $Q = (q, 1-q]$ for some $q < 1/2$ is necessarily such that its complement is also a robust convention for Q.

3.3 Tight Sets and Atoms are Distinct

Is there an easy way to characterize atoms? There are many subtleties.

A first observation is that conventions can have subsets that still form a convention. Thus, one way to find atoms might be to look at minimal conventions. So, let us define minimal conventions.

We say that a group $S \subset N$ is called *q-tight* if it is q-cohesive and q-closed, *and* has no nonempty strict subsets that are q-cohesive and q-closed.

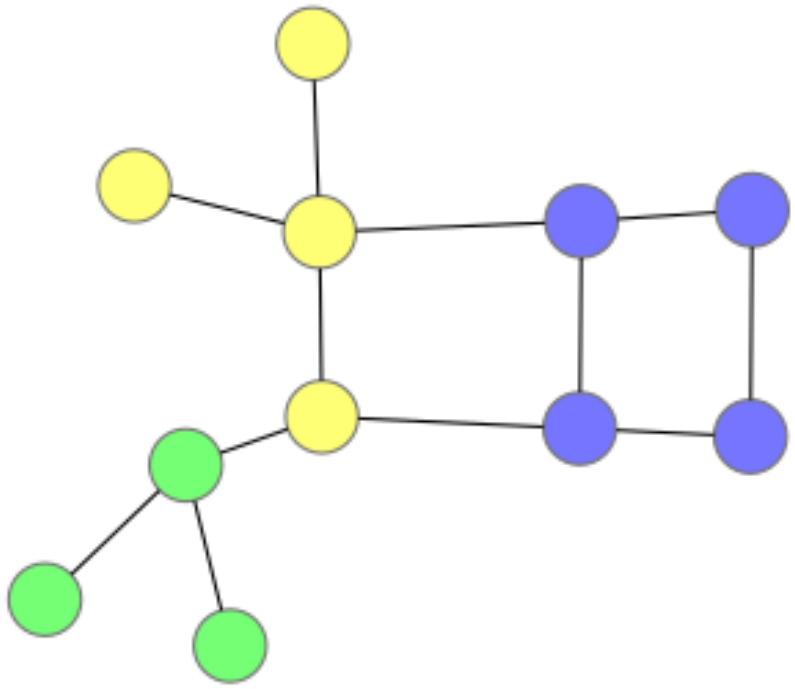


Figure 10: **Robust Behavioral Communities:** The behavioral communities for for $Q'' = (1/3, 2/3]$. Note the contrast with Figure 4.

A group S forms a convention with a threshold q and has no strict subsets that form a convention if and only if it is q-tight. It is then called a *tight convention*.

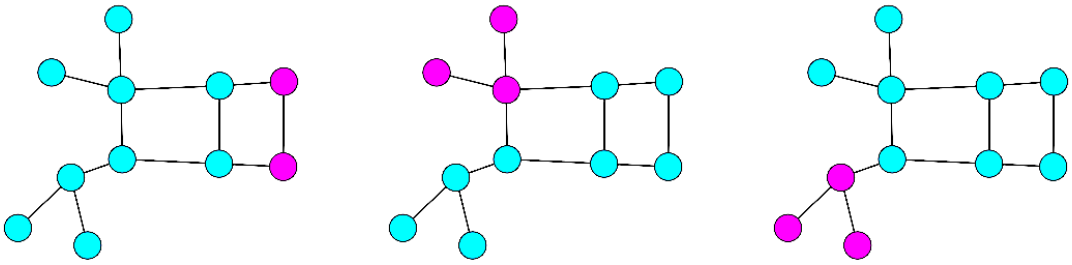


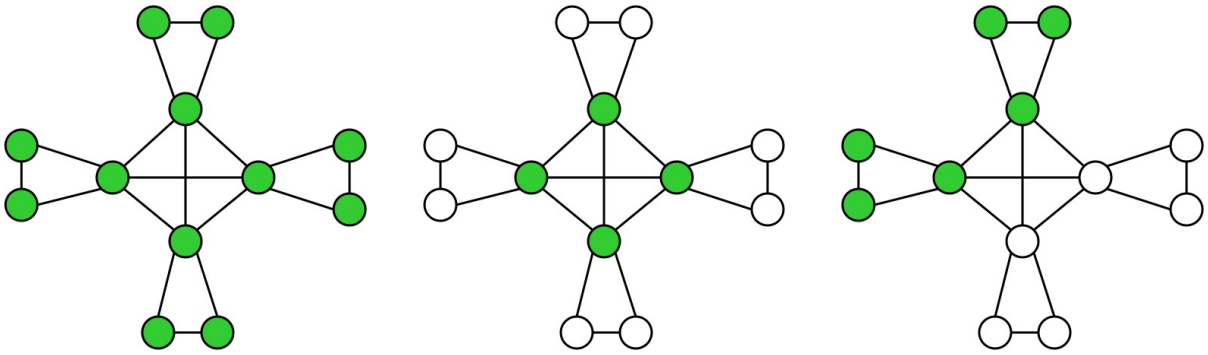
Figure 11: **Tight Conventions:** The tight conventions when $q = .4$.

It seems logical to conjecture that q-tight groups (at least when $q > 1/2$) would form atoms. The example pictured in Figures 12 and 13 shows the first subtlety: atoms can not only divide, but can actually subdivide q-tight groups.

The example from Figures 12 and 13 shows a related second subtlety: there are networks for which no atom is a convention.

A third subtlety is that not only is it possible for tight groups not to be contained in any atom, but it is also possible to have completely disconnected nodes form an atom.

For instance, pairs of nodes that are structurally equivalent (have the same friends) are



(a) A convention.

(b) A convention with the clique of four nodes in the center.

(c) A convention involving the upper left half of the nodes.

Figure 12: **Splitting a tight group:** Even though the central four nodes form a clique and are $q = .6$-tight as a group, they can be split apart in some conventions.

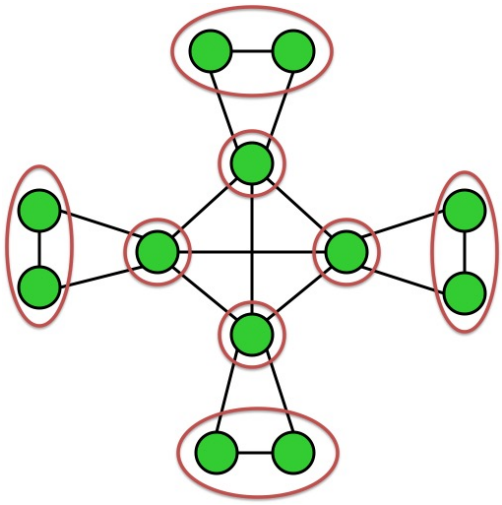
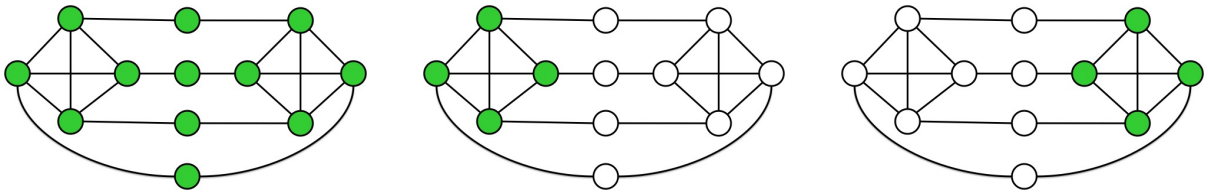


Figure 13: The partition into communities when $q = .6$. Even though the middle group forms a clique and is $q = .6$ tight, it is not an atom and its members are split apart in the community partition.

in the same atom regardless of whether they are friends with each other. In fact, one can go even further: there are atoms in which none of the nodes in the atom have any neighbors in common. The example in Figures 14 and 15 shows such a situation, in which there are a number of nodes that are similar to each other in terms of positions in the network and so end up in the same atom, but none of the nodes in the atom even have a friend in common.[6]

Of course, there are some obvious properties that atoms have. For instance, they must be finer than components, so there is some limitation on how disconnected atoms can be. This

[6]Thus, they are not structurally equivalent in the sense of Lorrain and White (1971) - these nodes satisfy a weaker form of similarity in position.



(a) All nodes infected. (b) One atom infected (c) The other atom infected.

Figure 14: **An atom that is never a convention:** The three nonempty conventions when $q = 3/4$.

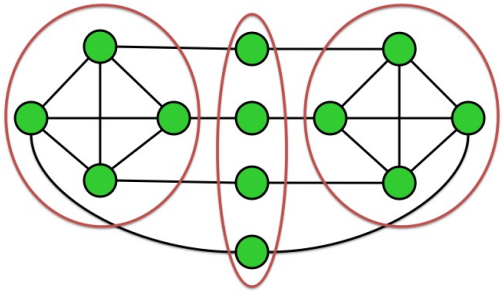


Figure 15: The partition into communities when $q = 3/4$. Note that there are three atoms, but the middle atom is never a convention on its own. In fact, none of the nodes in the atom are even connected to each other, nor do they even have any neighbors in common. However, those four nodes are structurally similar to each other in terms of having similar sorts of connections to the other atoms, which means that they always end up acting in exactly the same way.

observation follows since if two nodes are not path connected, then their different components are each a convention. Thus, they sometimes behave differently, and so are not in the same atom.

There are more things that we can deduce about atoms.

We say that a set S is *q-isolated* if it does not intersect any different q-tight group.

The following lemma provides one fact about atoms.

LEMMA **2** *If a set S is q-tight and q-isolated, then it forms an atom relative to q.*

Lemma 2 is a corollary to an even more straightforward lemma (whose proof is thus omitted).

Let us say that two nodes are *structurally similar* (relative to q) if whenever one is in some q-cohesive and q-closed group then so is the other.

LEMMA **3** *The sets of structurally similar nodes form equivalence classes which are the atoms.*

Lemma 2 follows from Lemma 3 since all of the nodes in a tight and isolated set are necessarily structurally similar, and cannot be structurally similar to any node outside of that set.

Let an atom be called *internal* relative to q if there exists a convention which coincides with the atom. In Figures 14 and 15, the two atoms on the left and the right are internal atoms.

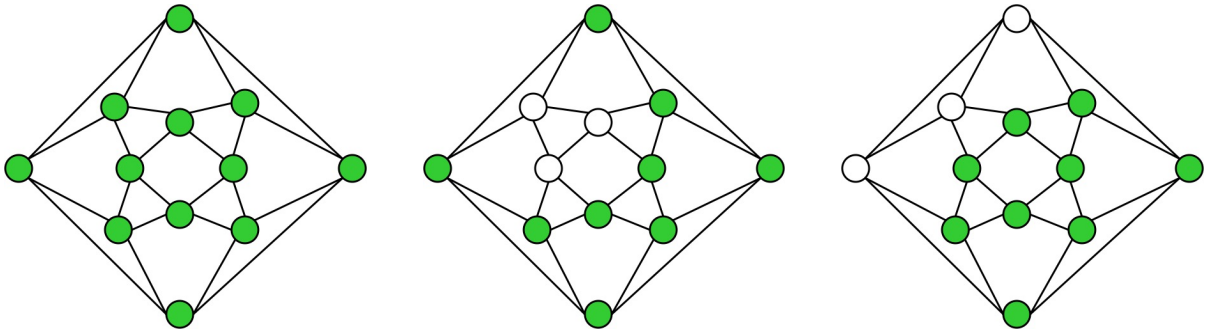
The internal atoms are 'intuitive' ones that are minimal conventions. The idea is that the play of people in that atom have to be the same because of the contagion of behavior *within* the atom.

Note that the internal atoms are exactly the q-tight and isolated sets of nodes. Thus, one class of atoms has an intuitive characterization.

The remaining atoms are *external* ones. In Figures 14 and 15, the middle atom is an external atom. It is an atom whose structure is determined by structural similarity and positions relative to other nodes that are sufficiently similar that the nodes have to play the same way.

It is not simply that nodes that can be mapped onto each other in some permutation must be in the same atom, as we see in the next example. There is complete symmetry in the network and yet no two nodes are in the same atom.

The example in Figures 16 and 17 is an interesting one since it has features different from either of the previous examples. There is a full symmetry to the network and lots of equivalences between nodes, and many tight groups, and yet the partition is into atoms of singletons. This is the last subtlety that we point out.



(a) A convention.

(b) A convention excluding one triple.

(c) A convention excluding a different triple.

Figure 16: **:** The conventions when $q = 3/4$ are either the whole network, or that excluding some triple that form a triangle .

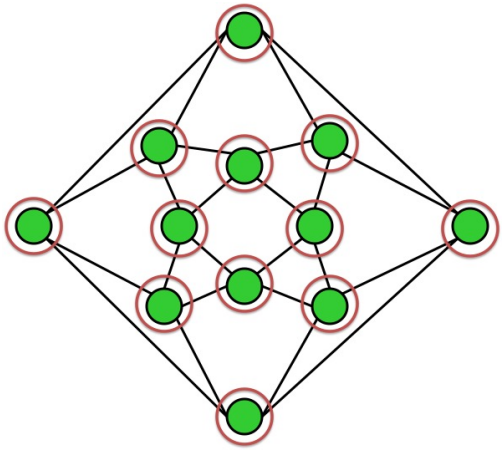


Figure 17: The partition into communities when $q = 3/4$. Even though the conventions are generally large subgroups of the community, and the network has strong symmetry properties, all atoms are singletons.

4 Community Structures in some Random Network Models

As we see above, atoms can be quite interesting sets. Beyond the intuitive interior atoms, there are exterior atoms that are dependent upon the graph structure in complex ways, which makes an abstract characterization difficult. Nonetheless, there is a lot that we can deduce about them in random graphs.

To get more of a feeling of how atoms behave, we prove some results about how they act in a basic and well-known class of random networks that exhibit natural divides between nodes: a stochastic block model. Such a model has different 'types' of nodes (the blocks) and allows for tunable levels of homophily in the network for various combinations of types of nodes. These models are natural ones for modeling and analyzing homophily, for instance (e.g., see Golub and Jackson (2012)).

We show that under 'natural' choices of q, the atoms are subsets of the blocks, while for any fixed t, the atoms coalesce to coincide with the whole network. This shows that the blocks capture natural divisions in behavior and norms when behavior depends on *relative* ratios of friends taking behavior; but if there is only some minimal threshold needed then behavior necessarily spreads across blocks.

4.1 A Stochastic Block Model

Let n index a sequence of random graph models, tracking the number of nodes in the society.

The society is partitioned in different types of people or nodes indexed by $J(n)$, with

generic indices of types j, j', and cardinalities $j(n)$. These might refer to demographic characteristics like age, religion, gender, ethnicity, profession, etc. So a particular type might be female Hispanic lawyers.

Let $\Pi(n)$ denote the associated partition of the nodes by types.

The probability that any node of type $j \in J(n)$ is linked to a node of type $j' \in J$ is given by some $p_{jj'}(n) = p_{j'j}(n)$. Links are independent across all pairs of nodes.

Let $d_{jj}(n) = p_{jj}(n)(j(n) - 1)$ and $d_{jj'}(n) = p_{jj'}j'(n)$ denote the expected links of a type $j \in J(n)$ to types j and j', respectively, in society n, and $d_j(n) = \sum_{j' \in J(n)} d_{jj'}(n)$ be the overall expected degree of a type j node.

Let g^n denote a random network generated on the n nodes.

4.2 Behavioral Communities in Stochastic Block Models

4.2.1 Fraction-based Threshold Communities

We first show that q-based – 'fractional' – communities are subsets of the blocks in a block model when the blocks are sufficiently well-separated, in a sense which we make precise below.

THEOREM 1 *Consider a growing sequence of stochastic block networks for which there exists $f(n) \to \infty$ such that $d_{jj'}(n) > f(n)\log(n)$ for all jj', and for which there there exists some $\overline{q} > \underline{q} > 0$ such that:*

- *$d_{jj}(n)/d_j(n) > \overline{q}$ for each $j \in J(n)$, and*
- *$d_{jj'}(n)/d_j(n) < \underline{q}$ for each $j' \neq j$.*

Then, for any compact Q, $Q \subset (\underline{q}, \overline{q})$ or $Q \subset (1-\overline{q}, 1-\underline{q})$, the partitions corresponding to $C(Q, g^n)$ are weakly finer than $\Pi(n)$ with a probability going to 1 as n grows.

Thus, behavior-based communities related to behaviors that depend on the fraction of one's neighbors adopting a behavior are a subset of the blocks.

Theorem 1 implies that if there is homophily in a network, so that there are some characteristics that drive differences in linking probabilities, then the behavioral atoms will uncover that homophily and possibly more.

In particular, it is worth noting that Theorem 1 admits the possibility that behavioral atoms will sometimes be finer than the breaks in the graph associated with differences in underlying linking frequencies. Those are often the communities found via the popular modular methods (e.g., the Louvain method of Blondel et al. (2008)), and so our atoms will sometimes be finer than those blocks/modules. In the case of block models, many modular methods will pick up the blocks and not subsets of those blocks (e.g., see Copic et al. (2009)). Those methods maximize some measure of the density of links within a community to those going out of a community. The reason that behavioral communities will sometimes be finer is

that the cohesiveness condition is a node-by-node condition and not a total weight condition, and similarly, closure is a requirement that no node have more than a certain connectivity with a set, not an average condition. Thus, there can be fractures within a block that allow for a convention to divide that block.

This is an important distinction – even in cases where the graph is generated with well-defined blocks, it can end up with additional purely randomly generated and subtle structure that actually splits nodes apart in terms of their behaviors. The fineness of the behavioral atoms is picking up additional divisions in the network that have real consequences but are not captured in the basic blocks.

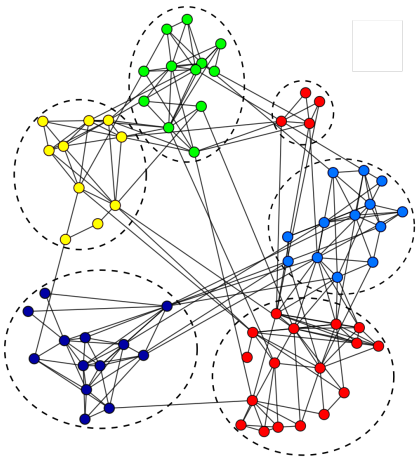
It is important to emphasize that the additional fineness of behavioral atoms capture true divides in the realized network that allow conventions to split blocks. The additional fineness is capturing real behavioral divides that go beyond the block structure and thus are not defects in terms of defining atoms but real information that behavioral atoms convey that other methods do not. [7]

A Comparison of Behavioral Communities and Modularity-Based Community Detection Algorithms

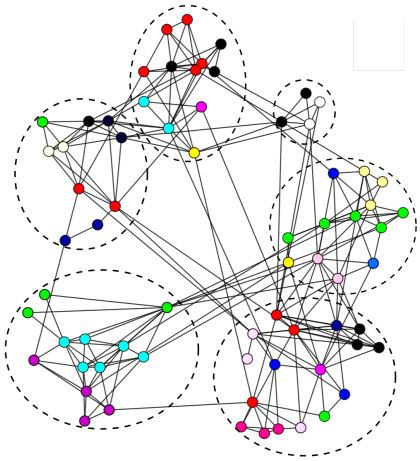
We demonstrate that behavioral atoms can subdivide communities defined by standard community detection algorithms, such as modular methods that work off of relative link densities. We use the high school social network from Figure 7.

Figure 18 depicts the communities identified by a modularity-minimizing algorithm (specifically, the Louvain method for community detection, Blondel et al. (2008)) via the dotted outlines (and color in subfigure (a)), while the colors in the other subfigures represent the behavioral atoms.

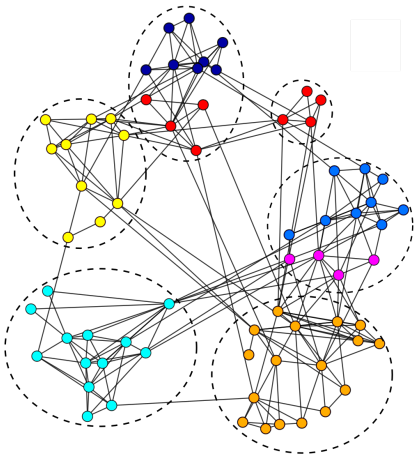
[7] Another distinction is that behavioral communities change with the threshold and Q, while modular methods are just looking for higher internal than external density, rather than requiring specific conditions on those densities.



(a) The communities found via modularity minimization



(b) The $q = 1/2$ behavioral atoms (color) compared with the modular communitities (outline)

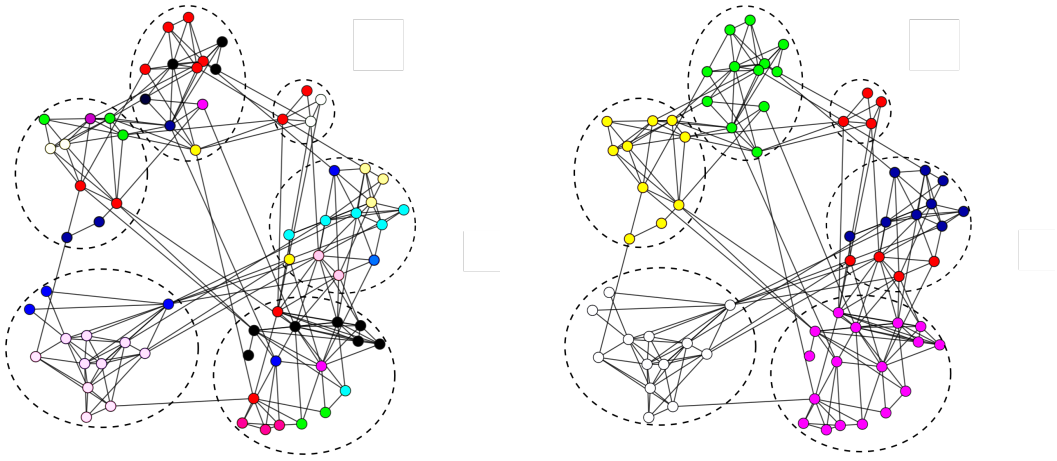


(c) The $Q = 1/2 \pm \varepsilon$ robust behavioral atoms (color) compared with the modular communities (outline)

Figure 18: **:** Comparing the behavioral atoms to the communities found via the Louvain method of modularity minimization (Blondel et al., 2008).

In Figure 18 we see that the $q = 1/2$ atoms fracture the modules quite a bit, but partly because of the exact nature of the fraction 1/2 with can leave some nodes exactly at indifference in some situations. If we examine a robust version (plus or minus $\varepsilon < 1/n$), then we get a clearer picture. The behavioral atoms at this $Q = 1/2 \pm \varepsilon$ level are subsets of the modules and in two cases split the modules/blocks into parts. This means that within those blocks it would be possible to have a robust convention that cuts across those blocks and has part of the block acting differently from the other part.

This effect becomes even starker when we look at $q = .45$, as in Figure 19 Here we get a more dramatic fragmentation of the modules. We also see that increasing the range of robustness makes a difference, and if the range of Q is increased enough then our atoms coincide with the modules, except for one module/block that can still support multiple subdividing robust conventions.

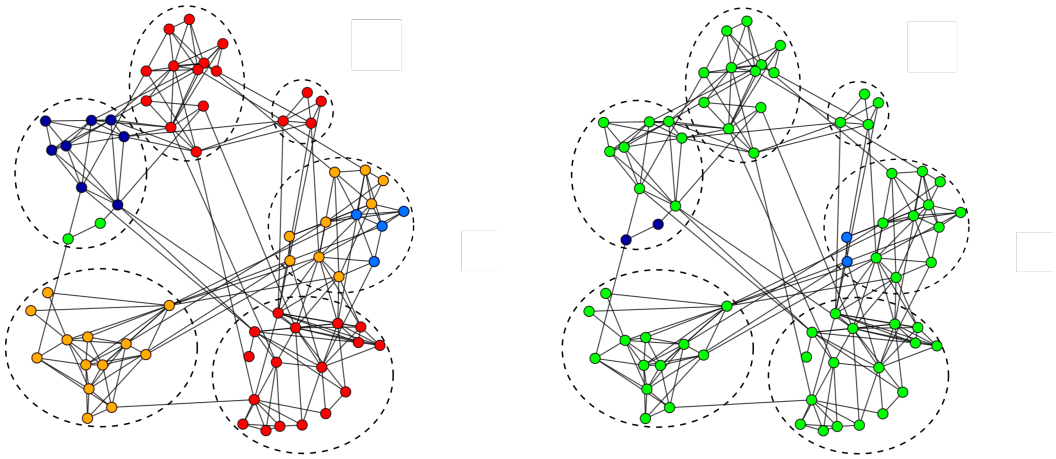


(a) The $Q = 0.45 \pm \varepsilon$ behavioral atoms (color) compared with the modular communitites (outline)

(b) The $Q = 0.45 \pm 1/10$ robust behavioral atoms (color) compared with the modular communities (outline)

Figure 19: **:** Comparing the behavioral atoms to the communities found via modularity minimization.

Interestingly, when we lower the q, as in Figure 20, then we see that the atoms actually go *across* modules/blocks. This happens for values of q that are lower than the frequency of some cross-block links. For behaviors with low thresholds, contagion will occur across blocks and so it is impossible to have atoms that are restricted to some blocks. For this network, a behavior that has a $q = 1/4$ threshold cannot be contained by any blocks – and the atomic structure reflects that.



(a) The $q = 1/3 \pm \varepsilon$ behavioral atoms (color) compared with the modular communitites (outline)

(b) The $Q = 1/4 \pm \varepsilon$ robust behavioral atoms (color) compared with the modular communities (outline)

Figure 20: **:** Comparing the behavioral atoms for lower q to the communities found via modularity minimization. With low q, communities necessarily connect across blocks. It is only when q is above the cross-block density that atoms fall within blocks.

The difference between, for instance, a modular method and the behavioral method is that a modular method only works off of *average* density. Instead, *individual* link patterns determine behavior rather than averages. If one pays attention to individual patterns then finer fissures appear that are not captured by standard algorithms that operate on aggregate statistics, and the possibility of contagion across blocks emerges.

Coincidence of Atoms and Blocks in Strongly-Distinguished Block Models There are some conditions under which the atoms coincide with the blocks. The blocks have to be sufficiently distinguished so that contagion does not occur across blocks, and have a high enough density so that they coalesce and have no internal fissures. Under those conditions, for an appropriate Q, the atoms will coincide with the blocks, as captured in the following corollary.

COROLLARY **1** *Consider a growing sequence of stochastic block networks satisfying the conditions of Theorem 1 and for which* $3\overline{q} > 1 + \underline{q} + 2\varepsilon + \frac{j(n)}{d_j(n)}$ *for some* $\varepsilon > 0$*. The partition corresponding to* $C([\underline{q} + \varepsilon, \overline{q} - \varepsilon], g^n)$ *coincides with* $\Pi(n)$ *with a probability going to 1 as n grows.*

The sufficient condition that $3\overline{q} > 1 + \underline{q} + 2\varepsilon + \frac{j(n)}{d_j(n)}$ for each j – a 'sparse-dense block model' – requires a high level of connectedness within blocks, and is a sufficient condition. The blocks become robust conventions themselves and all robust conventions are supersets

of the blocks. Thus, with a probability going to 1 in n, observing as few as $\#J(n)$ different robust conventions allows one to recover $\Pi(n)$.

The condition on the density within blocks prevents blocks from fragmenting under a behavior has a threshold of no more than $\overline{q}-\varepsilon$. Having the behavior's threshold be at least $\underline{q}+\varepsilon$ ensures that the behavior does not leak across blocks and a convention can be sustained within a block.

The conclusions in Theorem 1 contrast with the atomic structure associated with behaviors derived from some absolute threshold t. There, as n increases, even with a relatively slowly growing and sparse density, behavior is contagious within and across blocks, so that the whole network becomes the only convention, as we show next.

4.2.2 Absolute Thresholds

Behaviors based on absolute thresholds generally have contagion across blocks – provided that blocks have some minimal numbers of interconnections (just above the sharp threshold for inter-connectedness at a rate t).

THEOREM **2** *Consider some finite $t>0$ and a growing sequence of stochastic block networks. If there exists $\varepsilon>0$ for which $p_{jj'}(n)>(1+\varepsilon)\left(\frac{t\log(n)}{n}\right)^{1/t}$ for each j',j, then with a probability going to 1 as n grows the partition corresponding to $C(t,g^n)$ is the degenerate one generated by the atom of N.*

Clearly, if $p_{jj'}$ falls substantially below the threshold for enough jj' pairs then the graph will fragment. Provided the probabilities internal to blocks is large enough to generate conventions within the blocks, then the community structure will be non-degenerate.

We prove Theorem 2 by proving the following theorem about random graphs, which is of independent interest.

We could not find Theorem 3 in the graph-theory literature, so we have proven it directly. It's proof methodology also appears to be new and may also be of independent interest.

4.3 k-cores of Random Graphs

Following standard definitions, a k-core of an undirected graph g is a maximal subgraph, including fewer than n nodes, such that all nodes in the subgraph have degree of at least k within the subgraph.

Note that when a nonempty k-core exists, then it must be that the k-core forms a convention for threshold $t=k$.

We define a *weak* k-core to be a nonempty subgraph, including fewer than n nodes, for which all nodes in the subgraph have degree at least k within the subgraph and for which no single node could be added and have degree at least k.

Note that the set of weak k-cores are exactly the set of conventions (other than the whole set N) for threshold $t = k$.

A k-closed set is a nonempty subgraph, which has at least k nodes and fewer than n nodes, for which all nodes outside of the subgraph have fewer than k connections to the nodes in the subgraph.

Note that every (weak) k-core is k-closed, but the converse is not true as being k-closed does not require that nodes in the subgraph have degree at least k.

We use the standard notation $G(n, p)$ to indicate an Erdos-Renyi random graph on n nodes with a i.i.d. probability $p(n)$ of any link existing.

THEOREM **3** *Consider a growing sequence of Erdos-Renyi random graphs $G(n, p)$.*

- *If $p(n) > (1 + \varepsilon)\left(\frac{k \log(n)}{n}\right)^{1/k}$ for any $\varepsilon > 0$, then the probability that a k-closed set exists goes to 0 (and thus so does the probability that there exist any weak k-cores or k-cores, or (tight) conventions that have $k = t$ and involve less than all nodes).*

- *Conversely, if $p(n) < (1 - \varepsilon)\left(\frac{k \log(n)}{n}\right)^{1/k}$ for any $\varepsilon > 0$, then the probability that a k-closed set exists goes to one.*

$\left(\frac{k \log(n)}{n}\right)^{1/k}$ is what is known as a *sharp threshold* in random graph theory.

When $k = 1$ it reduces to the threshold for connection in an Erdos-Renyi random graph.

Note that even though Theorem 3 is about Erdos-Renyi random graphs, and Theorem 2 is about stochastic block models, stochastic block models can be constructed by starting from an Erdos-Renyi random graph, and than adding additional links within blocks (and possibly across some pairs of blocks). Thus, Theorem 2 is a corollary.

Once $p(n) < (1-\varepsilon)\left(\frac{k \log(n)}{n}\right)^{1/k}$ the number of k closed sets of sized k, as well as $k+1$..., goes to infinity. In that case, the existence of a (weak) k-core just requires that the probability is large enough so that out of that infinite sequence of such sets at least one forms a clique. Thus, as long as the link probability does not drop so low that such cliques disappear, there will exist a (weak) k-core.

The proof technique that we use is based on showing that the probability of having a k closed set of k nodes at the threshold of $\left(\frac{k \log(n)}{n}\right)^{1/k}$ can be bounded by the probability that there exists an isolated node at the threshold of $\frac{\log(n)}{n}$. This is useful given that the behavior of this other event is well-known, while the first event is not and involves more intricate correlations. These bounds turn out to be tight and so are useful in proving this theorem. We have not seen this technique, of bounding the probability of one class of events in one random graph model by bounding it by the probability of a different class of events in a different random graph model, used before.

5 Additional Empirical Examples

Here we show the atomic structures for some other network data sets, and compare those partitions to partitions of the network based on node demographics.

5.1 Indian Villages

Figure 21 pictures the network of social and financial relationships in an Indian village (data taken from Banerjee, Chandrasekhar, Duflo, and Jackson (2013)). The shape of each node represents its caste, and the number in the node designates its subcaste.

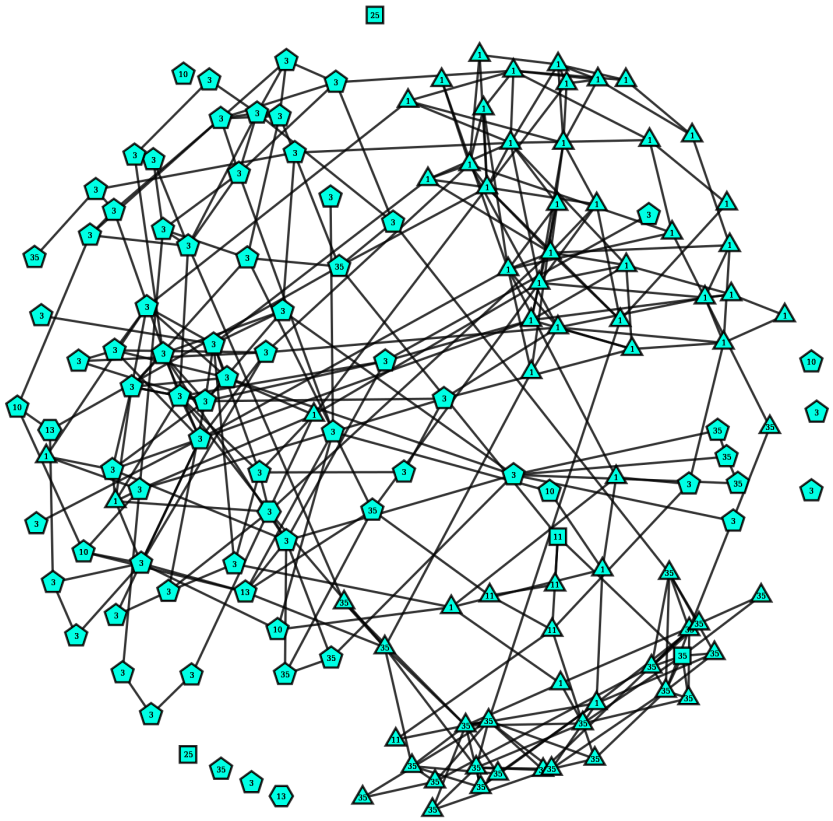
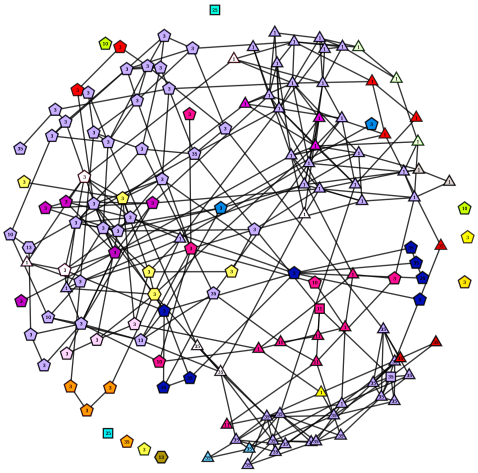
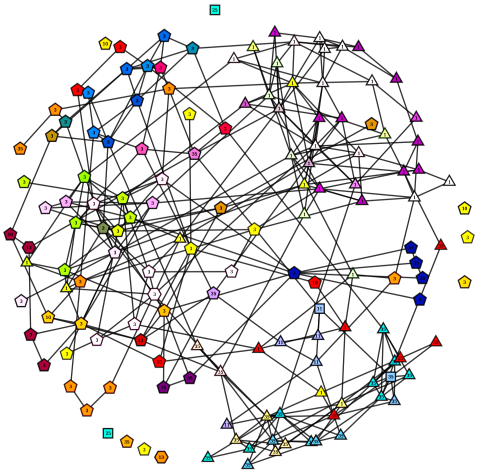


Figure 21: Network of Social Relationships in an Indian Village

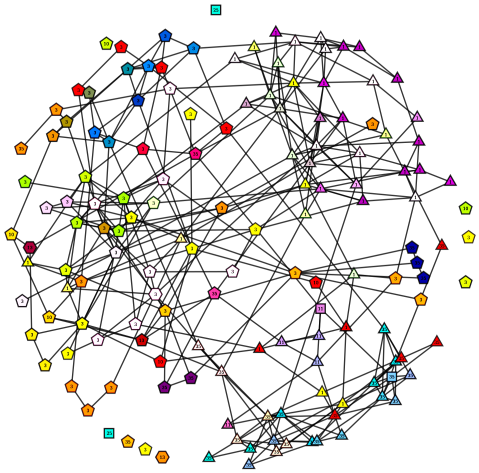
Figure 22 pictures some q atomic partitions for this social network.



(a) q=1/4



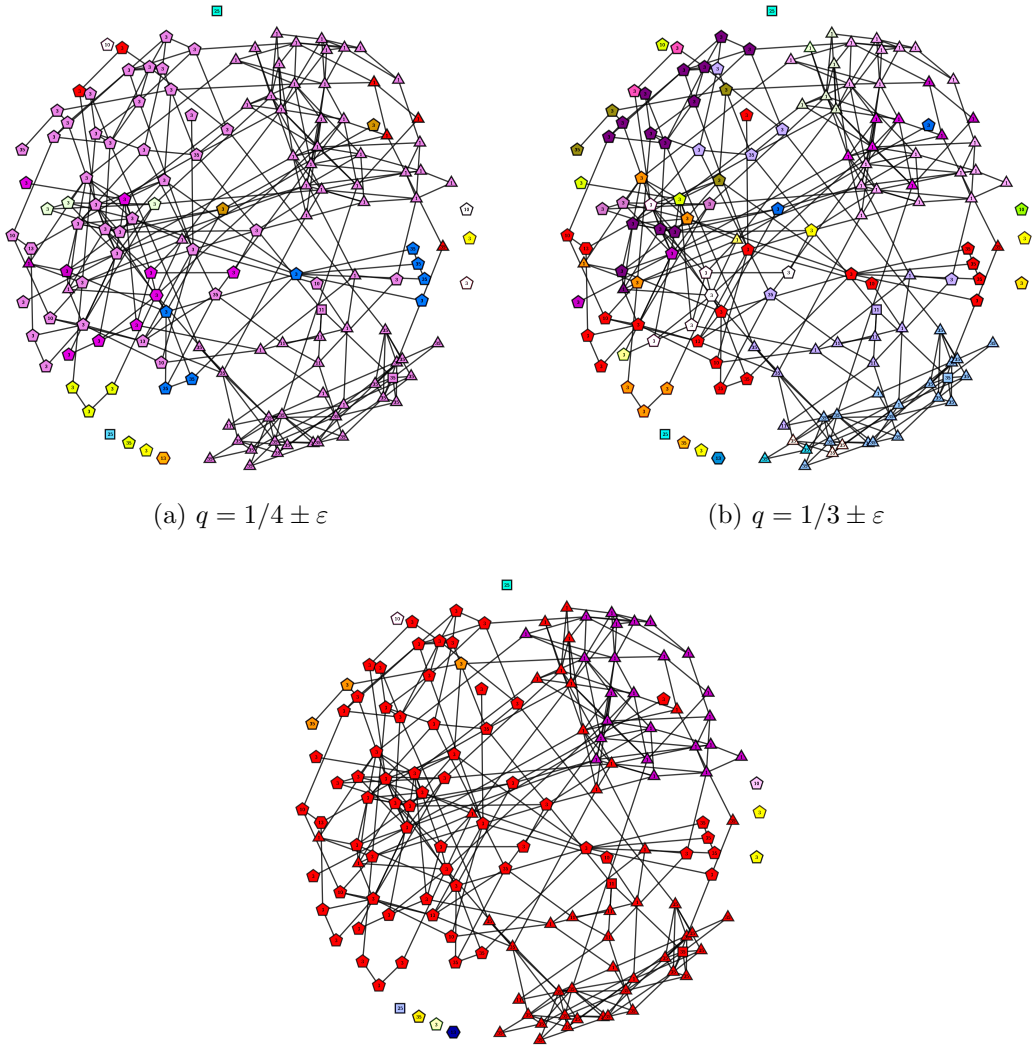
(b) q=1/3



(c) q=1/2

Figure 22: **:** The atomic partitions for an Indian village for various q's.

When we expand from the single q's to small robust intervals about these points, the partitions coarsen noticeably, as pictured in Figure 23.



(a) $q = 1/4 \pm \varepsilon$ (b) $q = 1/3 \pm \varepsilon$

(c) $q = 1/2 \pm \varepsilon$

Figure 23: **:** The atomic partitions for an Indian village as we expand about the single q points above.

Observe that the $1/2 \pm \varepsilon$ division picks up the subcaste division in the triangle caste: the triangles in subcaste 35 are in the same behavioral atom as the majority hexagon caste (the red atom), while most of the triangles in subcaste 1 are in their own separate behavioral atom (in purple). This suggests that behavioral atoms can pick up subtle social divisions.

In this particular network, expanding the Q ranges by any more than $1/20$ forces the entire main component to coalesce into a single atom.

6 Finding the Atoms

In this section, we describe a polynomial-time approximation algorithm that nevertheless recovers the atomic decomposition with high probability for some types of random graphs. We describe the algorithm for the q (fractional threshold) setting, but the logic for the t (absolute threshold) setting is the same.

Finding the decomposition of a network into its behavioral atoms by exhaustively checking whether each subset of nodes forms a convention is computationally infeasible, even for moderately sized networks, as the number of subsets is exponential in the number of nodes. Indeed, finding atomic decompositions appears to be np-hard problem. Nonetheless, an algorithm works well in many settings in which the network has a robust atomic structure to it.

The algorithm starts by forming the collection of connected subsets of nodes of size less than some fixed k. The number of such subsets is at most $\binom{n}{k}$, and we will show later that k can be fixed independent of n, so that the number of subsets formed is polynomial in n. The algorithm then selects a subset and generates its minimal Q-closed superset. If that superset also happens to be Q-cohesive, then it is a convention, and the algorithm stores the convention and moves on to the next subset in the working collection. If not, the algorithm adds to the superset the node whose addition most increases the node-wise minimum level of cohesion in the resulting subset, and then reiterates the preceding procedure with this new subset in place of the original subset. Since each such iteration adds at least one node to the subset, the algorithm terminates in at most $\binom{n}{k} \times n$ steps.

Under the conditions of Theorem 1 and Corollary 1, when the atoms turn out to be the blocks (which in that case happens with a probability converging to 1), then the algorithm will identify those blocks as atoms. Picking a k-node subset of a block will necessarily result in the block being found as a convention by the way the algorithm is designed, and so the blocks will be identified (with a probability going to 1).

7 Optimal Seeding

The behavioral atoms tell us how coordination behaviors can be distributed across a social network. It is thus natural that the atomic partition also provides useful information for efforts to influence those behaviors. In this section we explore the how knowledge of behavioral atoms helps in seeding diffusion.

Given a social network g, a behavioral threshold q, and an initial budget of k nodes to be chosen as "seeds," the optimal seeding problem is to find a set of k initial adopters which leads to the largest number of nodes in g eventually adopting the behavior, assuming that all nodes other than the initial adopters sequentially best respond to their neighbors' adoption decisions.

A solution to this problem is challenging because the fully optimal seeding depends

on details of the network structure: even adding or removing a single link can change the optimal set of nodes completely. While we cannot offer a fully optimal solution to the seeding problem, which is a variation on ones that are known to be NP-hard (Kempe, Kleinberg, and Tardos (2003, 2005)), we can show that the behavioral atoms inform an intuitive heuristic for the seeding problem that offers significant improvements over random seeding.

This contrasts with straight contagion settings ($t = 1$), where random seedings can do quite well (e.g., Mønsted et al. (2017)). Here, the seeding problem is complicated by the fact that behavior depends on several neighbors taking an action rather than just one. Thus, placing several seeds close together is necessary to get this sort of contagion process started, while dispersing them randomly is better in a simple contagion process.

In a network with distinct behavioral communities, the seeding problem is even more complex since spreading may not cross boundaries, as our analysis above shows. Thus, our behavioral atoms provide a natural boundary for this problem: we have to seed enough nodes in an atom to get the atom to adopt, and at the same time we have to hit enough distinct atoms to spread the behavior. This motivates the following "greedy" atom-wise seeding heuristic:

1. Find the q-atoms of G. Order them $A_1, A_2, ...A_m$.

2. For each atom A_i, find the minimal number of seeds needed to turn the entire atom on. Call this the cost of the atom, c_i.

3. Greedily seed the atoms in decreasing order of the size-to-cost ratio $|A_i|/c_i$ until we have used all k seeds (skipping over any atoms that have seeding cost in excess of k) or there are not atoms left that can be seeded with the remaining number of seeds.

4. If there are seeds left over, select seeds uniformly at random from the set of nodes which are not in the q-closure of the set of seeds already selected.

(Note that step 2 can be accomplished by brute force computation if none of the atoms is too large.) For moderately sized-networks, we can compute the optimal seeding by brute-force to check how close this heuristic comes to the optimum spread. We can also compare the heuristic's results to those attained by random seeding. The table below shows this comparison for randomly generated networks obtained by starting with an Erdos-Renyi graph and then randomly closing triangles to obtain a target clustering coefficient of 0.2. Behavior spreads by iterative best response with $q = 1/2$.

n	Linking Probability	k	Heuristic Average Share of Optimum	Random Seeding Average Share of Optimum
20	0.1	5	0.73	0.50
20	0.3	5	0.81	0.47
20	0.5	5	0.86	0.41
40	0.1	8	0.86	0.42
40	0.25	8	0.80	0.35
40	0.3	8	0.82	0.30
40	0.1	15	0.75	0.60
40	0.25	15	0.78	0.52
40	0.3	15	0.78	0.44

We see that the greedy heuristic tends to achieve around 80 percent of the optimal spread, while random seeding only manages about 50 percent. These results suggest that the behavioral atoms do in fact provide useful information for the seeding process.

We also perform similar simulations in the stochastic block model, and obtain similar results:

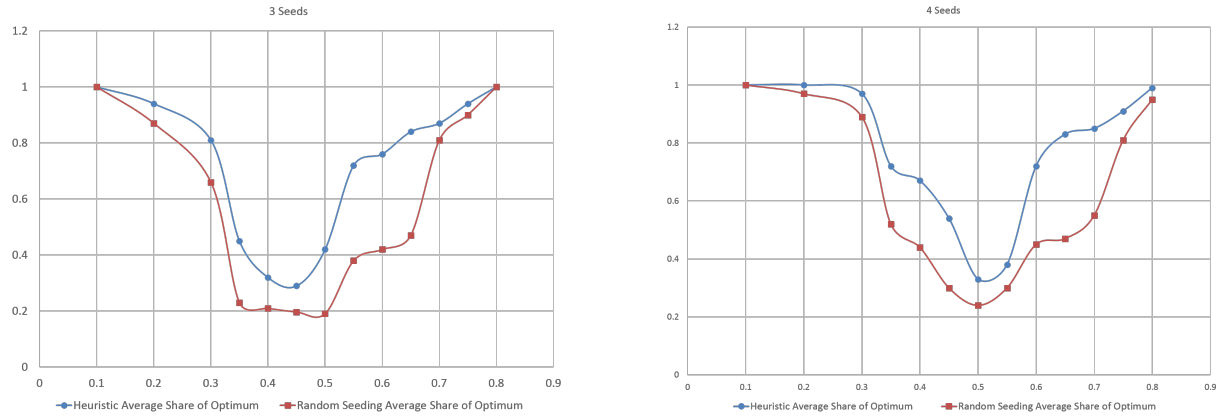
Num Blocks	Block Size	Density within Blocks	Density Across Blocks	k	Heuristic Average Share of Optimum	Random Seeding Average Share
2	20	0.2	0.1	5	0.77	0.49
2	20	0.5	0.1	5	0.73	0.38
2	20	0.2	0.1	8	0.81	0.62
2	20	0.5	0.1	8	0.79	0.51
2	30	0.2	0.1	5	0.72	0.44
2	30	0.5	0.1	5	0.68	0.59
2	30	0.2	0.1	8	0.77	0.54
2	30	0.5	0.1	8	0.82	0.59
3	20	0.2	0.1	5	0.72	0.42
3	20	0.5	0.1	5	0.76	0.38
3	20	0.2	0.1	8	0.81	0.49
3	20	0.5	0.1	8	0.78	0.56
3	30	0.2	0.1	5	0.68	0.37
3	30	0.5	0.1	5	0.7	0.32
3	30	0.2	0.1	8	0.81	0.44
3	30	0.5	0.1	8	0.84	0.5

We also examine how the comparison between random and heuristic seeding varies with q. For very low values of q the two are quite similar, as concentrating seeds makes little difference and behavior spreads widely. When q is very high even the optimal seeding with a few seeds does not do well, and so either method does similarly. The divergence between the methods appears in the middle ranges of q, where cohesiveness is important for spreading a behavior and concentrating seeds within an atom becomes important. We see this in Figure 24. Figure 24 below plots the average share of the maximum possible spread attained by our heuristic and by random seeding as we vary q for Erdos-Renyi networks with 35 nodes and density 1/10. We see that the middle range q values are also exactly where our heuristic gives the largest improvement over random seeding.

8 Estimating Behavioral Thresholds

To close, we discuss how to infer the behavior threshold q (or t) from observation of a network and agents' decisions to adopt a behavior.

Many applications are such that we not only observe a network, but we also observe



(a) 3 Seeds

(b) 4 Seeds

Figure 24: Comparing the average share of optimal spread attained by our heuristic vs. random seeding as we vary q in Erdos-Renyi networks on 35 nodes with density 1/10.

which agents have chosen some given behavior. For instance, we have seen the adoption of a new technology, or who cheats on their taxes, or who smokes (as in our illustration below), etc.

From these data we can recover the q that is most likely to have generated that behavior (or similarly, t) under the assumptions of the model.

As is well-known (e.g., see Aral et al. (2009); Bramoullé et al. (2009); Goldsmith-Pinkham and Imbens (2013b); Jackson et al. (2017)), homophily can confound behavior, and so working from an empirically observed convention might confound a behavior upon which atoms are defined.

This does not impact any of the definitions up to this point in the paper, since the objective of our atomic analysis was to define atoms and communities based on what behaviors can be supported in settings with strategic complements. Instead, if one is trying to estimate a q from observed behavior, then what is driving behavior matters. In the appendix we discuss methods of distinguishing behavior from homophily. In Subsections 8.1 and 8.2, for the purposes of illustrating potential techniques, we presume that the conventions represent peer-influenced behavior, and defer fuller discussion of including homophily in the appendix. These techniques could augment experiments (e.g., Centola (2011)) or suitable instruments (e.g., Aral and Nicolaides (2017)), to estimate preferences causally.

8.1 Estimating q

First, let us consider a case in which agents all have the same preferences and thus the same threshold q for adopting a behavior.

We observe a network and a convention and wish to estimate a q.

Let $N_{on} = \{i \mid \text{i adopts}\}$ be the set of adopting nodes and N_{off} be its complement.

For each agent i, let s_i be the share of i's neighbors who adopt the behavior. Let $S_{on} = \{s_i | \text{i adopts}\}$ be the distribution of realized on-neighbor shares for the players adopting the behavior, and let S_{off} be the analog for the players not adopting.

There will be perfect separation of the distributions, with q satisfying:

$$\max S_{off} < q \leq \min S_{on}$$

To illustrate, the network depicted below is an Erdos-Renyi network with 200 nodes and the nodes are labeled as pink if they have adopted with a threshold of $q = 2/5$:

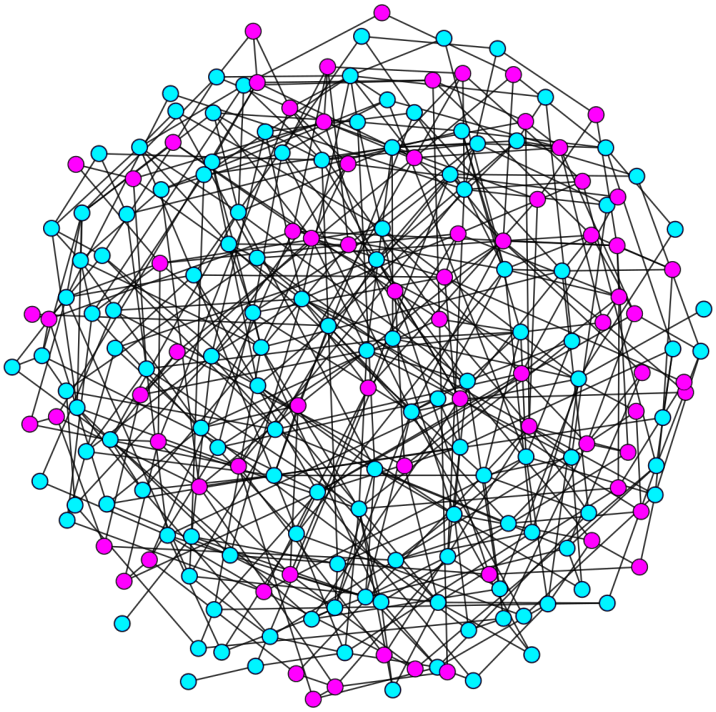


Figure 25: A randomly generated convention on an Erdos-Renyi Network

This equilibrium generates the distributions S_{on} and S_{off} shown below. Observe that $q = 2/5$ perfectly separates the distributions.

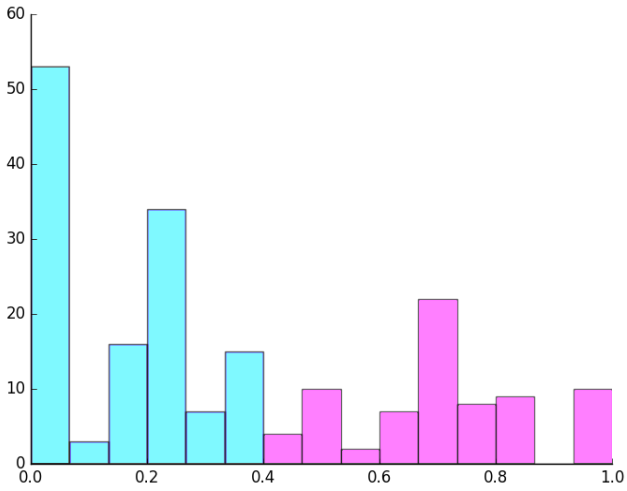


Figure 26: Distributions of on-neighbor shares for nodes adopting (pink) and not adopting(blue)

Of course, in most empirical applications, there is likely to be some heterogeneity in preferences and noisy behavior, so that the observed set of adopters may not form a convention for any precise q.

We can adapt the model to account this in a number of different ways.

The simplest is to introduce a probability $\alpha > 0$ with which each agent chooses his or her adoption decision independently (with equal probability of adopting or not),[8] ignoring other agents. Those agents not making their decisions independently then play the coordination game, with a threshold q, including best responding to their neighbors who made their decisions independently.

The network below shows an outcome of this process on the network from figure 25 with $\alpha = 1/5$ and $q = 2/5$:

[8]One can also introduce a parameter for the probability of adopting independently, or a distribution of such parameters and make that part of the estimation procedure. It could be identified as those who are not matching neighbors are choosing independently. Here we stick with the simple case for illustration.

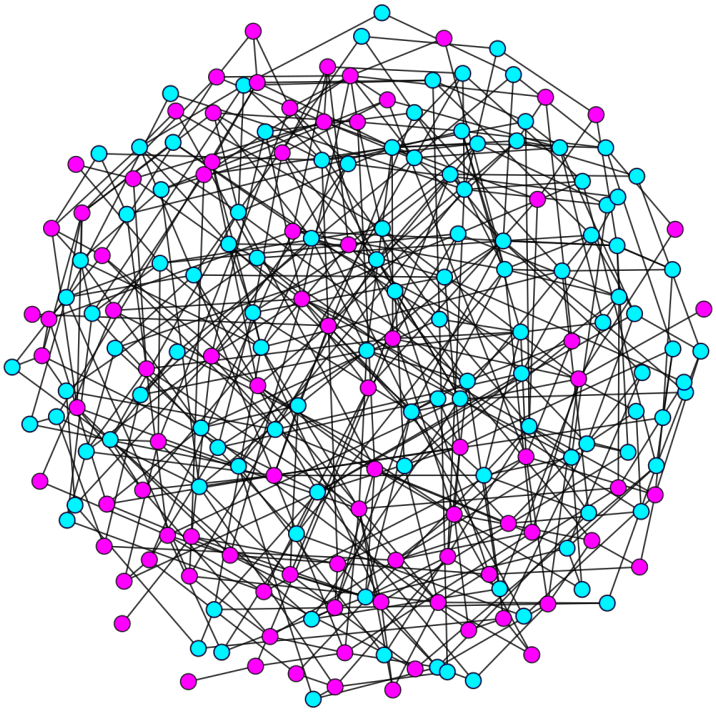


Figure 27: Adoption Decisions when some nodes independently choose their adoption.

This outcome generates the distributions S_{on} and S_{off} shown below. Because of the independent adopters, we cannot perfectly separate the two distributions. However, the true behavioral threshold $q = 2/5$ still approximately separates the distributions, in that most of S_{off} lies below q and most of S_{on} lies above it.

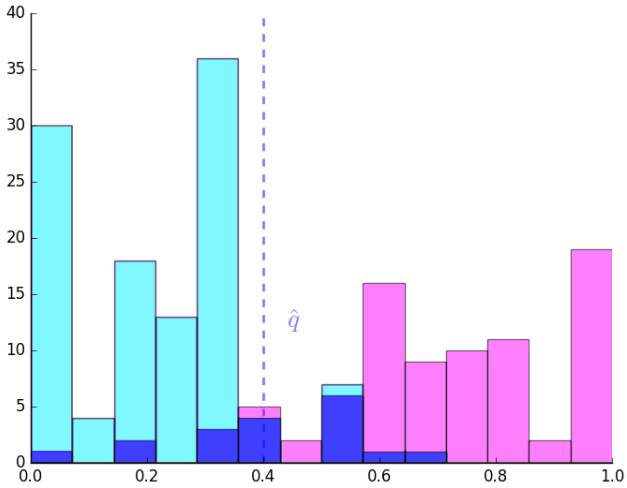


Figure 28: Distributions of 'on'-neighbor shares for nodes adopting in pink and those not adopting in blue.

There are many ways to estimate q for this setting. A very simple and direct way is to minimize the number of nodes that must be acting independently. This effectively corresponds to the maximum likelihood estimator (for α values that are not too high).

In particular, consider a statistic that for each q counts how many nodes must be acting independently (i.e., are not acting in accordance with q):

$$T(q) = |N_{on} \cap \{i : s_i < q\}| + |N_{off} \cap \{I : s_i \geq q\}|.$$

Choosing a $\widehat{q}$ that is a minimizer of $T(q)$.

Though we have only discussed estimation in the relative threshold q case, the approach above extends directly to the absolute threshold (t) case by substituting the number of i's friends taking the action rather than the share (replace s_i with $s_i \times d_i$) and then substituting t for q:

$$T(t) = |N_{on} \cap \{i : s_i \times d_i < t\}| + |N_{off} \cap \{I : s_i \times d_i \geq t\}|.$$

8.1.1 An Illustration of Estimating q

We illustrate the above technique by applying it to a famous example from the community detection literature: Zachary's Karate club.

This club spit into two pieces, with some of its members breaking off to form a new club. The split is pictured in Figure 29.

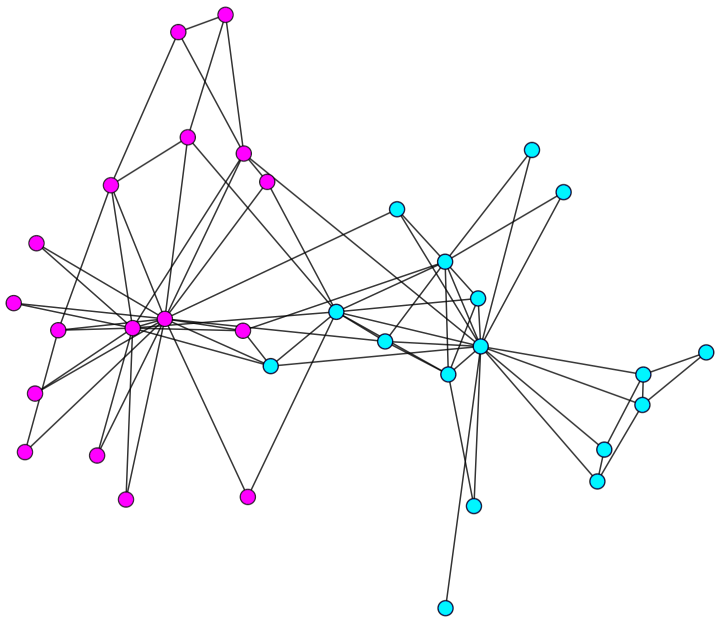


Figure 29: Zachary's Karate Club: the pink nodes are those who split off from the club while the blue nodes stayed.

We can consider the coordination game in which people prefer to split off if and only if at least a share q of their friends do. We can then estimate q using our technique from above.

This is pictured in Figure 30.

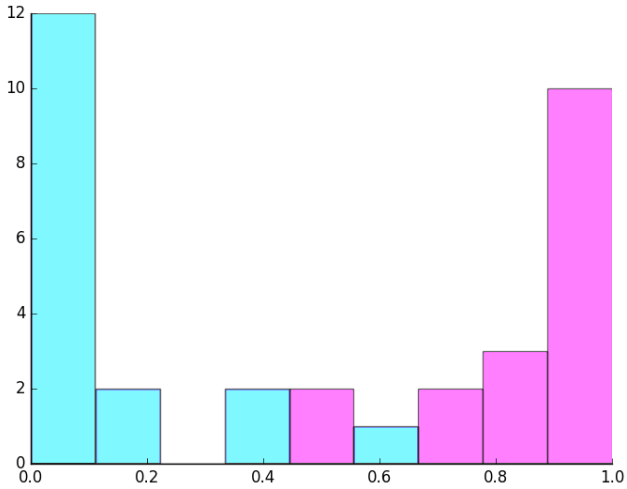


Figure 30: The distribution of neighbors who split off from the club: Those who split off had most of their friends also split, while those who stayed had fewer friends who split off.

To estimate q, we calculate the value of our "mistakes" statistic T at various potential thresholds:

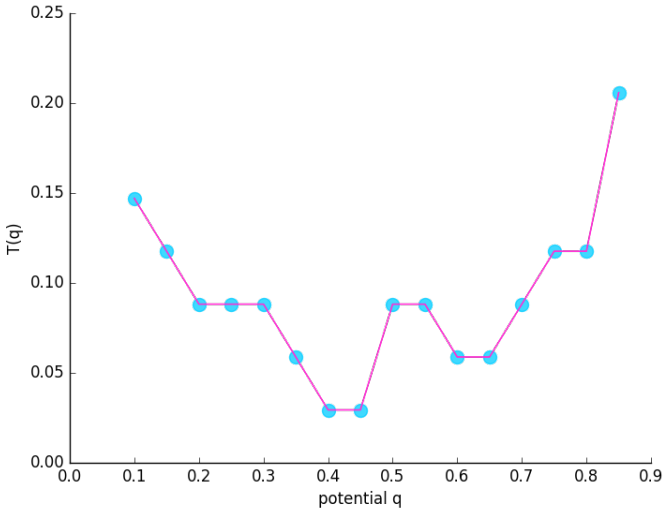


Figure 31: The fraction of nodes that would have to be deviating from coordination if the true threshold were q for various values of q.

We see that T is minimized for q between 0.4 and 0.45. The set of q that minimize T

will always be an interval, since "on"-neighbor shares must vary by discrete multiples of the inverse of the least common product of the nodes' degrees; but the size of the interval will shrink as n grows. This interval serves as an estimate.

8.2 Testing Relative vs Absolute Thresholds

In this subsection, we use the statistic $T(\cdot)$ defined above to evaluate whether a relative, q, or absolute, t, threshold model better fits the observed adoption decisions.[9]

If every person has the same degree, then the two models are interchangeable – every q has an equivalent t. However, in any network in which degree varies across people, then the two models diverge. In the relative model a higher degree person would require more friends taking the action than a lower degree person in order to also want to take the action. In the absolute threshold model, the number of friends needed to induce a person to take the action is independent of the degree. This is the key to identifying which model better fits the data.

To illustrate how the two models lead to different predictions and can be distinguished, we apply the procedure to data on smoking decisions on a high school social network from the AddHealth data.

Figure 32 pictures smoking in a high school social network: pink nodes represent students who said that they had smoked a cigarette in the past twelve months, while blue nodes represents student who had not.

[9]Previous work on testing for different models of peer influence have looked at whether it is information or influence (e.g., Banerjee, Chandrasekhar, Duflo, and Jackson (2013)) or whether it is a simple or complex contagion (e.g., $t = 1$ versus $t > 1$ Centola (2011); Mønsted et al. (2017)). Here we are testing different types of influence against each other.

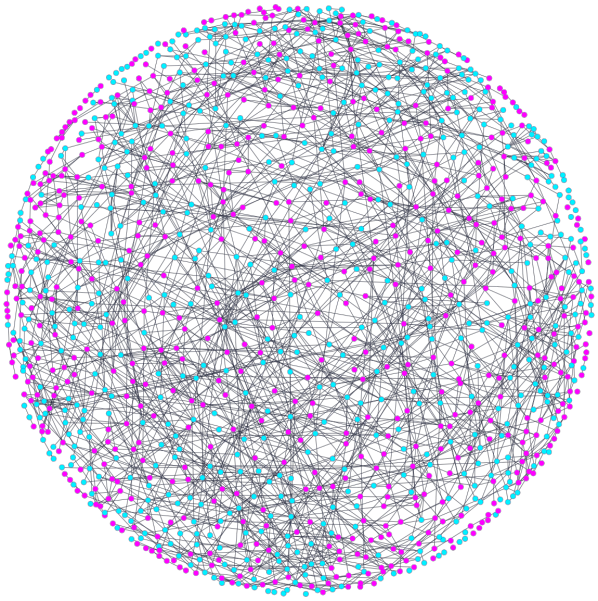
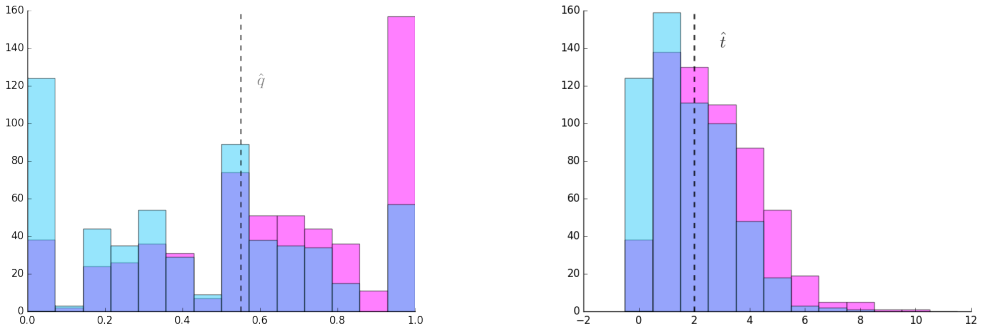


Figure 32: Smoking Adoption in a high school social network

We can compare how well the best-fitting absolute and relative thresholds divide the distributions of on-neighbor shares and numbers, respectively.



(a) Best relative threshold split

(b) Best absolute threshold split

Figure 33: **:** A comparison of how well a relative vs. an absolute threshold splits the on-neighbor distributions

We estimate a relative threshold $\hat{q} = 0.55$, which gives a value of $T(\hat{q})/n = 0.36$; and an absolute threshold of $\hat{t} = 2$, which gives a test-statistic value of $\hat{T}(\hat{t})/n = 0.4$.

In this network there are $n = 1221$ students. The relative threshold $\hat{q} = 0.55$ results in 438 students who are not acting according to the predicted action based on the best fitting threshold, while the absolute threshold $\hat{t} = 2$ has 486 students who are not acting according to the predicted action based on the best fitting threshold - or an extra 48 students, which is

four percent of the population. The relative threshold model thus better predicts behaviors than the absolute threshold.

Distinguishing the two models statistically can be thought of as follows. Letting z be the number of students who are not acting in accordance with the threshold, the chance of matching the observed behaviors is[10]

$$\left(1-\frac{\alpha}{2}\right)^{n-z}\left(\frac{\alpha}{2}\right)^{z}=\left(\frac{2-\alpha}{2}\right)^{n}\left(\frac{\alpha}{2-\alpha}\right)^{z}.$$

This probability decreases exponentially in the number of errors z. Thus, the likelihood is exponentially higher, by a factor $\left(\frac{2-\alpha}{\alpha}\right)^{z-z'}$ when the number of errors is $z' < z$. When $z - z' = 48$, then even for α that is very high, the likelihood is much higher under q than t. Then a log likelihood ratio test leads to a p-value of effectively zero.

If one did not use either model, then the best one could do would be to predict everyone to be a smoker. That would lead to an error rate of 0.49 (the fraction of non-smokers) or 600 students. That is 162 worse than the relative threshold and 114 worse than the absolute model. Again, a log likelihood ratio test leads to a p-value of effectively zero.

The fact that these strict q and t models are still mis-predicting more than a third of the students' actions is partly due to the fact that we are only using the network to predict actions and not including information from demographics. In a data set where one includes demographics, and allows the threshold to depend on demographics (e.g., race, income, gender, etc.), then one could further increase the predictiveness of the model as behavioral thresholds will adjust with demographics in many applications.

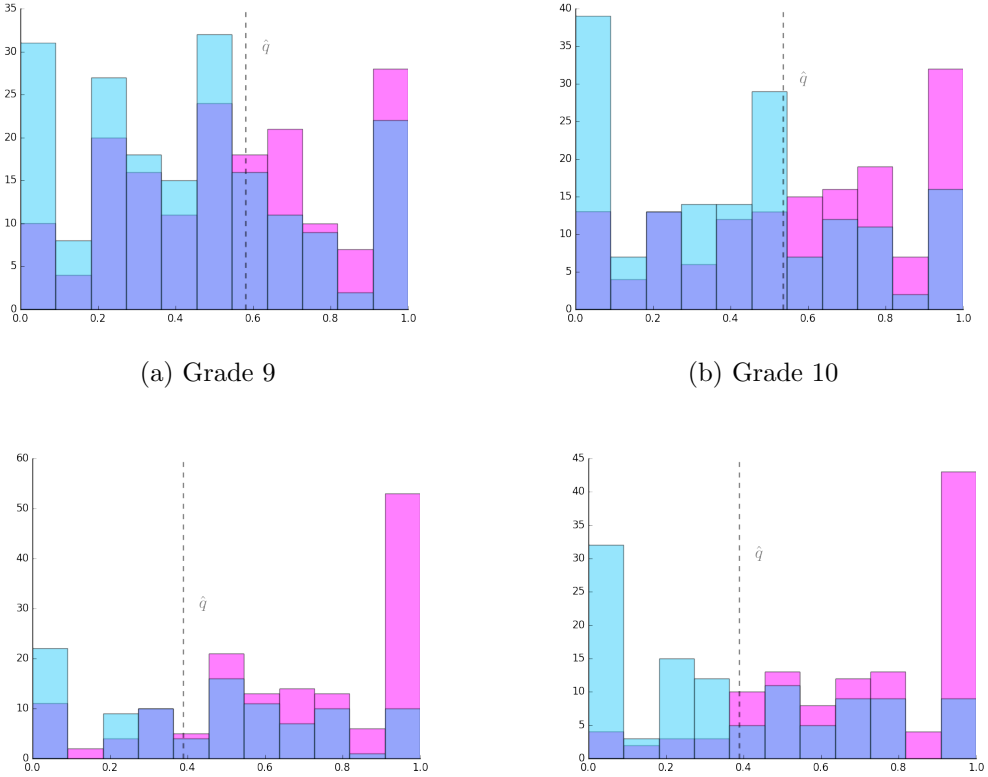
The technique we have outlined extends directly to allow for demographics. If X is the demographic information, then one can fit a function $q(X)$ to predict peoples thresholds and behaviors and then use

$$T(q(\cdot)) = |N_{on} \cap \{i : s_i < q(X_i)\}| + |N_{off} \cap \{I : s_i \geq q(X_i)\}|,$$

as the objective and select $q(\cdot)$ to minimize the function (and similarly for an absolute threshold function $t(\cdot)$).

To illustrate, we estimate $q(X)$ for the high school smoking example, where we take X to be a node's grade in school. Figure 34 below shows the on-node neighbor shares for smokers vs. non-smokers for each grade separately. (Note that we still include students in other grades when calculating the share of a node's neighbors who smoke.)

[10] $1-\alpha+\frac{\alpha}{2}=1-\frac{\alpha}{2}$ is the chance that an agent acting in accordance with the threshold's behavior is correct (the agent acts according to the model, or is random and happens to pick the right action with probability one half), and then $\frac{\alpha}{2}$ is the chance that an agent who is not acting in accordance with the threshold is correctly matched.



(a) Grade 9

(b) Grade 10

(c) Grade 11

(d) Grade 12

Figure 34: **:** Estimating $\hat{q}$ for each of the grades 9,10,11, and 12 separately.

We estimate $\hat{q}$ values of 0.58, 0.54, 0.39, and 0.39 for grade 9, 10, 11, and 12. So in this example, younger students are less easily influenced to smoke by the smoking behaviors of their friends. Using the grade specific thresholds, the share of agents of any gender whose behavior the relative threshold model mispredicts falls marginally, from 0.36 under the universal threshold above to 0.34 under the grade-specific thresholds, which corresponds to mispredicting 24 fewer students – which again leads to an exponential increase in likelihood, but is less dramatic. Still, it has a log-likelihood improvement p value of effectively 0.[11] There may be additional demographics that would lead to additional improvements in fit, the point here is simply to illustrate the potential approach.

9 Concluding Remarks

Although we have introduced the idea of 'behavioral communities', and the associated atomic structure of a network, in the context of settings of coordination on a behavior, it is clear

[11]The test statistic is $2 \times ln(\left(\frac{2-\alpha}{\alpha}\right)^{24})$, which for values of α all the way up to being well above 2/3, leads to a statistic above a χ^2 at a .999 level with four degrees of freedom.

that our approach provides a more general method of identifying communities in networks. For instance, one could extend this to settings in which many actions are possible, or in which behavior takes on a continuum of values. One would need to define when it is that a group would be considered a convention. There are many metrics that could be used, and even definitions that allow for overlapping conventions, which opens many avenues for future research.

Indeed, there is a long list of questions that our approach opens: Are there critical people in a community whose deletion would lead an atom to dissolve? How can a network be recovered from observation of communities? How can algorithms for detecting the atomic structure be improved? What do atoms look like in different types of networks? What are appropriate extensions in multiplexed networks?

References

Akbarpour, M., S. Malladi, and A. Saberi (2017): "Diffusion, Seeding and the Value of Network Information," *mimeo: Stanford University*.

Aral, S., L. Muchnik, and A. Sundararajan (2009): "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proceedings of the National Academy of Sciences*, 106, 21544–21549.

——— (2013): "Engineering social contagions: Optimal network seeding in the presence of homophily," *Network Science*, 1, 125–153.

Aral, S. and C. Nicolaides (2017): "Exercise contagion in a global social network." *Nature communications*, 8, 14753.

Banerjee, A., A. Chandrasekhar, E. Duflo, and M. O. Jackson (2013): "Diffusion of Microfinance," *Science*, 341, DOI: 10.1126/science.1236498, July 26 2013.

Blondel, V. D., J.-L. Guillaume, R. Lambiotte, and E. Lefebvre (2008): "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, 2008, P10008.

Blum, A., J. Hopcroft, and R. Kannan (2016): *Foundations of data science*, Mimeo: Carnegie Mellon University.

Bollobas, B. (2001): *Random Graphs*, Cambridge University Press.

Borgatti, S. P. and M. G. Everett (1992): "Notions of position in social network analysis," *Sociological methodology*, 1–35.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of peer effects through social networks," *Journal of Econometrics*, 150, 41–55.

Centola, D. (2011): "An experimental study of homophily in the adoption of health behavior," *Science*, 334(6060), 1269–1272.

Centola, D., V. M. Eguíluz, and M. W. Macy (2007): "Cascade dynamics of complex propagation," *Physica A: Statistical Mechanics and its Applications*, 374, 449–456.

Chen, W., Z. Liu, X. Sun, and Y. Wang (2010): "A game-theoretic framework to identify overlapping communities in social networks," *Data Mining and Knowledge Discovery*, 21, 224–240.

Clifford, P. and A. Sudbury (1973): "A model for spatial conflict," *Biometrika*, 60, 581–588.

Copic, J., M. O. Jackson, and A. Kirman (2009): "Identifying community structures from network data via maximum likelihood methods," *The BE Journal of Theoretical Economics*, 9.

Currarini, S., M. O. Jackson, and P. Pin (2009): "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica*, 77, 1003–1045.

——— (2010): "Identifying the roles of race-based choice and chance in high school friendship network formation," *Proceedings of the National Academy of Sciences*, 107, 4857–4861.

Currarini, S. and F. Mengel (2012): "Identity, Homophily and In-Group Bias," *mimeo.*

Fortunato, S. (2010): "Community detection in graphs," *Physics Reports*, 486:3-5, 75–174.

Goldsmith-Pinkham, P. and G. Imbens (2013a): "Rejoinder: Social Networks and the Identification of Peer Effects," *Journal of Business and Economic Statistics*, 31:3, 279–281.

——— (2013b): "Social Networks and the Identification of Peer Effects," *Journal of Business and Economic Statistics*, 31:3, 253–264.

Golub, B. and M. O. Jackson (2012): "How Homophily Affects the Speed of Learning and Best-Response Dynamics," *Quarterly Journal of Economics*, 127, 1287–1338.

Granovetter, M. (1978): "Threshold models of collective behavior," *American journal of sociology*, 83, 1420–1443.

Holland, P. W., K. B. Laskey, and S. Leinhardt (1983): "Stochastic blockmodels: First steps," *Social networks*, 5, 109–137.

Holley, R. A. and T. M. Liggett (1975): "Ergodic theorems for weakly interacting infinite systems and the voter model," *The annals of probability*, 643–663.

Jackson, M. O. (2008): *Social and economic networks*, Princeton: Princeton University Press.

——— (2013): "Comment: Unraveling Peers and Peer Effects (Comments on Goldsmith-Pinkham and Imbens' "Social Networks and the Identification of Peer Effects")," *Journal of Business and Economic Statistics*, 31:3, 270–273, DOI: 10.1080/07350015.2013.794095.

——— (2018, forthcoming): *The Human Network*, Pantheon Books: New York.

Jackson, M. O., B. W. Rogers, and Y. Zenou (2017): "The Economic Consequences of Social Network Structure," *Journal of Economic Literature*, 55:1, 49–95.

Jackson, M. O. and Y. Zenou (2014): "Games on Networks," *Handbook of Game Theory, Elsevier, edited by Young, H.P. and Zamir, S.*

Kempe, D., J. Kleinberg, and E. Tardos (2003): "Maximizing the Spread of Influence through a Social Network," *Proc. 9th Intl. Conf. on Knowledge Discovery and Data Mining*, 137–146.

——— (2005): "Influential Nodes in a Diffusion Model for Social Networks," *In Proc. 32nd Intl. Colloq. on Automata, Languages and Programming*, 1127–1138.

Kets, W. and A. Sandroni (2016): "A belief-based theory of homophily," *mimeo.*

Lorrain, F. and H. C. White (1971): "Structural equivalence of individuals in social networks," *The Journal of mathematical sociology*, 1, 49–80.

McPherson, M., L. Smith-Lovin, and J. Cook (2001): "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, 27, 415–444.

Mønsted, B., P. Sapieżyński, E. Ferrara, and S. Lehmann (2017): "Evidence of Complex Contagion of Information in Social Media: An Experiment Using Twitter Bots," *arXiv preprint arXiv:1703.06027.*

Morris, S. (2000): "Contagion," *Review of Economic Studies*, 67 (1), 57–78.

Sacerdote, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly journal of economics*, 116, 681–704.

White, D. R. and K. P. Reitz (1983): "Graph and semigroup homomorphisms on networks of relations," *Social Networks*, 5, 193–234.

Appendix

Proof of Theorem 1:

Let us begin by showing that with a probability going to 1, all nodes have at least $\overline{q}$ of their neighbors in their own block, and fewer than $\underline{q}$ of their neighbors outside of their block.

Consider some node i of type j. Let $d_i(j', n)$ be i's degree with nodes of type j'.

From Chernoff bounds it follows that

$$\Pr\left((1-\delta)d_{jj'}(n) < d_i(j',n) < (1+\delta)d_{jj'}(n)\right) > 1 - 2e^{-\delta^2 d_{jj'}(n)/3}.$$

Thus, looking across all j':

$$\Pr\left(\forall j' : [(1-\delta)d_{jj'}(n) < d_i(j',n) < (1+\delta)d_{jj'}(n)]\right) > (1 - 2e^{-\delta^2 \min_{j'} d_{jj'}(n)/3})^{|J|}.$$

It follows that

$$\Pr\left(\forall j' : [(1-\delta)d_{jj'}(n) < d_i(j',n) < (1+\delta)d_{jj'}(n)]\right) > 1 - 2|J|e^{-\delta^2 \min_{j'} d_{jj'}(n)/3}.$$

Therefore,

$$\Pr\left(\exists j' : [d_i(j',n) < (1-\delta)d_{jj'}(n) \text{ or } d_i(j',n) > (1+\delta)d_{jj'}(n)]\right) < 2|J|e^{-\delta^2 \min_{j'} d_{jj'}(n)/3}.$$

Thus the expected number of nodes for which their degree to some block lies outside of a δ bound around the expected degree that they should have with that block satisfies (letting j_i denote i's type):

$$E\left(\#\{i : \exists j', [d_i(j',n) < (1-\delta)d_{j_i j'}(n) \text{ or } d_i(j',n) > (1+\delta)d_{j_i j'}(n)]\}\right) < n2|J|e^{-\delta^2 \min_{jj'} d_{jj'}(n)/3}$$

Examining the right hand side of the inequality, it follows that

$$n2|J|e^{-\delta^2 \min_{jj'} d_{jj'}(n)/3} < n2|J|e^{-\delta^2 f(n)\log(n)/3} = n2|J|n^{-\delta^2 f(n)/3} = 2|J|n^{1-\delta^2 f(n)/3} \to 0$$

Thus, the probability that all nodes have all their degrees with all blocks within a factor of $(1 \pm \delta)$ times the expected degree goes to 1, for any δ.

Pick δ, such that

$$(1+\delta)\max_{j \neq j'} \frac{d_{jj'}(n)}{d_j(n)} < \min(Q) \quad \text{and} \quad (1-\delta)\min_{j} \frac{d_{jj}(n)}{d_j(n)} > \max(Q).$$

Consider the case in which $Q \subset (\underline{q}, \overline{q})$ [the other case holds by an analogous argument]. Thus, with a probability going to 1, all nodes have more than $\max(Q)$ of their links within their own block and less than $\min(Q)$ in any other block. It follows that with a probability going to 1, each block forms a Q-robust convention. Thus, $C(Q, g^n)$ is finer than $\Pi(n)$ with a probability going to 1. ∎

Proof of Corollary 1: Consider a growing sequence of stochastic block networks satisfying the conditions of Theorem 1 and for which $3\overline{q} - \varepsilon > 1 + \underline{q} + \varepsilon + \frac{j(n)}{d_j(n)}$ for some $\varepsilon > 0$.

By Theorem 1 we know that the partitions corresponding to $C([\underline{q} - \varepsilon, \overline{q} + \varepsilon], g^n)$ are finer than $\Pi(n)$ with a probability going to 1 as n grows. We show that with a probability going to 1 all conventions are coarser than the blocks. We simply need to show that with a probability going to 1 there is no convention that involves part of a block but not the whole block.

As in Theorem 1, the probability that all nodes have all their degrees with all blocks within a factor of $(1 \pm \delta)$ times the expected degree goes to 1, for any δ. So, again, we argue at the limit, but given the arbitrarily small $\delta > 0$, and the $\varepsilon > 0$, the argument extends.

We show that any convention that intersects some block most contain that block.

Consider some convention S and node $i \in S$ of type j. It has at most $(1 - \overline{q})d_j(n)$ of its neighbors outside of its own block, and so in order to have $(\overline{q} - \varepsilon)d_j(n)$ connections in the convention, it has to have $(2\overline{q} - 1 - \varepsilon)d_j(n)$ neighbors within the convention in its own block. This leaves $j(n) - (2\overline{q} - 1 - \varepsilon)d_j(n)$ nodes in its block that could be outside of the convention. These nodes outside of S but in j then have at least

$$\overline{q}d_j(n) - [j(n) - (2\overline{q} - 1 - \varepsilon)d_j(n)] = (3\overline{q} - 1 - \varepsilon)d_j(n) - j(n)$$

neighbors within S. It follows from the condition $3\overline{q} - \varepsilon > 1 + \underline{q} + \varepsilon + \frac{j(n)}{d_j(n)}$ that

$$(3\overline{q} - 1 - \varepsilon)d_j(n) - j(n) > (\underline{q} + \varepsilon)d_j(n)$$

and so the convention is not closed to those other nodes in the block, and hence must include all other nodes in the block. ∎

Proof of Theorem 2: We can decompose the network into an ER random graph with link probability $p(n) > (1 + \varepsilon)\left(\frac{t\log(n)}{n}\right)^{1/t}$, plus extra links. The result is then a corollary of Theorem 3. ∎

Proof of Theorem 3:

First we prove that if $p(n) \geq (1 + \varepsilon)\left(\frac{k\log(n)}{n}\right)^{1/k}$ then there is no k-closed set.

We prove this at the threshold of $(1+\varepsilon)\left(\frac{k\log(n)}{n}\right)^{1/k}$ and thus it also holds for any larger p since *not* having a k-closed set is a monotone property (e.g., see Bollobas (2001): if a graph has the property then adding more links maintains the property). Take $\varepsilon > 0$ to be small. Since k is fixed and ε is arbitrary, we equivalently work with $\left(\frac{(1+\varepsilon)k\log(n)}{n}\right)^{1/k}$.

First, note that the probability that some node has fewer than k connections to some set of k nodes is

$$1 - p(n)^k.$$

The probability of having some set B of cardinality $n_B \geq k$ be k-closed is at most

$$\left[\left(1 - p(n)^k\right)^{n-n_B}\right]^{\lfloor \frac{n_b}{k} \rfloor}. \tag{1}$$

This is because a necessary condition for B to be closed is having all of its sets of k nodes closed to nodes outside of B, and there are at least $\lfloor \frac{n_b}{k} \rfloor$ *disjoint* sets of k nodes in B and for which each of them being closed to nodes outside of B is an independent event So this only counts those disjoint sets being closed to nodes outside of B, and hence is a loose upper bound on the probability of B being closed.[12]

Next, note that the probability of having some set B of cardinality $n_B \geq k$ be 1-closed when the formation of a link happens with probability $p'(n)$ is

$$\left((1-p'(n))^{n_B}\right)^{n-n_B}.$$

We rewrite this as

$$\left[\left((1-p'(n))^{k}\right)^{n-n_B}\right]^{\frac{n_b}{k}}. \tag{2}$$

We show that when $p(n) = \left(\frac{(1+\varepsilon)k\log(n)}{n}\right)^{1/k}$ and $p'(n) = (1+\varepsilon)\frac{\log(n)}{n}$ then the expression in (1) is less than the expression in (2).

It is enough to show that

$$\left(1-p(n)^k\right) < \left[(1-p'(n))^k\right]^{\frac{n_b}{k}/\lfloor\frac{n_b}{k}\rfloor}. \tag{3}$$

Noting that $\frac{n_b}{k}/\lfloor\frac{n_b}{k}\rfloor \leq 2$, we verify that

$$1-p(n)^k < (1-p'(n))^{2k}.$$

At the designated values of $p(n)$ and $p'(n)$ this becomes

$$1-(1+\varepsilon)\frac{k\log(n)}{n} < \left(1-\frac{(1+\varepsilon)\log(n)}{n}\right)^2,$$

for which it is enough that

$$\frac{k(1+\varepsilon)\log(n)}{n} > 2\frac{(1+\varepsilon)\log(n)}{n}.$$

This establishes that the probability that any set of at least k nodes is k-closed under this $p(n)$ is less than the corresponding probability that the same set is 1 closed under this $p'(n)$.

This implies that the expected number of sets of at least k nodes that are k-closed under this $p(n)$ is less than the expected number of sets of nodes least k nodes that are 1-closed under this $p'(n)$.

[12] The possibility of overlap in the sized-k subsets induces correlation between the events that the subsets are closed to nodes outside of B and produces a much more complicated expression for the exact probability, but this loose upper bound is easier to calculate and suffices for our proof.

The second number is known to converge to 0 (see Bollobas (2001)[13]) Thus, the expected number of sets of at least k nodes that are k-closed under this $p(n)$ converges to 0. This implies that the probability of having any such sets converges to 0, as claimed.

To complete the proof, we show that if $p(n) = \left(\frac{(1-\varepsilon)k\log(n)}{n}\right)^{1/k}$ then the probability of having at least one k-closed set of size k goes to one.

The probability that any particular set of k nodes is closed is

$$(1-p^k)^{n-k}.$$

Thus, the expected number of closed sets of size k is

$$\sum_{B\subset N:\#B=k}(1-p^k)^{n-k} = \binom{n}{k}(1-p^k)^{n-k}.$$

At the threshold this is

$$\binom{n}{k}\left(1-\frac{(1-\varepsilon)k\log(n)}{n}\right)^{n-k}.$$

For fixed k, this is of the order of

$$n^k n^{-(1-\varepsilon)k} = n^{k\varepsilon} \to \infty.$$

Thus, the expected number of k-closed sets of k nodes goes to infinity.

To complete the proof, we show that the variance of the number of k-closed sets of k nodes compared to the mean is bounded. Once that is established, the fact that the expected number of k-closed sets of k nodes goes to infinity implies that the probability that there exists at least one k-closed set of k nodes exists goes to one can be proven using Chebychev's inequality by bounding the variance compared to the mean (See the proof on page 95 of Jackson (2008)).

Therefore, letting $X_{n,k}$ be the number of k closed sets of k nodes, we show that the variance of $X_{n,k}$ is a bounded multiple of $E[X_{n,k}]$.

First, note that we can write

$$Var(X_{n,k}) = E[X_{n,k}(X_{n,k}-1)] + E[X_{n,k}] - E[X_{n,k}]^2.$$

Next, note that $E[X_{n,k}(X_{n,k}-1)]$ is the expected number of ordered pairs of k closed sets of k nodes, and that

$$E[X_{n,k}(X_{n,k}-1)] \leq \binom{n}{k}(1-p^k)^{n-k}\sum_{k'=1}^{k}\binom{n-k}{k'}(1-p^{k'})^{n-k-k'}$$

[13]Usual statements of the threshold of connectivity of $G(n,p)$ are in terms of the probability of connectedness. Here we are using the stronger statement that the expected number of components is going to 0 above the threshold. This can be pieced together from results in Bollobas (2001), but one can also find direct treatments, for instance, see Section 4.5.2 in Blum, Hopcroft, and Kannan (2016).

Here k' is the number of nodes of the second set of nodes that does not overlap with the first, and $(1-p^{k'})^{n-k-k'}$ ignores any possible overlapping links between the two sets of k nodes, and so is a loose upper bound on the probability that a second set is closed conditional upon the first one being closed. Thus,

$$Var(X_{n,k}) \leq E[X_{n,k}] + \binom{n}{k}(1-p^k)^{n-k}\left(\sum_{k'=1}^{k}\binom{n-k}{k'}(1-p^{k'})^{n-k-k'}\right) - \binom{n}{k}^2(1-p^k)^{2n-2k}$$

and so

$$Var(X_{n,k}) \leq E[X_{n,k}] + E[X_{n,k}]\left(\sum_{k'=1}^{k}\binom{n-k}{k'}(1-p^{k'})^{n-k-k'} - \binom{n}{k}(1-p^k)^{n-k}\right).$$

This implies that[14]

$$Var(X_{n,k}) \leq E[X_{n,k}] + E[X_{n,k}]\left(\sum_{k'=1}^{k} E[X_{n-k,k'}] - E[X_{n,k}]\right).$$

Along the lines of our proof above, $\sum_{k'=1}^{k} E[X_{n-k,k'}] - E[X_{n,k}]$ is of the order of

$$\sum_{k'=1}^{k}(n-k)^{k'\varepsilon} - n^{k\varepsilon},$$

and

$$\sum_{k'=1}^{k}(n-k)^{k'\varepsilon} - n^{k\varepsilon} \leq \frac{(n-k)^{k\varepsilon}}{1-\frac{1}{(n-k)^{\varepsilon}}} - n^{k\varepsilon} = (n-k)^{k\varepsilon}\left(\frac{(n-k)^{k\varepsilon}}{(n-k)^{\varepsilon}-1} - \frac{n^{k\varepsilon}}{(n-k)^{\varepsilon}}\right) < 0,$$

which completes the proof. ∎

Distinguishing Homophily from Coordination

When estimating a threshold from an observed network and convention, there is an alternative hypothesis that could also potentially explain the data: homophily. This is a well-known confound of correlated behaviors (e.g., see Aral, Muchnik, and Sundararajan (2009); Jackson, Rogers, and Zenou (2017)). People with similar backgrounds and tastes tend to be linked to each other. Thus, there could be strong correlation patterns in neighbors' behaviors simply because they have the same tastes without any influence from one to another.

[14] Note that $\binom{n-k}{k'}(1-p^{k'})^{n-k-k'} < E[X_{n-k,k'}]$ since it is $p = \left(\frac{(1-\varepsilon)k\log(n)}{n}\right)^{1/k}$ rather than $\left(\frac{(1-\varepsilon)k'\log(n-k)}{n-k}\right)^{1/k'}$ that is used in the expression above, and that for large n, small ε, and $k' < k$, $\left(\frac{(1-\varepsilon)k'\log(n-k)}{n-k}\right)^{1/k'} < \left(\frac{(1-\varepsilon)k\log(n)}{n}\right)^{1/k}$.

With controlled experiments (e.g., random assignment of roommates as in Sacerdote (2001)) or instruments under some conditions (Bramoullé, Djebbari, and Fortin (2009), Aral and Nicolaides (2017)) one can test for influence directly as we discussed in the paper.

In the absence of such identification, faced with just observations of networks and behavior, it is more challenging to distinguish homophily from influence.

To fix terms, we use 'pure homophily' to describe a world in which the probability that two agents are linked to each other depends on some distance function of a vector of their characteristics and their adoption decisions are a stochastic function of that vector, but in which the adoption decision does not depend on the realized adoption decisions' of one's neighbors. In contrast, by 'pure coordination', we mean a model in which each agent in a network (one possibly formed under the influence of homophily) makes his adoption decision as a function only of whether or not the share of his neighbors adopting the behavior is above some common threshold q.[15]

Admitting homophily in our null hypothesis yields a more demanding test of coordination, since homophily will generate positive coordination between neighbors' adoption decisions just as coordination does. Testing against homophily thus requires a more subtle analysis of the structure that correlation takes.

It is impossible to entirely rule arbitrary models of homophily. For instance, suppose we observe a convention S that matches perfectly with some threshold q. Another model that could explain the data is that people who adopted are those who prefer to adopt, and the other people preferred not to adopt. The associated 'pure homophily' model of network formation is that all people who prefer to adopt also prefer to form links in such a way as to end up with a network in which a fraction of at least q (respectively, more than t) of their friends have the same preference; and conversely, all those who prefer not to adopt want to be in a network in which less than a fraction q (respectively, fewer than t) of their friends prefer to adopt. Beyond that, agents have no preferences over network structure.

This leads to the following straightforward observation:

Claim **1** *Given any convention S associated with some network g and adoption threshold q (or t), there is a pure homophily model of network formation that generates the same data.*

The claim means that ironclad testing for causation of behavior will require controlled experiments or valid instruments. Thus, making inferences from observed conventions requires caution, especially with respect to policy.

Of course, the 'pure homophily' model underlying the extreme version of the claim above is pathological.

If we restrict what are thought of as plausible models of homophily, then one can distinguish coordination behavior from homophily from an observed convention. The logic is as follows.

[15]Of course, there could be both homophily and coordination, but to fix ideas we discuss testing for one exclusively versus the other exclusively.

Consider the following simple, but fairly general model of pure homophily. Nodes have some real-valued unobserved characteristics denoted by θ_i and the probability that two nodes i, j are linked is random but strictly decreasing in the distance $|\theta_i - \theta_j|$, holding the rest of the network fixed.[16] The probability of adoption is an increasing function of θ_i.

This model of homophily is empirically distinguishable from pure coordination. Under this model of pure homophily, nodes' adoption decisions are correlated with their neighbors' adoption decisions *conditional on a share of at least q of those neighbors adopting*, for any choice of q, whereas with coordination this correlation disappears if we condition on q. The behavior of linked individuals is correlated under pure homophily, *and* that correlation remains after conditioning on fractions of neighbors adopting as types are correlated via network formation and seeing one of a pair's behavior is still informative even conditional upon others' behaviors.

We outline a formal test of pure coordination versus pure homophily with our simple version of homophily. Given a network G and a set of "on" nodes S, let $s_i = |N_i(G) \cap S|/N_i(G)$.Define the statistic R as a function of a threshold q as:

$$R(q) = \text{Corr}(s_i, 1_{i \in S} \mid s_i > q) + \text{Corr}(s_i, 1_{i \in S} \mid s_i < q)$$

In words, the statistic R is the sum of the correlations between adoption and on-neighbor shares conditional on nodes' on-neighbor shares being above and below q. The idea of the formal test is to examine the joint statistic (T, R). If there is no coordination, and also no homophily in characteristics correlated with adoption, then for all q $T(q)$ should be large while $R(q)$ should be small. If there is no coordination, but there is homophily in characteristics correlated with adoption, then for all q $\hat{R}(q)$ should be large. Only pure coordination should make it possible to find a q such that $T(q)$ and $R(q)$ are both small. This suggests the following procedure:

First estimate q by minimizing $H(q) = F(T(q), R(q))$ for some function H which is increasing in both T and R. Second, assuming some functional form for the linking probabilities (as a decreasing function of $|\theta_i - \theta_j|$) and adoption probabilities (as an increasing function of θ_i), use maximum likelihood estimation to estimate the parameters of these functions along with the values θ_i. Third, compare the minimized value $H(\hat{q})$ to the distribution of H generated by the maximum likelihood estimates (this may have to be bootstrapped depending on the functional form assumptions) to obtain a p-value.

One could also fit a hybrid model in which one simultaneously estimates the latent space of θ_i's and the q's as a function of that space. The idea would be that instead of simply choosing a best fit $q(X_i)$ function as we discussed above, one could fit a function of the form $q(X_i, \theta_i)$ together with the assignment of θ_i's based on the network and observed behaviors. This would be a variaiton on the approach of Goldsmith-Pinkham and Imbens (2013b) (see also Jackson (2013); Goldsmith-Pinkham and Imbens (2013a)) adapted to our setting.

[16] One can extend this to higher dimensional spaces of unobserved characteristics in the obvious manner, with a partial order on monotonicity of actions in types.

Seed Sets and Fragile Conventions

In this section we provide some alternative definitions of conventions and behavioral communities.

Conventions may have subsets that still form a convention.

So, let us define various sorts of minimal conventions.

9.1 Self-Sustaining Seed Sets

First, let us define a *q-self-sustaining seed set* to be a subset $S \subset N$ that is q-cohesive and has no nonempty strict subset that is q-cohesive.

So, a self-sustaining seed set is a minimal set of nodes such that if they all adopt the behavior then their best responses will continue to be to adopt the behavior. They may also spread the behavior beyond their set, but that is not necessary as part of the definition.

The reason to have the "self-sustaining" part, is that it might be possible to seed a convention with a smaller set if one can pay or force nodes to adopt a behavior and not change back. Self-sustaining means that none of the nodes would change back if they are all turned on, and so that requires that the group be q-cohesive.

Note that q-tight sets and self-sustaining seed sets can be distinct: there are clearly self-sustaining seed sets that are not tight since closure is not required of a self-sustaining seed set, and there are q-tight sets that are not self-sustaining seed sets since there can be a subset that is cohesive (just not closed). For example, a triad with $q < 1/2$ is tight, but any two nodes form a self-sustaining set.

Nonetheless, every q-tight set contains at least one self-sustaining seed set.

9.2 Fragile Conventions

Let us call a convention *fragile* if there exists some node in the convention, such that changing that node to not adopting the behavior and then iterating on best responses would lead *all* agents to change to not adopting the behavior - so changing just one node's behavior can completely eradicate the convention.

There are fragile conventions that are not tight, and vice versa.

For instance, the two conventions on the right side of Figure 3 are fragile but not tight. There is a single node in each of those conventions that when changed leads the convention to collapse, but there is still a subset of each convention that forms a convention (those on the left of the figure) and so the conventions are not tight.

To see a tight convention that is not fragile, again consider a triangle with $q < 1/2$. Changing any node's behavior will not alter the other two, and so it is not fragile.

A self-sustaining seed set that is a convention is necessarily fragile. All subsets of a self-sustaining seed set are not cohesive and so changing *any* node's behavior will unravel the convention.

Moreover, every fragile convention contains a self-sustaining seed set.

Any convention that contains at least two disjoint tight subsets is not fragile.

There are conventions that contain overlapping but distinct tight subsets that are fragile.

CLAIM **2** *Consider the set of nodes that can unravel a fragile convention: all of them need to be part of any self-sustaining seed set that generates the convention.*

proof: Any node that is taken out and can unravel a fragile convention is such that there is no cohesive subset left that excludes that node, so any cohesive set must include all such nodes.

9.3 Fragile and Seeding Community Structures

Let $C^F(Q, g)$ denote the finest σ-algebra whose elements are all coarser than all the fragile conventions that are robust relative to Q.

Here, instead of grouping two nodes together if they act the same way in every convention, they are grouped together if they act the same way in at least one fragile convention. Thus, their behavior is sometimes tied together by the network structure, but does not have to always be tied together.

The fragility of the convention is important as otherwise taking the union of two completely disjoint conventions would put people together without their behaviors really being tied down by each other.

The finest elements of $C^F(Q, g)$ are referred to as *fragile atoms*.

One can also define a version which also requires that incorporates a self-seeding requirement.

For instance, let $C^{FS}(Q, g)$ denote the finest σ-algebra whose elements are all coarser than all the fragile conventions that are robust relative to Q and which are q-self-sustaining seed sets for all $q \in Q$.