A Behavioral Micro-foundation for Cross-sectional Network Models[*]

Carter T. Butts[†]

Abstract

Models for cross-sectional network data have become increasingly well-developed in recent decades, and are widely used. Questions remain, however, about the connection between such cross-sectional models and the behavioral processes from which the corresponding networks were presumably generated. Here, we present a behavioral micro-foundation for such networks, based on a continuous time stochastic choice mechanism. As we show, the equilibrium behavior of this process under appropriate conditions can be expressed in exponential family form, allowing estimation of individual preferences using existing methods. We demonstrate the applicability of the approach via an analysis of advice-seeking behavior among managers within a high-tech firm, providing evidence of preferences for seeking advice from "upstream" sources net of covariate and other effects.

Keywords: social networks, stochastic choice, exponential random graphs, Markov chains

1 Background and Introduction

Models for cross-sectional network data have become increasingly well-developed in recent decades, particularly those based on the use of discrete exponential family modeling. At the same time, however, such models have been criticized for being difficult to interpret, and (more fundamentally) for lacking a principled connection with the behavioral processes from which the corresponding networks were presumably generated. Although explicitly dynamic models provide one response to this problem (see, e.g., Snijders (1996); Robins and Pattison (2001); Snijders (2005)), these require longitudinal data which is frequently impossible or impractical to obtain. The question thus arises of whether there exists a principled means of supplying a behavioral micro-foundation for at least some classes of cross-sectional network models. Here, we present such a foundation, based on a continuous time stochastic choice mechanism. As we show, the equilibrium behavior of this process under appropriate conditions can be expressed

1

in exponential family form, allowing estimation of individual preferences using existing methods. After deriving these equilibrium results, we demonstrate the applicability of the approach via an analysis of advice-seeking behavior among managers within a high-tech firm, providing evidence of preferences for seeking advice from "upstream" sources net of covariate and other effects.

1.1 Formal Preliminaries

The networks which concern us here can be represented as simple undirected and directed graphs of fixed order. We represent both by ordered pairs $G = (V, E)$, where V is the set of *vertices* and E is a set of *edges* on V. Throughout, we shall use $n = |V|$ to refer to the order of the graph being described (where $|\cdot|$ denotes cardinality). For simple graphs representing undirected relations, the elements of E are two-element subsets of V; in the directed case, the elements of E are ordered pairs of vertices. Frequently, we will represent graphs via their adjacency matrices. The *adjacency matrix* of a graph, G, is the $n \times n$ binary matrix, y, such that $y_{ij} = 1$ if $\{i, j\} \in E$ in the undirected case (or $(i, j) \in E$ in the directed case) and $y_{ij} = 0$ otherwise. Given a set of multiple graphs $G_1 = (V, E_1), \ldots, G_m = (V, E_m)$ on common vertex set V, we can easily extend this notation by defining the corresponding *adjacency array*, y, as the $m \times n \times n$ binary array such that $y_{ijk} = 1$ if $\{j, k\} \in E_i$ ($(j, k) \in E_i$ in the directed case) and $y_{ijk} = 0$ otherwise. The adjacency matrix for the ith graph in such an array clearly corresponds to $y_{i\cdot\cdot}$, and is referred to as the ith *slice* of y.

Adjacency matrices and arrays are particularly useful for describing *random graphs*, i.e., random variables whose state space consists of a graph set. We will here designate random graphs via their adjacency matrices, using capital letters to denote random variables and lower case letters to denote realizations. Thus, a random graph G may have (random) adjacency matrix Y, within which the presence of an i, j edge is denoted by the binary random variable (or *edge variable*) Y_{ij}. A corresponding realization of G would then be represented by the adjacency matrix y, with the state of the i, j edge indexed by y_{ij}. This notation extends directly to adjacency arrays. In some cases, we will also want to refer to edge variables within a random graph *excluding* a particular edge variable, or with the states of particular variables held constant. To this end, we define Y_{ij}^c to be the set of all edge variables of Y *other than* $\{i, j\}$ (directed case, (i, j)); likewise, y_{ij}^c refers to all elements of realization y other than the i, jth. A related situation arises when we wish to describe a graph structure in which a particular edge is forced to be either present or absent: Y_{ij}^+ and Y_{ij}^- are respectively employed to denote random matrices with $Y_{ij}^+ = 1$ and $Y_{ij}^- = 0$ and $Y_{kl}^+ = Y_{kl}^- = Y_{kl}$ for $\{i, j\} \neq \{k, l\}$ (directed case, $(i, j) \neq (k, l)$). Realization matrices y_{ij}^+ and y_{ij}^- are defined in like manner.

Finally, we note that when describing stochastic processes on graphs (i.e., random graph processes), we will employ sequences of random adjacency matrices or arrays indexed by parenthetical superscripts (e.g., $Y^{(1)}, Y^{(2)}, \ldots$). The same principle will be applied to realizations of such processes; hence, a real-

ization of the above process might be denoted $y^{(1)}, y^{(2)}, \ldots$. For processes on graph sets, we simply apply the same notation to the corresponding adjacency arrays.

2 Cross-sectional Network Models

Cross-sectional network models have been the subject of considerable research over the past several decades (see, e.g., Holland and Leinhardt (1981); Frank and Strauss (1986); Wasserman and Pattison (1996); Pattison and Robins (2002); Wasserman and Robins (2005); Snijders et al. (2006); Hunter and Handcock (2006) and related references), one result of which has been the emergence of a general family of formalisms for representing probability distributions on graphs. Specifically, let Y be the adjacency matrix of order-n random graph G, and let $\mathcal{Y}_n$ be the support of Y. We may then write the pmf of Y in exponential family form as

$$\Pr\left(Y = y \,|\, \rho\right) = \frac{\exp\left(\rho\left(y\right)\right)}{\sum_{y' \in \mathcal{Y}_n} \exp\left(\rho\left(y'\right)\right)} I_{\mathcal{Y}_n}(y), \tag{1}$$

where $\rho : \mathcal{Y}_n \mapsto \mathbb{R}$ is a *graph potential*, and $I_{\mathcal{Y}_n}$ is an indicator function for membership in $\mathcal{Y}_n$. Typically, the graph potential is expressed in canonical form as $\rho\left(y\right) = \theta^T t\left(y\right)$, with $t : \mathcal{Y}_n \mapsto \mathbb{R}^p$ being a vector of *sufficient statistics*, and $\theta \in \mathbb{R}^p$ a vector of *parameters*. Although sometimes (incorrectly) described as a model in its own right, the framework of Equation 1 is more properly regarded as a way of representing distributions on $\mathcal{Y}_n$; in particular, observe that by choosing t to correspond to a $|\mathcal{Y}_n|$-dimensional vector of indicators for the state of Y, any pmf on a graph set of fixed order may be written in this fashion. Notwithstanding the potential for confusion, it is both traditional and useful to refer to network models parameterized in exponential family form as *exponential random graph* (or ERG) models, and we follow this practice here.

Although most work with ERGs has focused on the single-network case, there is nothing in the above to force such a restriction. Indeed, simply taking Y and $\mathcal{Y}_n$ to refer to m-slice adjacency arrays allows Equation 1 to be applied to graph sets without additional difficulties. General issues relating to parameterization of such "multivariate" exponential random graph (or MERG) models have been discussed by Wasserman and Pattison (1996); Pattison and Wasserman (1999) and Robins et al. (1999), although their use dates back to the original work by Holland and Leinhardt (1981).

The attractiveness of ERGs as a general framework for cross-sectional network modeling stems from several sources. As already noted, the ERG formalism is *complete* on the set of finite order graphs and graph sets[1], in the sense that all models for such entities can be written (albeit not always parsimoniously) in ERG form. As special cases of the more general discrete exponential families (Brown, 1986; Barndorff-Nielsen, 1978), considerable inferential theory

---

[1] In fact, it is also complete on graphs of countable order, though such models are not guaranteed to have a finite parameterization.

exists for ERG models (see, e.g., Hunter and Handcock (2006); **?** for recent reviews); thus, fairly general methods of parameter estimation, hypothesis testing, model selection, and adequacy evaluation exist for models parameterized in this form. Similarly, general algorithms for simulation of ERG models exist (e.g., Snijders (2002); Hunter et al. (2008)) based on Markov chain Monte Carlo (MCMC) methods. Finally, the ERG form is well-suited to parameterization based on *dependence hypotheses*, a notion from spatial statistics (Besag, 1974) which can be adapted to produce families of network models satisfying particular assumptions regarding the interdependence of edges (see, e.g., Frank and Strauss (1986); Pattison and Robins (2002)).

As these observations suggest, the ERG framework acts as a sort of "lingua franca" for cross-sectional network modeling: models translated into ERG form can immediately exploit a broad family of inferential and simulation tools, and can moreover be compared and even combined on a principled basis. While the creation of this framework represents a tremendous advance, it does nothing to address the *content* of network models – this last must necessarily come from a combination of empirical investigation, prior substantive knowledge regarding the structures under study, and an enumeration of modeling objectives. With the exception of models based on the above-mentioned dependence hypotheses and certain obvious covariate effects, however, linking substantive theory with ERG models has proven difficult. The major portion of the text that follows is concerned with this latter problem. Before turning to this, however, we pause to consider a few other important prefatory matters.

2.1 Potentials, Conditional Probabilities, and ERGs

The ERG (in the sense of Equation 1) specifies the joint probability of an entire graph (or set thereof) as a single observation. Direct calculation of such probabilities are rarely feasible, due to the normalizing factor $Z\left(\rho, \mathcal{Y}_n\right) = \sum_{y' \in \mathcal{Y}_n} \exp\left(\rho\left(y'\right)\right)$. Z corresponds directly to the *partition function* of statistical mechanics, and (through its derivatives) implicitly contains a wealth of information regarding the behavior of the corresponding model. (See Strauss (1986); **?** for related discussion.) Regrettably, Z has closed-form expressions for only a handful of models, and is generally incomputable due to the large number of elements in the corresponding sum (and the roughness of the underlying function). The various computational strategies for coping with this issue do not concern us here. However, the problem of normalization does motivate the general interest in conditional probabilities and probability ratios vis a vis models in ERG form. For instance, let $y, y' \in \mathcal{Y}_n$ be elements of the support of Y. Then, from Equation 1,

$$\frac{\Pr\left(Y = y' \mid \rho\right)}{\Pr\left(Y = y \mid \rho\right)} = \frac{\exp\left(\rho\left(y'\right)\right)}{\sum_{y'' \in \mathcal{Y}_n} \exp\left(\rho\left(y''\right)\right)} \frac{\sum_{y'' \in \mathcal{Y}_n} \exp\left(\rho\left(y''\right)\right)}{\exp\left(\rho\left(y\right)\right)} \tag{2}$$

$$= \exp\left(\rho\left(y\right) - \rho\left(y\right)\right). \tag{3}$$

4

Thus, the log-odds of y' versus y reduces to their difference in potentials, with Z falling out of the equation. This observation leads immediately to an expression for the conditional probability of a single edge, given the rest of the graph:

$$\frac{\Pr\left(Y = y_{ij}^+ \left| y_{ij}^c, \rho\right.\right)}{\Pr\left(Y = y_{ij}^- \left| y_{ij}^c, \rho\right.\right)} = \exp\left(\rho\left(y_{ij}^+\right) - \rho\left(y_{ij}^-\right)\right) \tag{4}$$

$$\frac{\Pr\left(Y = y_{ij}^+ \left| y_{ij}^c, \rho\right.\right)}{1 - \Pr\left(Y = y_{ij}^+ \left| y_{ij}^c, \rho\right.\right)} = \exp\left(\rho\left(y_{ij}^+\right) - \rho\left(y_{ij}^-\right)\right) \tag{5}$$

$$\Pr\left(Y = y_{ij}^+ \left| y_{ij}^c, \rho\right.\right) = \frac{\exp\left(\rho\left(y_{ij}^+\right) - \rho\left(y_{ij}^-\right)\right)}{1 + \exp\left(\rho\left(y_{ij}^+\right) - \rho\left(y_{ij}^-\right)\right)} \tag{6}$$

$$= \frac{1}{1 + \exp\left(\rho\left(y_{ij}^-\right) - \rho\left(y_{ij}^+\right)\right)} \tag{7}$$

$$= \text{logit}^{-1}\left(\rho\left(y_{ij}^+\right) - \rho\left(y_{ij}^-\right)\right). \tag{8}$$

Here again, we are left with a simple expression in terms of potential differences – in this case, the conditional probability of an edge is seen to be the inverse logit of the change in potential associated with that edge (holding the rest of the graph constant). This is, of course, the basis for the well-known conditional logistic interpretation of the ERG model, and for the standard ERG pseudo-likelihood estimator (Frank and Strauss, 1986; Strauss and Ikeda, 1990). Our use for the expression is rather different: as we shall see, it forms an important element in our linkage of behavioral choice models with network structure. Before coming to this, however, we round out this section of the paper by reviewing the connection between cross-sectional ERGs and a basic family of generative processes.

2.2 Generative Processes

Although we have thus far discussed ERGs entirely within a cross-sectional framework, they are also associated with various generative processes. Most generically, it is obviously the case that a temporal aggregation or cross-section of any fixed-order network process can described in the form of Equation 1; this is a trivial consequence of the already-noted fact that all distributions on $\mathcal{Y}_n$ can be written in this form. Such an observation is not especially helpful in linking ERGs with generative processes, however, since it provides no indication of how the appropriate potential, ρ, might be found. Moreover, graph potentials for aggregations or cross-sections of dynamic processes will not in general be invariant to sampling time, length of aggregation window, or even the initial conditions of the underlying process. Thus, one cannot hope to find a completely generic way to map generative processes onto ERGs (much less the reverse).

Despite this caveat, there are certain families of generative processes which do give rise to cross-sectional distributions in a predictable manner. Typically, these processes have been studied in the context of statistical simulation of ERG draws, e.g. in support of inference or model assessment (e.g. Crouch

et al. (1998); Snijders (2002); Hunter et al. (2008)), with a focus on properties relating directly to such goals (e.g. simulation accuracy and efficiency). Our interest is somewhat different: we will ultimately use the generative process to link micro-level behavior with cross-sectional network structure. Nevertheless, it happens that we can exploit this pre-existing work in furthering our theoretical objectives.

Our focus here is on a particularly simple process, namely the single-update Gibbs sampler (Snijders, 2002). Let us begin by positing some support for a process with cross-section Y, $\mathcal{Y}_n$, from which some initial graph $Y^{(0)}$ is drawn from an ERG with arbitrary potential ρ_0. Let $E^*(\mathcal{Y}_n)$ denote the set of edge variables of Y. We now form the sequence $Y^{(1)}, Y^{(2)}, \ldots$ via the following iterative procedure:

1. At the ith iteration, draw an edge variable Y_{jk} from a variable selection process S_i on $E^*(\mathcal{Y}_n)$;

2. With probability $Pr\left(\left(Y_{jk}^+\right)^{(i-1)} \middle| \left(Y_{jk}^c\right)^{(i-1)}, \rho\right)$, let $\left(Y^{(i)} = Y_{jk}^+\right)^{(i-1)}$. Otherwise, let $Y^{(i)} = \left(Y_{jk}^-\right)^{(i-1)}$.

The following is then a standard result (e.g., Gilks et al. (1996)):

Proposition 1. *Let ρ be a finite potential on $\mathcal{Y}_n$, and let S_i be a process on $E^*(\mathcal{Y}_n)$ such that*

1. *S_i is independent of $Y^{(1)}, Y^{(2)}, \ldots, Y^{(i)}$; and*

2. *$\sum_i I(S_i = Y_{jk}) \rightarrow \infty$ as $i \rightarrow \infty$ almost surely, for all j, k such that $Y_{jk} \in E^*(\mathcal{Y}_n)$.*

Then the sequence $Y^{(0)}, Y^{(2)}, \ldots, Y^{(i)}$ approaches $ERG(\rho)$ in distribution as $i \rightarrow \infty$.

As this is a special case of the well-known random scan Gibbs sampler, we do not provide a proof of this proposition here. Interested parties are referred to the above review chapter and included references for details.

Proposition 1 holds because the iterative procedure forms a Markov chain on the elements of $\mathcal{Y}_n$ whose equilibrium distribution is $ERG(\rho)$. Irreducibility and aperiodicity of the chain stems from the fact that the conditional probability of setting an edge to be present or absent is bounded away from 0 or 1 (itself a consequence of the finiteness of ρ combined with Equation 8), along with the properties of S; that the resulting equilibrium distribution is the correct one is less obvious, but is a known result for chains of this type.

For our purposes, several points bear emphasis. First, the above procedure provides an example of a generative process based on local structural adjustments, which ultimately gives rise to a known cross-sectional distribution. Second, the local adjustments on which the process is based are Bernoulli trials,

6

whose probabilities are inversely logistic in the potential difference between successive graphs. Third, the selection of the initial graph state is wholly arbitrary, in the sense that it does not impact the equilibrium distribution of the process. Finally, the asymptotic cross-sectional properties of the graph process do not depend upon the details of the process by which edge variables are selected for updating – so long as the specified conditions are met, edge variables may be updated in any order (and, indeed, irregularly) without affecting the long-run behavior of the chain. Taken together, these properties suggest an interesting possibility: if it were possible to map a plausible behavioral process to the sampling algorithm, then one could in principle use its equilibrium distribution to predict the long-run cross-sectional behavior of the corresponding social system. It is precisely this notion that we explore in the section which follows.

3 Cross-sectional Distributions from Stochastic Choice

In the previous section of this paper we focused on the definition of cross-sectional network models, closing with the description of a simple family of stochastic processes whose equilibria correspond to specified ERG distributions. In this section, we build on these results by demonstrating a family of behaviorally reasonable micro-processes that belong to the above class, and specify sufficient conditions for relating their resulting equilibria to individual preferences. We begin by describing a simple stochastic choice process for social networks, subsequently showing how this process can lead to a specified equilibrium distribution in the case of unilateral edge control (i.e., relations for which any given edge is controlled by a single party). Next, we generalize this to the more complex case of bilateral edge control (i.e., relations for which any given edge is jointly controlled by two parties), closing the section with a framework for the general multilateral case.

3.1 Agent Choice Framework

The core of our behavioral micro-process is a stochastic choice model, through which individuals are presumed to make decisions which impact (but may or may not entirely determine) a set of relationships. We begin by assuming a *manifest relation*, Y, on some fixed vertex set V. The edges of Y are presumed to arise from the actions of one or more *agents,* who are collectively denoted by the set A. It is important to note that the set of agents need not be the same as (nor even have a non-empty intersection with) the set of vertices, although $V = A$ is a useful special case. The connection between agents and the edges they affect is formalized via a set of *control lists*, which specify the agents whose behaviors govern any given edge. Specifically, for each edge variable Y_{ij}, we define the corresponding control list $c_{ij} \subseteq A$ to be the minimum lexically ordered ℓ-tuple of agents whose behaviors determine the state of Y_{ij}. (The assumption

7

of lexical ordering is not technically necessary, but considerably simplifies subsequent notation.) The control list length, ℓ, is referred to as the *control number* for Y, and is of considerable substantive importance. Networks for which $\ell = 1$ are said to exhibit *unilateral edge control*, as the state of each edge is determined by the behavior of a single agent. Similarly, networks with $\ell = 2$ are said to exhibit *bilateral edge control*, with $\ell > 2$ reflecting the more general multilateral case. Particular issues arise in each of these settings, and we therefore treat them in greater detail below. In passing, we note that one can in principle posit cases of *mixed control*, in which ℓ varies across edge variables. Although we here limit ourselves to the "pure" cases in which ℓ is constant, our framework generalizes to mixed control in a fairly natural way.

Given an agent set and associated control structure, we must next posit a framework for representing agent choices regarding his or her associated edges. Here, we propose to do this via a set of ℓ latent graphs, each of which encodes the current decision of a set of agents regarding the edges of Y. Specifically, we define the *prosphoric array* for Y to be an $\ell \times N \times N$ adjacency array, P, whose i, j, k cell contains the current decision state of the ith agent in c_{jk} regarding Y_{jk}. The manner in which the cells $P_{.jk}$ determine Y_{jk} – and hence the behavioral interpretation of the associated "decision" – is relation-specific, and will be discussed in detail below. Intuitively, however, it is often useful to think of the prosphoric array as representing "offers" or "overtures" for the creation or maintenance of ties, whose appropriate coincidence (depending on both ℓ and on the properties of the relation) may lead to the presence of the associated manifest relation. While this analogy is the basis for the terminology employed here ("prosphoric" from the Greek for "offer" or "overture"), it should be stressed that we do not presume that such sub-relational interactions are actually taking place; rather, we consider the prosphoric array as a formal device for aggregating agent choices, which may or may not have any direct relationship to the physical or social details of the interactive process through which the ties of Y are actually created. In particular, it should be noted that the slice structure of P is determined by the lexical ordering of A, and is thus semi-arbitrary. That said, the joint structure of P *is* meaningful in terms of its relationship to the manifest relation, Y, and it is to this that we now turn.

3.1.1 Unilateral Edge Control

To express the relationship between individual decisions and the resultant graph structure, we posit a *resolution function*, r, which takes the set of possible prosphoric arrays into $\mathcal{Y}_n$ (the set of possible manifest networks). In the simplest case of unilateral edge control (i.e., $\ell = 1$), r is simply the identity function: that is, $Y_{jk} = P_{1jk}$. As this implies, a manifest edge under unilateral control exists if and only if its controlling agent chooses it, and hence the meaning of a decision within the prosphoric array is simply that of whether or not to maintain the corresponding relationship.

Unilateral edge control is an obvious model for directed relations such as advice seeking, in which each individual has direct control over his or her outgoing

ties. In this case (which we will revisit empirically in a subsequent section) it is reasonable to let $V = A$, and to set $c_{ij} = (j)$ for all (i, j) pairs. While such examples are easy to come by, they are not the only phenomena which can be modeled in this way. For instance, consider an organizational design problem, in which managers within different divisions of an organization have discretion to define the formal reporting relationships for their own divisions (with the CEO defining reporting relationships among the division heads). In this case, $A \subset V$, and c_{ij} for a given relationship would be the identity of the controlling manager. One can easily conceive of related applications on undirected relations (e.g., design of communication channels rather than reporting relations), relations for which the set of agents is not contained within the set of vertices (e.g., design of physical systems, such as road networks), etc. What all of these cases have in common is that edges are contingent upon the behavior of a single decision maker, and that the corresponding decision is simply one of whether or not to maintain the edge in question.

3.1.2 Bilateral Edge Control

Although many relations are unilateral, there are likewise many examples of relations which depend upon the actions of two parties. In such cases, $\ell = 2$, and more than one substantively distinct mapping from P to Y is possible. Of the possible variants, two forms are of clear sociological interest and are discussed in detail here. First, we have the case in which a manifest edge exists if and only if both controlling parties agree to it; we refer to these as *symphonic relations* (from "symfonia," or agreement), and define the corresponding resolution function as $r(P)_{jk} = P_{1jk}P_{2jk}$. The second case is that in which a manifest edge exists if and only if *some* controlling party agrees to it. We refer to such relations as *epibolic* (from "epiboli," or imposition), expressing the intuition that either party can "impose" the relationship upon the other. The resolution function for an epibolic relation is $r(P)_{jk} = 1 - (1 - P_{1jk})(1 - P_{2jk})$, i.e., a logical OR on the elements of $P_{.jk}$. Although one can identify other cases as well (e.g., a logical XOR), symphonic and epibolic relations constitute the bulk of those currently studied within the literature. Given this, we take a moment to consider each in turn before proceeding to the case of general multilateral relations.

Symphonic Relations   As with unilateral relations, bilateral symphonic relations are obvious models for a wide array of interpersonal networks. For instance, many notions of friendship (e.g., the *philos* of Krackhardt (1992)) are defined in such a way as to require mutual assent, and such is of practical necessity for many forms of collaboration (e.g., in task performance following disasters). The typical scenario for such relations is an undirected graph, Y, with $A = V$ and $c_{jk} = (j, k)$ (where $j < k$ in lexical order). Although the above gives the impression that symphonic relations occur only in cases of positive interaction, this is not the case. Consider, for instance, the interaction between an infrastructure designer or maintainer (e.g., a civil authority) and an adversary

(e.g., a hostile military force) in the context of a transportation network. For an edge (e.g., a roadway) to be passable, it must both be built/maintained by the infrastructure designer *and* allowed to operate by the adversary. Obviously, the designer and adversary have very different goals, and their interaction is strongly negative. Nevertheless, such a situation can be treated as a symphonic relation with A consisting only of the designer and adversary, and c_{jk} being the ordered elements of A for all j, k.

As with the unilateral case, there is no necessary correspondence between the resolution mechanism and the directedness of the edges of Y, the positivity/negativity of the relation, nor the elements of A. Modeling Y as symphonic does, however, fix the meaning of the decision $P_{ijk} = 1$ to be "the ith controlling party of the j, k edge chooses to implement the edge, *if* the other party concurs." So long as one bears in mind that "concurrence" may be tacit (and need not connote a positive-valence interaction among the controlling parties), this provides a clear sense of the decision in question.

Epibolic Relations If friendship is an obvious example of a symphonic relation, interpersonal contact is a similarly common epibolic example. Contact is inherently symmetric, and may be instigated by either party to the relation; it is thus natural to model it by letting $A = V$, and $c_{jk} = (j, k)$. Interestingly, it may be appropriate to model the same underlying relationship as epibolic or unilateral, depending on what is assumed to drive the behaviors of the participants. For instance, an advising relation generally involves a two-way exchange of information, even if one party is the initiator of the interaction. If this exchange *per se* were conjectured to drive agent behavior (and not the properties of the unilateral advice-seeking edge), then it would be more appropriate to treat the interaction as undirected and epibolic rather than directed and unilateral. (Use of model comparison techniques to empirically disambiguate such cases is in principle possible, although this will not be treated here.) While contact and advice are directly interpersonal, one can naturally find epibolic relations in other contexts. Returning to our earlier transportation example, we could for instance consider a scenario in which two authorities (e.g., governments of neighboring city-states) have the capacity to build road segments within the same area (e.g., neutrally held territory adjacent to both). Since a road segment exists if implemented by either party (in the absence of malfeasance), the road network as a whole is an epibolic relation. The agents in this case are the two government authorities, and the control list for all edges is simply these two agents in lexical order.

The meaning of the decision $P_{ijk} = 1$ in the epibolic case is straightforwardly implied by the above examples: "the ith controlling party of the j, k edge chooses to implement the edge, *regardless* of whether the other party concurs." This is obviously close to the unilateral case, save in that neither party can prevent the edge from being established by the other. As before, it should be stressed that this is independent of the valance of Y, and could reflect either tacit or explicit sub-relational interaction.

3.1.3 General Multilateral Edge Control

Although most studied relations fall into the unilateral or bilateral cases, there is in principle no limit to the number of agents who may be involved in edge creation. For instance, **?** discuss games in which a committee of designers attempts to negotiate the structure of an organization; although their case is concerned with dynamics (i.e., plans for changing an organization over time), similar interactions may occur in negotiations over structure per se. Models for hypergraphs (which can be represented two-mode bipartite graphs, with appropriate restrictions on $\mathcal{Y}_n$) may also suggest a more general multilateral control structure.

In such cases, the number of possible resolution mechanisms is obviously quite large, and we do not attempt to enumerate them here. Symphonic and epibolic relations generalize naturally to the general context, with respective R functions $r(P)_{jk} = \prod_{i=1}^{\ell} P_{ijk}$ and $r(P)_{jk} = 1 - \prod_{i=1}^{\ell} (1 - P_{ijk})$. Other potential rules include thresholding mechanisms (i.e., some critical number k of controlling parties must choose an edge to activate it), which takes the "democratic" process of majority rule as a special case. As before, choice of r is a substantive modeling decision which indirectly determines the behavioral meaning of the prosphoric array, and as such should be guided by prior knowledge regarding the social system in question.

3.2 Edge Updating Events, and the Decision Model

Having constructed a framework for representing agent decisions regarding edges within a network, and having discussed the mechanisms by which those decisions are transformed into manifest ties, we are now in a position to discuss the choice process itself. Our approach is to model agents as boundedly rational decision makers, who myopically adjust their relationships so as to (on average) increase their utilities for the resulting manifest network. Adjustments are assumed to occur episodically in continuous time, such that adjustment opportunities (though not realized adjustments) are independent of the current network state. As we will subsequently show, this process (for certain families of utility functions) leads to cross-sectional behavior which can be directly parameterized in terms of individual utility functions.

3.2.1 Edge Updating Process

As noted above, we assume that relationships within Y are persistent (as opposed to instantaneous; see **?** for a discussion), with changes occurring in continuous time. Changes occur when provoked by agents' decisions (within P), as determined by the resolution function, r. As cognitively bounded individuals, we presume that agents are not capable of evaluating all controlled edges simultaneously, nor at all times. Rather, we presume that an unobserved saliency-governing process episodically leads each agent to evaluate his or her decision regarding a particular edge, at which time he or she revisits the decision within

the current structural context (as described below). Should this evaluation lead to a state change, P (and, if necessary, Y) are immediately updated, and all decisions are then fixed until the next opportunity arises.

We here refer to such "opportunities" as *edge updating events*, and the process which generates them as an *edge updating process*. Formally, we define this process, $X^{(1)}, X^{(2)}, \ldots$, as a sequence of random (j, k, l, t) tuples, where $a(X^{(i)}) = j$ is the updating agent, $e_s(X^{(i)}) = k$ is the sender of the hypothetical edge being updated, $e_r(X^{(i)}) = l$ is the receiver of the hypothetical edge, and $\tau(X^{(i)}) = t$ is the time at which the updating event occurs. Like P, we consider X to be fully latent, and we specify only that it satisfy the following two conditions. First, we require X to be independent of P (and hence of Y) – change opportunities (but not realized changes) do not depend on the state of the network. Second, we require that all elements of P are updated at least occasionally, in the limited sense that $\sum_{x:\tau(x)<t} I\left(a\left(x\right) = i, e_s(x) = j, e_r(x) = k\right) \to \infty$ as $t \to \infty$ a.s. for all $\{j, k\}$ (directed case (j, k)) in $E^*(\mathcal{Y}_n)$ and all $i \in c_{jk}$. Thus, in the limit, all edges will be updated infinitely many times, though this process need not happen in any particular order. While very weak, we will see that these conditions are sufficient to yield the desired result.

3.2.2 Stochastic Choice Mechanism

When an edge updating event $X^{(i)} = (j, k, l, t)$ occurs, the evaluating agent $a = c_{klj}$ revisits his or her decision regarding the state of the j, k edge (i.e., the value of P_{jkl}). Let $P^{(i-1)}$ be the current state of the prosphoric array; a must then choose between $P^{(i)} = \left(P^{(i-1)}\right)^+_{jkl}$ and $P^{(i)} = \left(P^{(i-1)}\right)^-_{jkl}$. We here assume that a's preferences can be expressed by a finite utility function on the states of Y, i.e., $u_a : \mathcal{Y}_n \mapsto \mathbb{R}$. Given this, a myopically chooses the state of P_{jkl} based on its immediate relational consequences. Within such a setting, a frequently used behavioral choice model is the logistic choice model (McFadden, 1973), so-called because of its functional form. Specifically, the logistic choice model predicts that the probability of choosing a given option (versus a single alternative) is equal to the inverse logit of the difference in utilities between the alternatives. When one option is decisively superior, the model converges towards the rational actor solution (strict utility maximization), alternately converging towards equal probability of selection at the indifference point; between these two extremes, the decision maker is predicted to favor the superior option, with the degree of favoring increasing with the difference in utilities. Applying this framework to

12

a prosphoric choice in the above case gives us:

$$
\Pr\left(P_{jkl}^{(i)} = \left(p^{(i-1)}\right)_{jkl}^{+} \,\middle|\, \left(P^{(i-1)}\right)_{jkl}^{c} = \left(p^{(i-1)}\right)_{jkl}^{c}, u_a \right)
$$

$$
= \frac{\exp\left[u_a\left(r\left(\left(p^{(i-1)}\right)_{jkl}^{+}\right)\right)\right]}{\exp\left[u_a\left(r\left(\left(p^{(i-1)}\right)_{jkl}^{+}\right)\right)\right] + \exp\left[u_a\left(r\left(\left(p^{(i-1)}\right)_{jkl}^{-}\right)\right)\right]} \tag{9}
$$

$$
= \mathrm{logit}^{-1}\left[u_a\left(r\left(\left(p^{(i-1)}\right)_{jkl}^{+}\right)\right) - u_a\left(r\left(\left(p^{(i-1)}\right)_{jkl}^{-}\right)\right)\right]. \tag{10}
$$

Thus, a's choice is seen to be governed by the difference in his or her utilities for the states of Y resulting from his or her decision, as desired.

One important special case of the above occurs when there exists some potential function, ρ, such that $\rho\left(y_{kl}^{+}\right) - \rho\left(y_{kl}^{-}\right) = u_a\left(y_{kl}^{+}\right) - u_a\left(y_{kl}^{-}\right)$ for all $a \in c_{kl}$ and all $\{k,l\}$ (or (k,l) in the directed case). In such a setting, Equation 10 reduces to

$$
\Pr\left(P_{jkl}^{(i)} = \left(p^{(i-1)}\right)_{jkl}^{+} \,\middle|\, \left(P^{(i-1)}\right)_{jkl}^{c} = \left(p^{(i-1)}\right)_{jkl}^{c}, \rho \right) =
$$

$$
\mathrm{logit}^{-1}\left[\rho\left(r\left(\left(p^{(i-1)}\right)_{jkl}^{+}\right)\right) - \rho\left(r\left(\left(p^{(i-1)}\right)_{jkl}^{-}\right)\right)\right], \tag{11}
$$

and all decisions depend only on the difference in potentials (which is not agent-specific). Considered in game theoretic terms, the existence of such a ρ identifies the corresponding behavioral model as a *potential game* (Monderer and Shapley, 1996), and we likewise say that ρ is a potential for the tuple $(A, \ell, c, \mathcal{Y}_n)$.

3.3   Equilibrium Distribution of Relations Under the Behavioral Model

Given the above, we are now ready to state our primary result:

Theorem 1. *Let Y be the adjacency structure arising from the behavioral model specified by $(\mathcal{Y}_n, A, \ell, c, r, u)$ under edge updating process X, and let $Y^{[t]}$ be the state of Y at time t. If ρ is a potential for $(A, \ell, c, \mathcal{Y}_n)$, and X is such that*

1. *X is independent of P; and*

2. *$\sum_{x:\tau(x)<t} I\left(a\left(x\right) = i, e_s(x) = j, e_r(x) = k\right) \to \infty$ as $t \to \infty$ a.s. for all $\{j,k\}$ (directed case (j,k)) in $E^*(\mathcal{Y}_n)$ and all $i \in c_{jk}$,*

then $Y^{[t]}$ converges in distribution to $\Pr\left(Y^{[t]} = y\right) = |\{p : r(p) = y\}|\, \frac{\exp[\rho(y)]}{\sum_{p' \in \mathcal{P}_n} \exp[r(p')]}$ on support $\mathcal{Y}_n$ as $t \to \infty$.

Proof. We begin by noting that the independence of X and P implies that both P and Y are embedded Markov chains; without loss of generality, then, we may restrict our attention to the sequences $Y^{(1)}, \ldots, Y^{(i)}, \ldots$ and $P^{(1)}, \ldots, P^{(i)}, \ldots$

13

arising from the set of realized edge updating events $x^{(1)}, \ldots, x^{(i)}, \ldots$. Since ρ is a potential for $(A, \ell, c, \mathcal{Y}_n)$, we have from Equation 11 that $\Pr\left(P_{jkl}^{(i)} = \left(p^{(i-1)}\right)_{jkl}^+ \left| \left(P^{(i-1)}\right)_{jkl}^c = \left(p^{(i-1)}\right)_{jkl}^c, \rho \right.\right) = \text{logit}^{-1}\left[\rho\left(r\left(\left(p^{(i-1)}\right)_{jkl}^+\right)\right) - \rho\left(r\left(\left(p^{(i-1)}\right)_{jkl}^-\right)\right)\right]$ (where $j = a(x_i)$, $k = e_s(x_i)$, and $l = e_r(x_i)$). From Equation 8, however, this is exactly equal to the conditional probability that $P_{jkl} = pk_{jkl}$ for a multivariate ERG process on P with potential $\rho(r(P))$. Now, consider the sequential edge updating process formed by the realized updating events. By assumption, $x^{(i)}$ is independent of P. Further, $\sum_{x:\tau(x)<t} I\left(a\left(x\right) = i, e_s(x) = j, e_r(x) = k\right) \to \infty$ as $t \to \infty$ a.s. clearly implies that $\sum_{i=1}^{\infty} I\left(a\left(x\right) = i, e_s(x) = j, e_r(x) = k\right) \to \infty$ a.s. as well. This implies that the sequence $P^{(0)}, P^{(1)}, \ldots$ satisfies the conditions of Proposition 1, and hence $P^{(0)}, P^{(1)}, \ldots, P^{(i)}$ approaches $\Pr\left(P^{(i)} = p \,|r, \rho\right) = \exp\left[\rho\left(r\left(p\right)\right)\right] / \left[\sum_{p' \in \mathcal{P}_n} \exp\left[\rho\left(r\left(p'\right)\right)\right]\right]$ in distribution as $i \to \infty$. Since the summation condition on $x^{(i)}$ also implies that the total number of updating events must approach infinity almost surely as $t \to \infty$, this limit must hold in t as well.

It now remains to find the limiting distribution of $Y^{(i)}$. We proceed by first integrating over p, and then substituting from our previous results:

$$\Pr\left(Y^{(i)} = y \,|\rho\right) = \sum_{p \in \mathcal{P}_n} \Pr\left(Y^{(i)} = y \,\Big| P^{(i)} = p, \rho\right) \Pr\left(P^{(i)} = p \,|r, \rho\right) \quad (12)$$

$$= \sum_{p:r(p)=y} \Pr\left(P^{(i)} = p \,|r, \rho\right) \quad (13)$$

$$= \sum_{p:r(p)=y} \frac{\exp\left[\rho\left(r\left(p\right)\right)\right]}{\sum_{p' \in \mathcal{P}_n} \exp\left[\rho\left(r\left(p'\right)\right)\right]} \quad (14)$$

$$= \|\{p : r(p) = y\}\| \frac{\exp\left[\rho\left(y\right)\right]}{\sum_{p' \in \mathcal{P}_n} \exp\left[\rho\left(r\left(p'\right)\right)\right]}. \quad (15)$$

Since this limit also holds in t, the limiting distribution of $Y^{[t]}$ must be as derived. □

Theorem 1 provides us with an expression for the long-run cross-sectional behavior of Y, in the sense that a cross-sectional "snapshot" of Y taken at a random time will (in the large-t limit) have the specified distribution. This suggests, in practice, that models for cross-sectional data drawn from sources for which the above assumptions are reasonable may be (in some cases, at least) be given a behavioral interpretation. We now illustrate the use of these principles with a simple example, specifically involving advice-seeking behavior in a high tech firm.

4   Example: Advice-Seeking Among Managers

We demonstrate the utility of the behavioral framework via an illustrative analysis of a classic data set of Krackhardt (1987), consisting of self-reported advice-seeking behavior among 21 managers in a high-tech firm. In addition to the
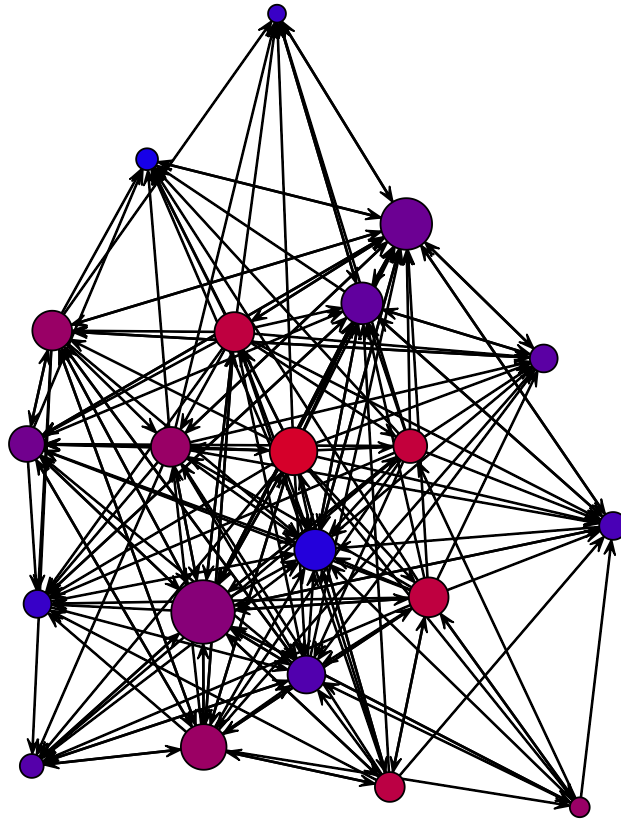
Figure 1: Krackhardt Advice Network Sociogram, Scaling and Shading by Degree

advice network, the data set includes attributions of friendship, as well as reporting ties (i.e., i reports to j as part of his or her regular duties within the firm). The networks here are cross-sectional, reflecting a snapshot of the social system at one point in time (see Figure 1). Nevertheless, it seems reasonable to presume that the observed ties arise from an underlying dynamic process, by which managers within the firm decide from whom they will seek advice given both the broader relational context (e.g., personal relationships and the formal organizational hierarchy) and the choices made by others.

To model the Krackhardt data, we begin by assuming that advice-seeking is a unilateral relation for which each individual chooses his or her outgoing edges. This leads immediately to a very simplified equilibrium distribution: since the prosphoric array in this case is exactly equivalent to the manifest relation, the equilibrium arising from Theorem 1 (assuming the relevant conditions are met)

	Deviance	Model df	AIC	Rank
Edges	578.43	1	580.43	7
Edges+Sender	441.12	21	483.12	4
Edges+Covar	548.15	3	554.15	5
Edges+Recip	577.79	2	581.79	8
Edges+Sender+Covar	385.88	23	431.88	2
Edges+Sender+Recip	405.38	22	449.38	3
Edges+Covar+Recip	547.82	4	555.82	6
Edges+Sender+Covar+Recip	**378.95**	**24**	**426.95**	**1**

Table 1: Model Selection Statistics for Krackhardt Advice Network, Independent Dyad Models

is simply a standard ERG with potential ρ. Assuming that the observed data is truly a snapshot from the long-run behavior of such a process, we may then estimate ρ (and hence the actor utilities, up an affine transformation) by existing methods.

To that end, we start by fitting models for standard independent dyad effects, interpreting them for our purposes in choice-theoretic terms. Table 1 shows basic goodness-of-fit information for models based on various combinations of edge formation preferences (homogeneous or inhomogeneous); edgewise covariate preferences (i.e., differential payoffs for seeking advice from friends and/or supervisors); and reciprocity preferences (i.e., differential payoffs for seeking advice from those who currently come to you for advice). As the table indicates, the favored model under the AIC is that including all effects (with heterogeneous sending preferences). Thus, we find evidence within the data not only for interpersonal differences in advice-seeking propensity, but also for attention to relational context and reciprocity in seeking out others for professional advice.

Next, we consider effects involving dependence (Table 1). In the present case, we fit models including two types of terms: differential preferences for edges which would induce three-cycles (i.e., seeking advice from those seeking advice from your current advisees); and differential preferences for edges which would produce transitive completions (i.e., seeking advice from your current advisors' advisors). Here too, the joint model wins out, although its goodness-of-fit versus the next best alternative (the same model without the transitive triad term) is only marginally superior. Thus, the data suggests that actors within this network are attentive to broader network features, above and beyond reciprocity and covariate effects.

To interpret the effects themselves, we now consider the estimated effects within the AIC-favored model (Table 3). Both the signs and magnitudes of the estimated effects are intuitive. Advice-seeking is generally aversive for most managers, although less so when seeking advice from friends. Seeking advice from managers is generally *preferred* (the total payoff is typically positive), reflecting the standard norm that one should gain input from one's supervisor

	Deviance	Model df	AIC	Rank
Edges+Sender+Covar+Recip	378.95	24	426.95	4
Edges+Sender+Covar+Recip+CycTriple	361.61	25	411.61	2
Edges+Sender+Covar+Recip+TransTriple	368.81	25	418.81	3
Edges+Sender+Covar+Recip+CycTriple+TransTriple	**358.73**	**26**	**410.73**	**1**

Table 2: Model Selection Statistics for Krackhardt Advice Network, Dependent Dyad Models
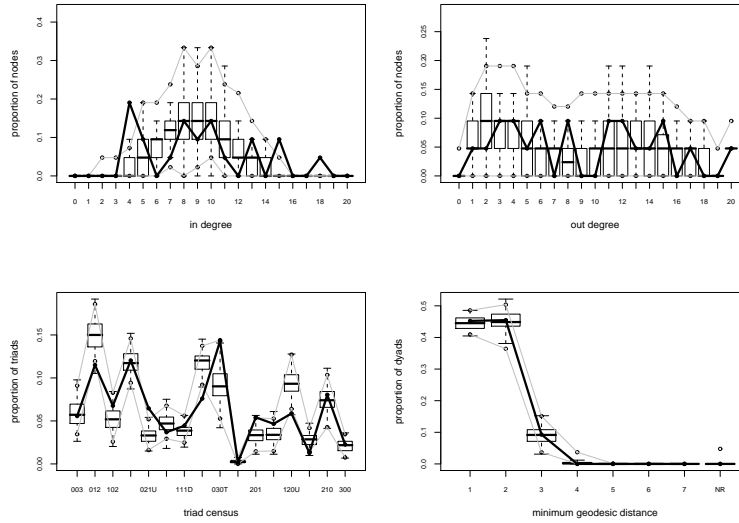
Figure 2: Model Adequacy Checks for Krackhardt Advice Network, AIC-Selected Model

regarding work-related matters. Reciprocal relationships reduce the cost of advice-seeking, but cyclic triads do not; this can be most easily explained by the notion that such activity takes place within a status hierarchy, and those significantly below one (one's advisees' advisees) are poor sources of information. Consistent with this intuition, we find that seeking advice from those to whom one has two-paths (one's advisors' advisors) is a higher-payoff activity. The data thus suggests not only attentiveness to triadic properties, therefore, but more generally a tendency towards status-seeking behavior within the network.

As a final note, it is also worth asking how well these simple factors do at accounting for broader network structure. To this end, Figure 2 shows model adequacy checks based on degree, triad census, and the distribution of geodesic distances within the graph. On the whole, the fitted model does fairly well: most simulated statistics cover the observed data, including those (like geodesics) based on global features not contained within the model. The relative success of these local properties in capturing other aspects of the data suggests that it is not necessary to posit that managers consider global network properties when making advice-seeking decisions. Of course, this does not rule out the possibility that such information is used in more subtle ways, nor that some other process than the one described here is at work. Nevertheless, the data does not suggest that such a thing is necessary to explain the bulk of the observed structure, while the simple behavioral model appears adequate to account for much of what is observed.

18

Effect	$\hat{\theta}$	s.e.	$\Pr(>|Z|)$		Effect	$\hat{\theta}$	s.e.	$\Pr(>|Z|)$	
Edges	**−1.022**	**0.137**	**0.0000**	$***$	**Sender14**	**−1.513**	**0.231**	**0.0000**	$***$
Sender2	**−2.039**	**0.637**	**0.0014**	$**$	**Sender15**	**16.605**	**0.336**	**0.0000**	$***$
Sender3	0.690	0.466	0.1382		**Sender16**	**−1.472**	**0.232**	**0.0000**	$***$
Sender4	−0.049	0.441	0.9112		**Sender17**	**−2.548**	**0.197**	**0.0000**	$***$
Sender5	0.355	0.495	0.4734		**Sender18**	**1.383**	**0.214**	**0.0000**	$***$
Sender6	**−4.654**	**1.540**	**0.0025**	$**$	**Sender19**	**−0.601**	**0.190**	**0.0016**	$**$
Sender7	−0.108	0.375	0.7726		Sender20	0.136	0.161	0.3986	
Sender8	−0.449	0.479	0.3486		Sender21	0.105	0.210	0.6157	
Sender9	0.393	0.496	0.4281		**Reciprocity**	**0.885**	**0.081**	**0.0000**	$***$
Sender10	0.023	0.555	0.9662		**Edgecov (Reporting)**	**5.178**	**0.947**	**0.0000**	$***$
Sender11	**−2.864**	**0.721**	**0.0001**	$***$	**Edgecov (Friendship)**	**1.642**	**0.132**	**0.0000**	$***$
Sender12	**−2.736**	**0.331**	**0.0000**	$***$	**CycTriple**	**−0.216**	**0.013**	**0.0000**	$***$
Sender13	**−0.986**	**0.194**	**0.0000**	$***$	**TransTriple**	**0.090**	**0.003**	**0.0000**	$***$

Null Dev 582.24; Res Dev 358.73 on 394 df

Table 3: Parameter Estimates for Krackhardt Advice Network, AIC-Selected Model

5 Conclusion

In this paper, we have provided a basic micro-foundation for cross-sectional network models, based on a stochastic choice process. While not applicable in all circumstances, behaviorally plausible conditions do exist for which this process will give rise to well-defined equilibrium behavior, and for which said equilibrium can be expressed in terms of individual utilities. The presence of such a relationship allows for estimation and comparison of models for agent preferences from cross-sectional data, without explicit knowledge of the underlying dynamic process. It is hoped that the results shown here will lead to further investigation of the relationship between choice processes and social networks, and to a clearer understanding of what can (and cannot) be inferred from cross-sectional network data.

6 References

Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory.* John Wiley and Sons, New York.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 36(2):192–236.

Brown, L. D. (1986). *Fundamentals of Statistical Exponential Families, with Applications in Statistical Decision Theory.* Institute of Mathematical Statistics, Hayward, CA.

Crouch, B., Wasserman, S., and Trachtenburg, F. (1998). Markov chain Monte Carlo maximum likelihood estimation for $p*$ social network models. Paper presented at the XVIII International Sunbelt Social Network Conference, Sitges, Spain.

Frank, O. and Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association*, 81:832–842.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). Intoducing Markov chain Monte Carlo. In Gilks, W. R., Richardson, S., and Spiegelhalter, D. J., editors, *Markov Chain Monte Carlo in Practice*, pages 1–20. Chapman and Hall, London.

Holland, P. W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs (with discussion). *Journal of the American Statistical Association*, 76(373):33–50.

Hunter, D. R. and Handcock, M. S. (2006). Inference in curved exponential family models for networks. *Journal of Computational and Graphical Statistics*, 15:565–583.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., and Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, 24(3).

Krackhardt, D. (1987). Cognitive social structures. *Social Networks*, 9(2):109–134.

Krackhardt, D. (1992). The Strength of Strong Ties: the Importance of Philos in Organizations. In Nohria, N. and Eccles, R., editors, *Networks and Organizations: Structures, Form, and Action*, pages 216–239. Harvard Business Press, Boston, MA.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*. Academic Press.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, 14:124–143.

Pattison, P. and Robins, G. (2002). Neighborhood-based models for social networks. *Sociological Methodology*, 32:301–337.

Pattison, P. and Wasserman, S. (1999). Logit models and logistic regressions for social networks: II. multivariate relations. *British Journal of Mathematical and Statistical Psychology*, 52:169–193.

Robins, G., Pattison, P., and Wasserman, S. (1999). Logit models and logistic regressions for social networks, III. valued relations. *Psychometrika*, 64:371–394.

Robins, G. L. and Pattison, P. (2001). Random graph models for temporal processes in social networks. *Journal of Mathematical Sociology*, 25:5–41.

Snijders, T. A. B. (1996). Stochastic actor-oriented models for network change. *Journal of Mathematical Sociology*, 23:149–172.

Snijders, T. A. B. (2002). Markov Chain Monte Carlo estimation of exponential random graph models. *Journal of Social Structure*, 3(2).

Snijders, T. A. B. (2005). Models for longitudinal network data. In Carrington, P. J., Scott, J., and Wasserman, S., editors, *Models and Methods in Social Network Analysis*, pages 215–247. Cambridge University Press, New York.

Snijders, T. A. B., Pattison, P. E., Robins, G. L., and Handcock, M. S. (2006). New specifications for exponential random graph models. *Sociological Methodology*, 36.

Strauss, D. (1986). On a General Class of Models for Interaction. *SIAM Review*, 28(4):513–527.

Strauss, D. and Ikeda, M. (1990). Pseudolikelihood Estimation for Social Networks. *Journal of the American Statistical Association*, 85(409):204–212.

Wasserman, S. and Pattison, P. (1996). Logit models and logistic regressions for social networks: I. an introduction to Markov graphs and $p*$. *Psychometrika*, 60:401–426.

Wasserman, S. and Robins, G. (2005). An introduction to random graphs, dependence graphs, and $p*$. In Carrington, P. J., Scott, J., and Wasserman, S., editors, *Models and Methods in Social Network Analysis*, chapter 10, pages 192–214. Cambridge University Press, Cambridge.