Bistability Through Triadic Closure

Peter Grindrod* Desmond J. Higham† Mark C. Parsons‡

December 12, 2011

Abstract

We propose and analyse a class of evolving network models suitable for describing a dynamic topological structure. Applications include telecommunication, on-line social behaviour and information processing in neuroscience. We model the evolving network as a discrete time Markov chain, and study a very general framework where, conditioned on the current state, edges appear or disappear independently at the next timestep. We show how to exploit symmetries in the microscopic, localized rules in order to obtain conjugate classes of random graphs that simplify analysis and calibration of a model. Further, we develop a mean field theory for describing network evolution. For a simple but realistic scenario incorporating the triadic closure effect that has been empirically observed by social scientists (friends of friends tend to become friends), the mean field theory predicts bistable dynamics, and computational results confirm this prediction. We also discuss the calibration issue for a set of real cell phone data, and find support for a stratified model, where individuals are assigned to one of two distinct groups having different within-group and across-group dynamics.

Keywords calibration, conjugacy, dynamic network, mean field theory, random graph, stochastic process, temporal network, triangulation, voice call data.

1 Preliminaries

A diverse range of application areas give rise to large, complex interaction patterns. In the field of network science, classes of random graph have been proposed and tested as models to capture the structure of these interactions [24]. In many cases, the model may be viewed as an iterative procedure that builds a network sequentially by randomly

*Department of Mathematics, University of Reading, UK
†Department of Mathematics and Statistics, University of Strathclyde, UK
‡Department of Mathematics, University of Reading, UK

rewiring an existing structure [25, 30, 31] or by adding nodes and links in order to 'grow' a network [3]. In these cases the object of interest, however, is the final, static network. Our work differs in that we wish to model a network structure that is inherently dynamic, with edges appearing and disappearing between a fixed set of nodes. Such a scenario arises naturally in many modern, data-rich, applications; for example, telecommunication (who phoned whom each day), on-line social interaction (who interacted with whom in a chat room), on-line retailing (people who bought this book also bought) [17]. Attention has recently been paid to the issue of extending traditional graph theoretical concepts such as paths to the time-dependent setting [5, 13, 16, 18, 28, 29] and related computational complexity issues [1, 7, 8]. There has also been interest in the waiting times between link changes [2, 32] and the emergence of communities [4, 23]. Also, *link prediction*—estimating likely new connections a short time ahead—is being recognized as an important task [10, 20, 22]. However, in this work we focus on the fundamental issue of modeling and analysing such networks directly and from first principles; that is, prescribing reasonable 'laws of motion' and studying the potential behaviors that can arise. Our target applications are digitally generated communication or on-line social interaction networks and our interest lies in the changes of the connectivity structure itself. Related work on *adaptive networks* [14] has studied scenarios where the nodes are involved in their own dynamical system that is coupled to the dynamic topology—for example, in an epidemiological SIS model, susceptible (S) nodes may seek to avoid links with nodes that are currently infected (I).

The main novel contributions in our work are (a) to introduce the concept of conjugate graphs, which can play an important role in understanding and analysing a model, (b) to derive a mean field theory approach to summarizing long term behaviour, (c) to introduce a simple but realistic nonlinear network evolution model driven by the concept of triadic closure from social science and to show that it admits bistable behaviour, and (d) to consider the issue of model calibration and show that the use of a stratified model improves the fit for a voice call data set.

The presentation is organised as follows. In the next section, we introduce a general stochastic framework and show how modelling and simulation can be simplified by focussing on the dynamics of individual edges. We also illustrate the ideas in the particular case where triadic closure is encouraged—friends of friends tend to become friends. Section 3 then introduces the concept of conjugacy, which can be used to describe inherent symmetries in a model. A mean field approach is described in section 4 and applied to the triadic closure model in section 5. This model, and a more general stratified version, is fitted to voice call data in section 6, and conclusions are given in section 7.

2 Preliminaries

Throughout this work we focus on undirected graphs on n vertices with no loops, which may be represented by symmetric, binary $n \times n$ adjacency matrices. Here $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ has $a_{ij} = a_{ji} = 1$ if there is an edge from node i to node j and has $a_{ij} = a_{ji} = 0$

otherwise, with all $a_{ii} = 0$. Let S_n denote the set of all such adjacency matrices and let $\mathbf{1}$ denote the adjacency matrix for the n-vertex clique. If $A \in S_n$ then $\mathbf{1} - A \in S_n$ is the adjacency matrix for the complementary graph to that represented by A. Let R_n denote the set of symmetric $n \times n$ real matrices with all elements taking values in $[0, 1]$, with zeros on the main diagonal. So $S_n \subset R_n$. For real $n \times n$ matrices M_1 and M_2 the Hadamard (or Schur) product, denoted by $M_1 \circ M_2$, is the matrix obtained by element-wise multiplication, so that

$$(M_1 \circ M_2)_{ij} = (M_1)_{ij}(M_2)_{ij}.$$

If $A_1,\ A_2 \in S_n$ then $A_1 \circ A_2 \in S_n$ represents the adjacency matrix for the graph of common edges.

Following the treatment in [12], we use the phrase *evolving network model* to describe a stochastic rule that generates a sequence of networks represented by a sequence of adjacency matrices, $\{A_k\}_{k=0}^{K}$, where each $A_k \in S_n$. We say that A_k represents the state of the evolving network at the kth time step, t_k, where $t_0 < t_1 < \ldots < t_K$ are equally spaced points in time.

It is natural and analytically convenient to focus on Markovian models. We therefore consider the case of a *first order evolving network model* characterized by the conditional probability distribution for A_{k+1} given A_k, denoted by $P(A_{k+1}|A_k)$, defined for all pairs $A_{k+1},\ A_k \in S_n$. The expected value of A_{k+1} given A_k will be written as

$$\langle A_{k+1}|A_k\rangle := \sum_{A_{k+1}\in S_n} A_{k+1}P(A_{k+1}|A_k).$$

By construction $\langle A_{k+1}|A_k\rangle \in R_n$ and the (i, j)th element of $\langle A_{k+1}|A_k\rangle$ contains the conditional probability that the corresponding edge is present in A_{k+1}.

We will say that the first order evolving network model, $P(A_{k+1}|A_k)$, is *edge independent* if, given A_k, information about the existence of any particular edges in A_{k+1} has no effect on the probability that any other edge is in A_{k+1}.

We also note that under this assumption of edge independence, any element $W \in R_n$ defines a random graph. The (i, j)th element of W contains the probability that the corresponding edge is present and we can construct the associated probability distribution over S_n, say $P_W(A)$:

$$P_W(A) = \prod_{i=1,j=i+1}^{n} (W)_{ij}^{(A)_{ij}}\ (1 - (W)_{ij})^{1-(A)_{ij}}.$$

For example, if we have $p \in (0, 1)$ then $p\mathbf{1} \in R_n$ represents a classical Erdös-Rényi/Gilbert random graph (usually denoted $G(n, p)$), where each edge exists with independent probability p [24]. Let $CL_{(n,k)} \in S_n$ represent the adjacency matrix for a circular lattice on n vertices, with each vertex connected to its k nearest clockwise and its k nearest anti-clockwise neighbours. Then for any constants $0 < q \leq p < 1$, $pCL_{(n,k)}$ is a partial lattice

with every edge present with independent probability p, while $pCL_{(n,k)} + q(\mathbf{1} - CL_{(n,k)})$ is a partial lattice with uniform "short cuts" in the spirit of the classical Watts-Strogatz small world model [25, 31].

In restricting ourselves to first order edge independent evolving network models we may simply consider matrix valued functions

$$\mathcal{F} : S_n \to R_n. \tag{1}$$

Any such mapping $\mathcal{F}$ generates a first order edge independent evolving network model for which

$$< A_{k+1} | A_k >= \mathcal{F}(A_k). \tag{2}$$

A particularly useful form for $\mathcal{F}(A_k)$ is

$$\mathcal{F}(A_k) = (\mathbf{1} - \omega(A_k)) \circ A_k + \alpha(A_k) \circ (\mathbf{1} - A_k), \tag{3}$$

where $\alpha(A_k)$ and $\omega(A_k)$ are given mappings $S_n \to R_n$, representing conditional birth rates and death rates respectively, as introduced in [12]. It is straightforward to compute a path for such a Markov chain; that is, a particular network sequence whose transitions respect the relevant edge birth and death rates. The following pseudo-code summarizes this approach, given an initial adjacency matrx, A_0.

for k = 0,1, 2, ...

Compute $\alpha(A_k), \omega(A_k) \in R_n$

for all disjoint pairs $i \neq j$

if $(A_k)_{ij} = 0$ then set

$(A_{k+1})_{ij} = 1$ with prob. $\alpha(A_k)_{ij}$ (birth)

$(A_{k+1})_{ij} = 0$ with prob. $1 - \alpha(A_k)_{ij}$ (no change)

else we have $(A_k)_{ij} = 1$, so set

$(A_{k+1})_{ij} = 0$ with prob. $\omega(A_k)_{ij}$ (death)

$(A_{k+1})_{ij} = 1$ with prob. $1 - \omega(A_k)_{ij}$ (no change)

end if

end for all pairs

end for k

We finish this section by introducing a novel example that fits into this framework. We consider the case where the death rate is a constant for all edges

$$\omega(A_k) \equiv \widetilde{\omega}\mathbf{1}, \quad \text{for } \widetilde{\omega} \in (0,1), \tag{4}$$

while the birth rates for edges not present in A_k are given by

$$\alpha(A_k) = \delta\mathbf{1} + \epsilon\mathbf{1} \circ A_k^2, \tag{5}$$

for some constants δ and ϵ. We assume that $0 < \delta \ll 1$ and, to guarantee probabilities in the range $[0,1]$, that $0 < \epsilon(n-2) < 1 - \delta$.

This model reflects a situation where the more common adjacencies two nonadjacent vertices have in A_k, the more likely they are to become adjacent in A_{k+1}. In other words, somebody who is not currently your friend, but who is currently a friend of many of your current friends, has an enhanced chance of becoming your friend at the next step. Forming new associations by the process of triangulating current adjacencies is very natural within a number of applications. In social networks two peers may be likely to become introduced through common friends; in this context, the mechanism is often referred to as triadic closure [11, 19, 27]. In developing cognitive processing capability in the brain, triangulation increases efficiency of communication and resilience, and overabundance of triangles has been observed in both anatomical and functional studies [6, 15, 21].

In Figure 1 we show the evolution of a network path from this model with $n = 100$ nodes and parameter values

$$\widetilde{\omega} = 0.01, \quad \epsilon = 0.0005, \quad \delta = 0.0004. \tag{6}$$

A dot in row i and column j denotes an edge from node i to node j. The initial network was a sample of an Erdös-Rényi graph with expected edge density $p = 0.3$—each possible edge exists with independent probability p. The figure shows the adjacency matrix at times t_k for $k = 50, 100, 150, \ldots, 750$. We see that the density of edges increases with time. Figure 2 gives a histogram of the nodal degrees at the final time, $k = 750$. The structure appears to be approximately Poisson, and the edge density

$$\widehat{p}_k := \frac{1}{n(n-1)/2} \sum \sum_{i>j} (A^{[k]})_{ij}, \tag{7}$$

at the final time was $\widehat{p}_{750} = 0.712$.

In Figure 3 we repeat the experiment with the same parameter values, but start with an Erdös-Rényi graph with a lower expected edge density of $p = 0.15$. In this case we see that the edge density decreases over time. The final time value was $\widehat{p}_{750} = 0.051$. The degree distribution, shown in Figure 4, again appears to be approximately Poisson.

The analysis that we develop in sections 3 and 4 will explain the difference in behavior that we see between these two cases, pointing out for the first time in this context that bistability arises naturally.

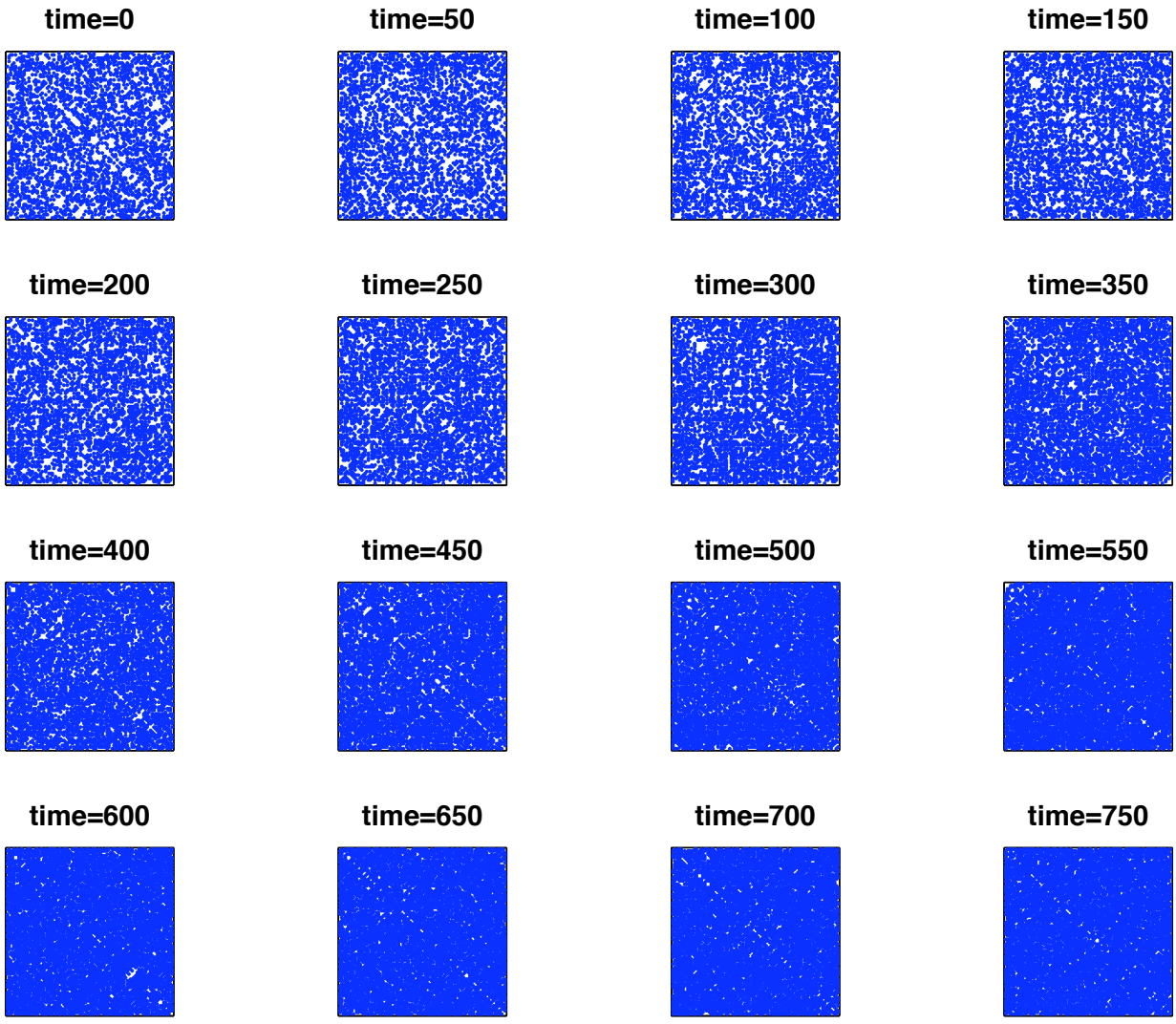

Figure 1: Network evolution from the triangulation model (4)–(5) with $n = 100$ nodes and parameter values from (6). Initial network is an Erdös-Rényi sample with expected edge density of $p = 0.3$.

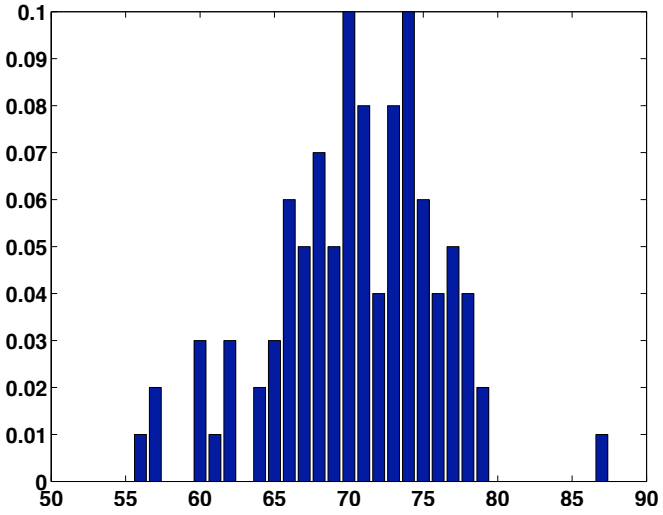

Figure 2: Degree distribution for the final network (after 750 iterations) in Figure 1.

3 Classes of Conjugate Random Graphs

In this section we show how symmetries in a model give rise to a natural definition of conjugacy. Assume that the mapping $\mathcal{F}$ in (1) characterising a first order edge independent evolving network is subordinate to a mapping $\mathcal{F} : R_n \to R_n$. We will see in section 4 how this can be rather natural in some examples when we wish to replace A_k by its own expected value. Let λ denote a (finite or infinite dimensional) parameter ranging over a domain Λ within some suitable space. Let $W : \Lambda \to R_n$. We will say that $W(\lambda)$ is a *class of conjugate random graphs* for $\mathcal{F}$ if for each $\lambda \in \Lambda$ there exists a unique element $g(\lambda) \in \Lambda$ such that

$$W(g(\lambda)) = \mathcal{F}(W(\lambda)).$$

Hence the parameterized set $\{W(\lambda)|\lambda \in \Lambda\} \subset R_n$ is positively invariant under $\mathcal{F}$. Moreover the action of $\mathcal{F}$ on R_n can in this case be reduced to the action of g on Λ.

Suppose that $\mathcal{F} : R_n \to R_n$ also possesses some symmetries. That is, suppose that there is a subgroup of $n \times n$ permutation matrices, $H = \{Q_r\}$, such that $\mathcal{F}$ is invariant under each of these these permutations; so that

$$\mathcal{F}(A) = Q_r \mathcal{F}(Q_r^T A Q_r) Q_r^T, \quad A \in S_n, \quad Q_r \in H.$$

Then $\mathcal{F}$ uses no a priori (extra) information about the vertices that distinguishes one permutation in H from another. Now let $\mathcal{W} \subset R_n$ denote the subset of random graphs in R_n that are invariant under the symmetries in H, so that

$$Q_r^T W Q_r = W,$$

for all $W \in \mathcal{W}$, $Q_r \in H$. It then follows that

$$Q_r^T \mathcal{F}(W) Q_r = \mathcal{F}(Q_r^T W Q_r) = \mathcal{F}(W).$$

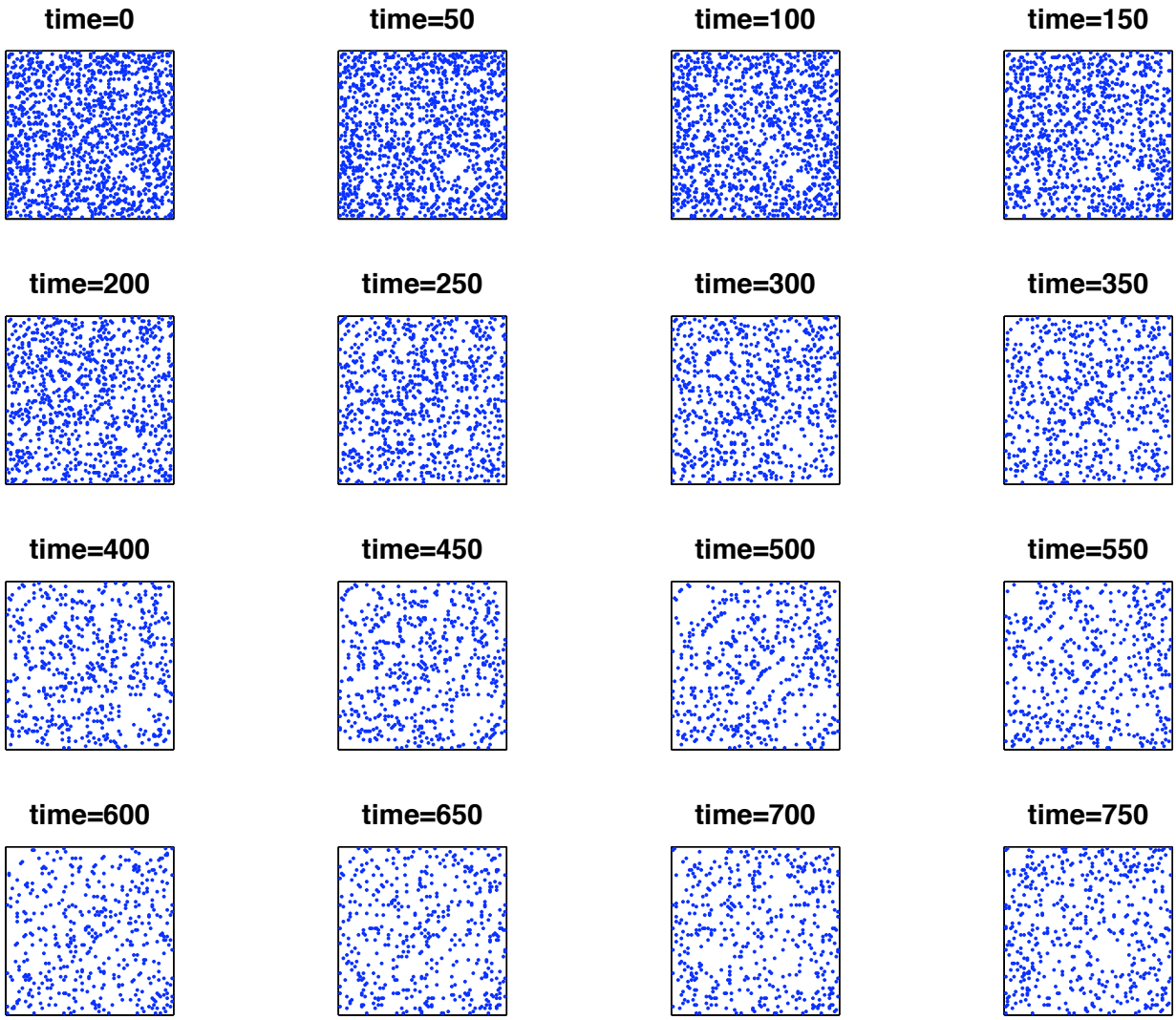

Figure 3: As for Figure 1, with the initial network as an Erdös-Rényi sample with expected edge density of $p = 0.15$.

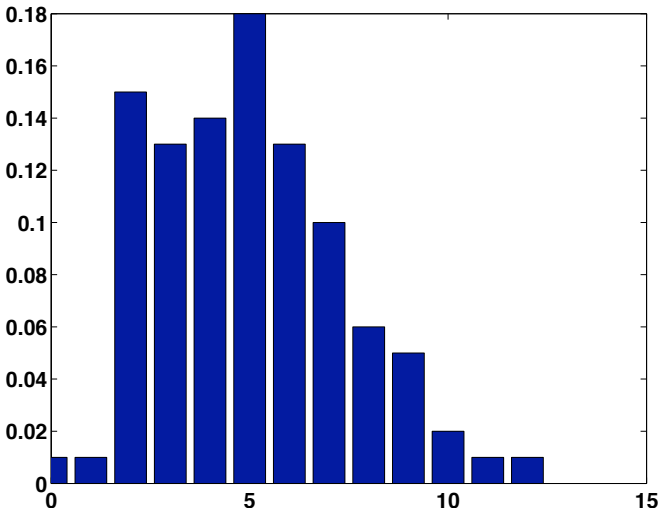

Figure 4: Degree distribution for the final network (after 750 iterations) in Figure 3.

Hence $\mathcal{F}$ maps $\mathcal{W}$ into $\mathcal{W}$. So, under a suitable parametrization, with some $\lambda \in \Lambda$, the subset $\mathcal{W}$ is a possible class of conjugate random graphs for $\mathcal{F}$.

In the simple case where $\mathcal{F}$ is invariant under all possible permutations on n vertices, we know that no subset of nodes is distinguished in any way. For example this will certainly be so whenever $\mathcal{F}(W) = Q_1(W) \circ Q_2(W) \circ ... \circ Q_S(W)$, where each of the $Q_s(W) : R_n \to R_n$ is a polynomial, $Q_s(W) = \beta_0 \mathbf{1} + \beta_1 W + \beta_2 W^2 + ... + \beta_\mu W^\mu$, say, for suitable nonnegative constants $\beta_0, ..., \beta_\mu$. The random graph $W = p\mathbf{1}$ is invariant under all such permutations; $\mathcal{W}$ is the set of such graphs, parameterized by $p \in [0, 1]$. This situation represents a kind of egalitarian or democratic scenario where every node is the same and none is distinguished based on any a priori information. So we must treat all edges and all edge birth dynamics according to the same model.

Alternatively, we may have a *stratified model* where some extra prior information divides the nodes onto two disjoint subsets. (The generalization to a finer partition is immediate.) This could represent a social network between individuals from two sexes for instance, where it was posited that male-male, female-male and female-female interactions follow different birth/death dynamics. Similarly, there may be an elite and non-elite (officers and troops), or cultural, or functional splitting of the vertices. To be concrete, suppose that all edges between any of the first n_1 vertices satisfy a given identical dynamic; all edges between any of the last $n_2 = n - n_1$ vertices satisfy a distinct given identical dynamic; and all edges between any of the first n_1 vertices with any of the last n_2 vertices satisfy a third given identical dynamic. Then $\mathcal{F}$ will be invariant under all permutations Q that permute the two subsets separately, and that do not swap any vertices between them. In that case W is invariant with respect to all such permutations

if it has a symmetric block structure: say

$$\begin{aligned}(W)_{ij} &= 0 \text{ if } i = j\\ (W)_{ij} &= p \text{ if } i \neq j \text{ and } i, j \leq n_1\\ (W)_{ij} &= q \text{ if } i \neq j \text{ and } i, j > n_1\\ (W)_{ij} &= r \text{ if } \min\{i, j\} \leq n_1 \text{ and } \max\{i, j\} > n_1.\end{aligned}$$

Then $\mathcal{W}$ is the set of such graphs, parameterised by three constants: $\lambda = (p, q, r) \in [0, 1]^3 = \Lambda$.

For such a stratified random graph, $W \in \mathcal{W}$, the expected number of edges is given by

$$\frac{n_1(n_1 - 1)}{2}p + \frac{n_2(n_2 - 1)}{2}q + n_1 n_2 r.$$

We also recall that the Watts-Strogatz clustering coefficient for a node is defined as the ratio of links between the vertices within its neighbourhood divided by the number of links that could possibly exist between them [24, 31]. The expected Watts-Strogatz clustering coefficient then has the form

$$\frac{\frac{n_1!}{2!(n_1-3)!}p^3 + \frac{n_2!}{2!(n_2-3)!}q^3 + \frac{n_1 n_2!}{2!(n_2-2)!}r^2 q + \frac{n_2 n_1!}{2!(n_1-2)!}r^2 p + n_1(n_1 - 1)n_2 p r^2 + n_2(n_2 - 1)n_1 q r^2}{\frac{n_1!}{2!(n_1-3)!}p^2 + \frac{n_2!}{2!(n_2-3)!}q^2 + \frac{n_1 n_2!}{2!(n_2-2)!}r^2 + \frac{n_2 n_1!}{2!(n_1-2)!}r^2 + n_1(n_1 - 1)n_2 p r + n_2(n_2 - 1)n_1 q r}.$$

This expression has six terms in each of the denominator and the numerator, representing the expected number of open jaws (pairs of edges from a *central* vertex to two other distinct vertices), and the expected number of those open jaws that are complete triangles, respectively. Six terms arise because the *central* vertex (of the open jaw) may be in either of the stratified subsets, while the other two vertices may be such that none, one, or both lie within the same subset. The expression simplifies to yield

$$\frac{3(n_1 - 1)n_1 n_2 p r^2 + 3n_1(n_2 - 1)n_2 q r^2 + (n_1 - 2)(n_1 - 1)n_1 p^3 + (n_2 - 2)(n_2 - 1)n_2 q^3)}{(2(n_1 - 1)n_1 n_2 p r + 2n_1(n_2 - 1)n_2 q r + (n_1 - 1)n_1 n_2 r^2 + n_1(n_2 - 1)n_2 r^2 + (n_1 - 2)(n_1 - 1)n_1 p^2 + (n_2 - 2)(n_2 - 1)n_2 q^2)}.$$

In section 6 we compare unstratified and stratified models on voice call data.

4 A Mean Field Approximation for Evolving Networks

Our aim in this section is to develop a heuristic approach for analysing the behaviour of an edge independent first order evolving network (2). We begin with the simplifying assumption that A_k has the properties of its own mean (expected) graph, $\langle A_k | X \rangle$, given any prior information X, rather than including all of the details of any particular value for A_k. Then we may use this expected value to calculate the consequent expectation, $\langle A_{k+1} | \langle A_k | X \rangle \rangle$.

In reality of course each of the edges in A_k is there or not, taking a binary value. Hence by employing the expectation for A_k, the analysis is only an approximation. It will be particularly poor at times when the probability of edges appearing is relatively small and/or the existence and distribution of just a few edges has a critical effect: they cannot in reality be *smeared out*.

Alternatively, we may only be given $\langle A_k|X\rangle \in R_n$ as a random graph itself, conditional on any previous information X, and we may wish to use our evolving graph model to calculate an estimate for $\langle A_{k+1}|X\rangle$, and so on. We should calculate

$$\langle A_{k+1}|X\rangle = \sum_{A_k \in S_n} \mathcal{F}(A_k) P(A_k|X).$$

Instead we might calculate the approximation

$$\langle A_{k+1}|X\rangle \approx \mathcal{F}(\sum_{A_k \in S_n} A_k P(A_k|X)) = \mathcal{F}(\langle A_k|X\rangle). \tag{8}$$

There is equality here if the only nonlinearities in $\mathcal{F}$ involve the multiplication of independent stochastic variables. In most cases we will have to consider the expected number(s) of some combinations of edge being present. Here the edge independence assumption is exactly what we need. For example, the expected value of the number of mutual adjacencies (for any give pair of vertices) involves a sum over all pairs of edges connecting to the possible mutual adjacent vertex. These are mutually independent, and of course the two necessary edges within each term are mutually independent. Hence for this type of $\mathcal{F}$, with each term in the range involving only sums over independent events each of which itself is a product over individual edges, (8) is exact. We refer to (8) as a *mean field approximation* for the evolving graph.

Suppose that we may represent $< A_k|X >$ by some random graph, say $W_k \in R_n$. Then using the mean field approximation we simply iterate with $\mathcal{F}$:

$$W_{k'+1} = \mathcal{F}(W_{k'}), \quad k' = k, k+1, k+2, \ldots$$

to obtain $\langle A_{k'}|X\rangle = W_{k'}$ for all $k' = k, k+1, k+2, \ldots$. This iteration generates a sequence of expected values for the evolving network at all future time steps, given the approximation for A_k, but using the mean field approximation.

Now let us assume that $W(\lambda)$ is a conjugate random graph for $\mathcal{F}$. If we have $W_k = W(\lambda_k)$, for some $\lambda_k \in \Lambda$, then we can iterate with g to produce a sequence

$$\lambda_{k'+1} = g(\lambda_{k'}), \qquad k' = k, k+1, k+2, \ldots,$$

and it follows that

$$\langle A_{k'}|X\rangle = W_{k'} = W(\lambda_{k'}), \qquad k' = k, k+1, k+2, \ldots$$

Hence under the mean field approximation a conjugate random graph is particularly useful, as it descends from the mean field iteration over R_n to one over Λ.

5 Bistability Through Triadic Closure

We now apply this mean field theory to the triadic closure model (4)–(5), where

$$\mathcal{F}(A_k) = (1 - \widetilde{\omega})A_k + (\mathbf{1} - A_k) \circ (\delta \mathbf{1} + \epsilon A_k^2). \tag{9}$$

It is easy to see by symmetry of the model (there being no distinguished vertices nor differences in the way vertices and edges are treated) that the Erdös-Rényi graphs are possible conjugate random graphs for this mapping. Substituting $\langle A_k | X \rangle = p_k \mathbf{1}$ for A_k as the mean field approximation, we obtain the iteration $p_{k+1} = g(p_k)$ where

$$g(p) = (1 - \widetilde{\omega})p + (1 - p)(\delta + \epsilon(n-2)p^2). \tag{10}$$

At equilibrium $p_{k+1} = p_k \equiv p^\star$, where

$$p^\star = (1 - p^\star)(\delta + \epsilon(n-2)p^{\star 2})/\widetilde{\omega}. \tag{11}$$

In the limit $\delta \to 0$, there are three real roots $p^\star$ if and only if $\widetilde{\omega} < \epsilon(n-2)/4$. The smallest is $\delta/(\delta + \widetilde{\omega}) + O(\delta^2)$, representing a sparse graph with almost no triangulation, where the random birth rate, δ, equilibrates with death rate, $\widetilde{\omega}$; and the the larger roots are at $1/2 \pm \sqrt{1/4 - \widetilde{\omega}/\epsilon(n-2)} + O(\delta)$, where the nonlinear triangulation term equilibrates with the death rate, $\widetilde{\omega}$. Moreover $g'(p) \geq 0$ for all $p \in [0, 1]$ and $g(0) = \delta$. Hence, in this regime, the two outer steady states are stable for this iteration, whilst the middle root is unstable. Intuitively, with a low initial edge density the triangulation rule cannot get started and the network remains sparse, whereas for a sufficiently high initial edge density, the network evolves into an ϵ-dependent state. Figure 5 illustrates the case with $n = 100$ and model parameters from (6), where there are stable fixed points at 0.049 and 0.721 surrounding an unstable fixed point at $p^\star = 0.229$. These values are consistent with the experiments in Figures 1–4.

As a further test, Figure 6 shows the results of a single simulation with the same model parameters as in Figure 5. As the initial network, we sampled an Erdös-Rényi random graph with edge probability $p = 0.3$. The jagged curve in the figure shows the edge density (7) against t_k. The solid curve represents the mean field recurrence from (10), with $p_0 = 0.3$. We see that there is very good agreement when this macroscopic quantity is computed directly from the full microscopic simulation and from the mean field approximation.

Our mean field analysis predicts that the long term behaviour may be sensitive not only to the initial conditions but also to the transient stochastic fluctuations. Figure 7 confirms this effect by showing five separate simulations with the same model parameters, all starting out from the same graph, a sample of an Erdös-Rényi graph with edge probability $p = 0.23$. This value was chosen deliberately to be close to the unstable middle fixed point of the mean field equation. As in Figure 6, we plot the edge density $\widehat{p}_k$ at each time step. We see that four paths evolve towards the lowest mean field solution, $p^\star = 0.049$, and one breaks out towards the highest mean field solution $p^\star = 0.721$.

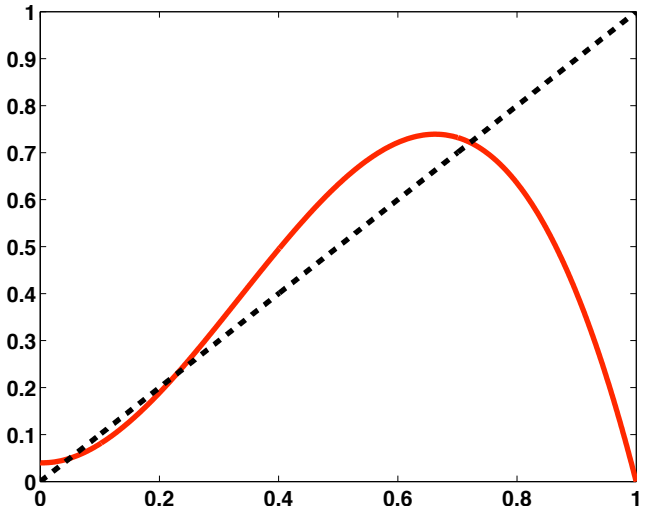

Figure 5: Graph of p (dashed) and $(1-p)(\delta+\epsilon(n-2)p^2)/\widetilde{\omega}$ (solid): the fixed points are at 0.049, 0.229 and 0.721. (Here $n = 100$ and parameter values are taken from (6).)

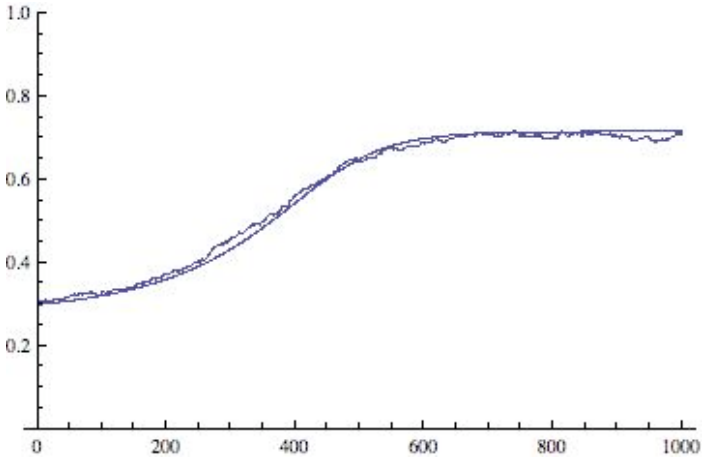

Figure 6: Comparison of $\widehat{p}_k$ in (7) estimated from a simulation with solution to the dynamical mean field equation (10).

At each t_k, we can calculate the average Watts-Strogatz clustering coefficient, C_k, defined as the clustering coefficient averaged over all nodes. It is interesting to compare its evolution with that of the edge density, $\widehat{p}_k$. Since the initial network and long term networks are Erdös-Rényi graphs we will have have $C_k = \widehat{p}_k$ there. For the simulation in Figure 8, we see that C_k increases slightly ahead of $\widehat{p}_k$, as initially random clusters strengthen, before they infill at the higher density. The individual, vertex-wise, values of the clustering coefficient increase in variance during the phase of rapid growth; see Figure 9.

To finish this section, we emphasize our belief that in order to understand network formation and forecast network behaviour it is essential to have a class of realistic dynamical models capable of replicating behaviour seen in evolving network structures. Indeed, if we merely observe evolving social and communication networks, then, as with any

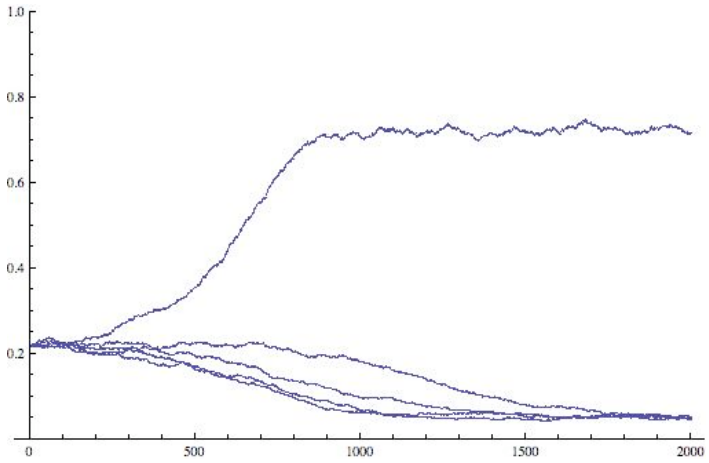

Figure 7: Five simulations starting from the same network: we show the fraction of edges $\widehat{p}_k$ in (10) at each time step (equivalent to the edge probability p for an evolving Erdös-Rényi random graph).

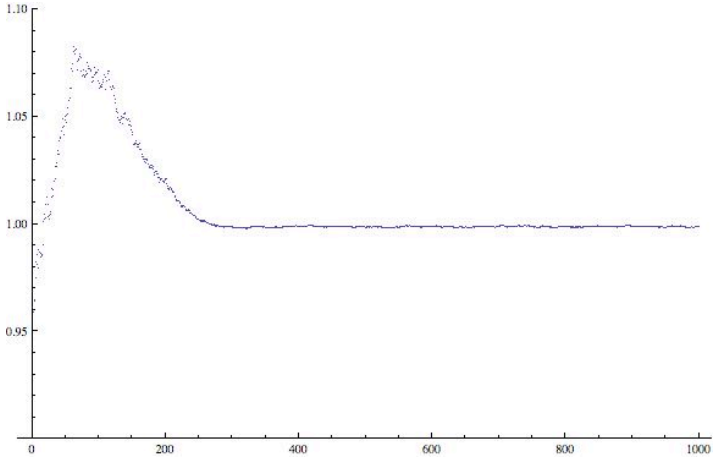

Figure 8: Ratio of average clustering coefficient to edge density: $C_k/\widehat{p}_k$.

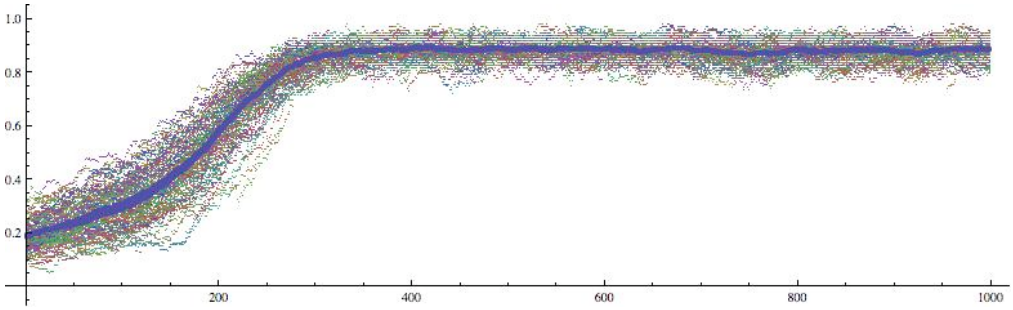

Figure 9: Vertex-level clustering coefficients as the network in Figure 8 evolves.

dynamical system, we only see some subset of stable behavior. The bistability effect shown in this section confirms that a single realisation can give a misleading picture, and also opens up the possibility of *timely and targeted intervention*—for example, a

service provider may wish to stimulate early activity to move the edge density into a region where the network will then self-organise into a profitable, well connected regime.

6 Application to Voice Call Data

In this section we consider an evolving network data set of pairwise mobile phone communications during a period of the year where connections were on the increase. We compare the unstratified (homogeneous) model and a stratified model introduced in section 3, by first calibrating them via the evolution of the appropriate edge densities, and then examining how they predict the corresponding possible evolution of the average Watts-Strogatz clustering coefficient.

Suppose that we observe an evolving random graph. By making an assumption about $\mathcal{F}$, and selecting an appropriate subgroup of permutations, Q, we may derive the mean field equations, over $\mathcal{W}$, which will involve the dynamical parameters from $\mathcal{F}$. From data we can calculate the evolution of the coordinates λ describing $\mathcal{W}$. By fitting these to the mean field model we can estimate the unknown dynamical parameters. Then the choice of model may be validated or invalidated by checking other evolving network metrics not used to do the calibration.

For example here we consider weeks 8 through 15 of data from the Reality Mining data set given in [9], showing voice call between around 150 people over time. We consider the weekly call networks, summarizing which pairs of people communicated during successive weeks.

Assuming initially that no people are distinctive in any way, we first fit the simple three parameter unstratified model in (9). Assuming a homogeneous population, since $\mathcal{F}$ is invariant under all permutations, the mean field dynamics act over $\mathcal{W} = \{p\mathbf{1} \,|\, p \in [0,1]\}$. So, as described in section 5 we have $p_{k+1} = g(p_k)$ for g in (10). From the data we can estimate p_k via $\widehat{p}_k$ in (7) , the density of edges present. We have

$$\widehat{p}_{k+1} = (1-\widetilde{\omega})\widehat{p}_k + (1-\widehat{p}_k)(\delta + \epsilon(n-2)\widehat{p}_k^2) + \mathrm{err}_k,$$

where the errors, err_k, arise as an average over $n(n-1)/2$ independent edge-processes each of which must take binary values. Thus a Gaussian approximation to the structure of the errors is reasonable and the parameters may be fitted with simple least squares. This results in estimates $(\delta, \epsilon, \widetilde{\omega}) = (0.02170, 0.00868, 0.18399)$. In Figure 10 we show $\widehat{p}_k$ as well as the 5th and 95th percentiles arising from 200 simulations with the full model, each starting out from A_8 and using the estimated dynamical parameters. We see that the calibrated model provides an ensemble of simulations about the actual data.

Now let us examine the performance of the model by considering the evolution of the average Watts-Strogatz clustering coefficient, C_k, from week to week. For an Erdös-Rényi graph $C_k = p_k$ in expectation. In Figure 11 we see that the model performs extremely

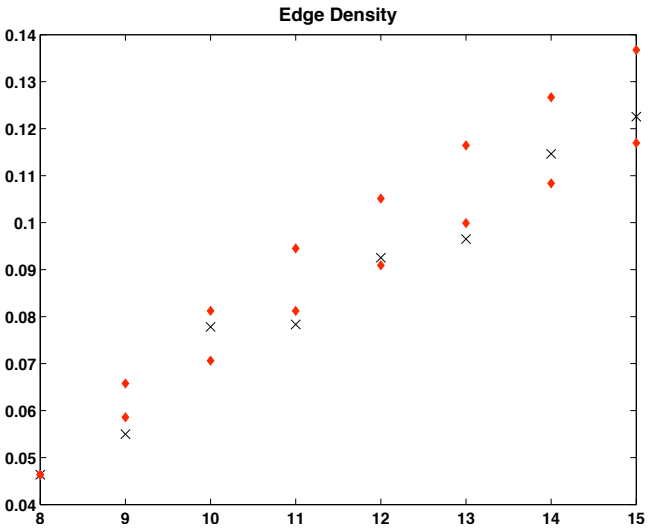

Figure 10: Evolution of edge density by week from data (crosses), and also the 5 and 95 percentile values (triangles) achieved via an ensemble of model simulations each using the fitted dynamical parameter values, and each starting from A_8.

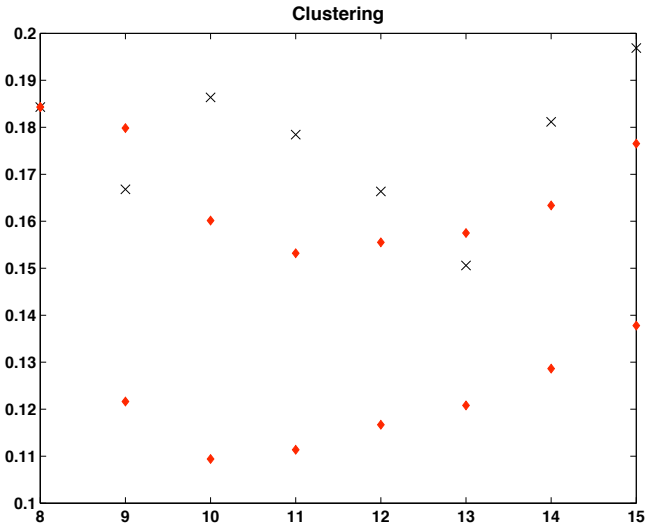

Figure 11: Evolution of C_k by week from data (crosses), and also the the 5 and 95 percentile values (triangles) achieved via an ensemble of model simulations each using the fitted dynamical parameter values, and each starting from A_8.

poorly since the observed values for C_k are much higher than those achievable under the conditioned Erdös-Rényi mean field evolution.

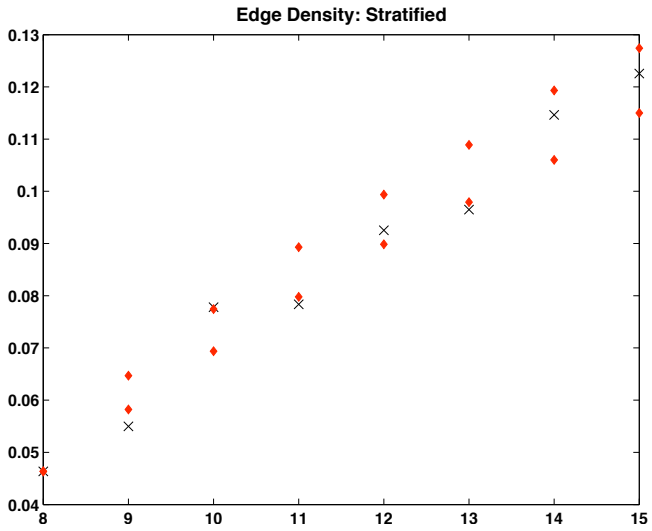

Figure 12: Evolution of edge density by week from data (crosses), and also the 5 and 95 percentile values (triangles) achieved via an ensemble of model simulations each using the fitted dynamical parameter values, and each starting from A_8.

One explanation for this poor fit is that the population is not homogeneous, which motivates us to consider a stratified random graph model of the type introduced in section 3, in order to increase the clustering within some subgroup, and thus increase the values for C_k overall, while keeping the overall edge density low. We partitioned the vertices into two sets; one of size $n_1 = 92$, and one of size $n_2 = 14$. This was done by first clustering the vertices using the sum of the adjacency matrices as a similarity matrix and adopting a Fiedler vector approach to give a spectral clustering [12, 26]. In other applications this task might be done a priori on grounds such as gender, functional role or responsibility of the individuals.

Assuming an identical birth and death dynamic for edges within each vertex subset and between the subsets, then, in the (p, q, r) notation of section 3, the mean field dynamic at the $(k+1)$th time step becomes

$$\begin{aligned} p_{k+1} &= (1-\widetilde{\omega}_p)p_k + (1-p_k)(\delta_p + \epsilon_p((n_1-2)p_k^2 + n_2 r_k^2)), \\ r_{k+1} &= (1-\widetilde{\omega}_r)r_k + (1-r_k)(\delta_r + \epsilon_r((n_1-1)p_k r_k + (n_2-1)q_k r_k)), \\ q_{k+1} &= (1-\widetilde{\omega}_q)q_k + (1-q_k)(\delta_q + \epsilon_q((n_2-2)q_k^2 + n_1 r^2)), \end{aligned}$$

involving nine parameters. These nine degrees of freedom can be fitted using the estimates for (p_k, q_k, r_k). We obtain the overall edge density evolution shown in Figure 12.

Again let us examine the performance of the stratified model by considering the evolution of the Watts-Strogatz clustering coefficient, C_k. In Figure 13 we see that the model performs far better than the unstratified version. This indicates that there well may be some hierarchical stratification (such as the one proposed here on the basis of a priori clustering) amongst the subjects observed in the data set.

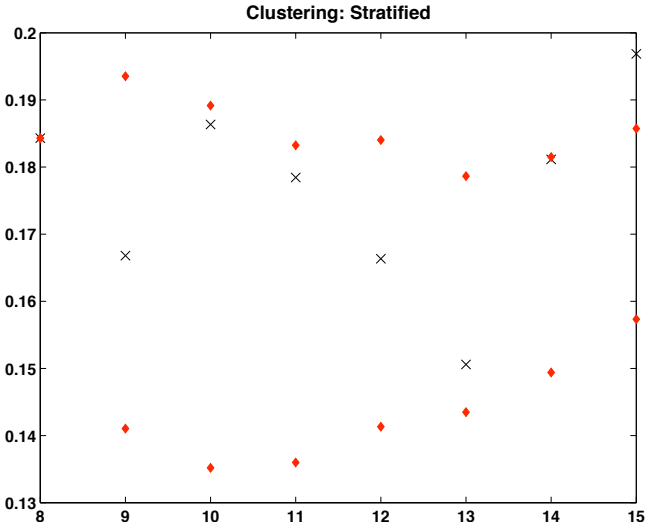

Figure 13: Evolution of C_k by week from data (crosses), and also the 5 and 95 percentile values (triangles) achieved via an ensemble of stratified model simulations each using the fitted dynamical parameter values, and each starting from A_8.

7 Discussion

The motivation for this work was to develop a framework for modelling and analysing dynamic connectivity structures. A successful model offers the potential to illustrate the range of possible behaviors and also to allow predictions under various 'what if' scenarios, such as spiking information (rumours, marketing, stimulus) or making direct perturbations (disabling or enhancing specific vertices or edges).

By introducing the concept of conjugate graphs and developing a mean field theory, we opened up the potential to approximate and calibrate the dynamics of observed networks. In particular, this can help us to identify when the network is close to an unstable rest point, where stochastic details will be important.

We illustrated the ideas on a simple but realistic nonlinear model based on a small set of parameters that is amenable both to analysis and to calibration. In this model, new edges are more likely to appear through successive triangulation (in addition to random births and deaths). This is a natural dynamic that is likely to enhance communication efficiency and resilience. However, our analysis showed that given a lack of early stimulus the networks will remain at a relatively sparse state—a sufficiently dense initial state is needed to produce well-clustered long time networks. There is, of course, great scope

for many other classes of evolving network mechanisms to be proposed, analysed and calibrated using the ideas and tools developed here.

Acknowledgment

The authors would like to thank the Engineering and Physical Sciences Research Council and the Research Councils UK Digital Economy programme for support through the MOLTEN (Mathematics Of Large Technological Evolving Networks) project, EP/I016058/1 and HORIZON, EP/G065802/1.

References

[1] C. Avin, M. Koucký, and Z. Lotker, *How to explore a fast-changing world (cover time of a simple random walk on evolving graphs)*, in ICALP '08: Proceedings of the 35th international colloquium on Automata, Languages and Programming, Part I, Berlin, Heidelberg, 2008, Springer-Verlag, pp. 121–132.

[2] A.-L. Barabási, *The origin of bursts and heavy tails in human dynamics*, Nature, 435 (2005), pp. 207–211.

[3] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science, 286 (1999), pp. 509–12.

[4] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, *Dynamic reconfiguration of human brain networks during learning*, Proc. Nat. Acad. Sci., 108 (2011), p. doi: 10.1073/pnas.1018985108.

[5] K. Berman, *Vulnerability of scheduled networks and a generalization of Menger's Theorem*, Networks, 28 (1996), pp. 125–134.

[6] E. Bullmore and O. Sporns, *Complex brain networks: graph theoretical analysis of structural and functional systems*, Nature Reviews Neuroscience, 10 (2009), pp. 186–198.

[7] A. E. Clementi, C. Macci, A. Monti, F. Pasquale, and R. Silvestri, *Flooding time in edge-Markovian dynamic graphs*, in Proceedings of the 27th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC'08), ACM Press, 2008, pp. 213–222.

[8] A. E. F. Clementi, F. Pasquale, A. Monti, and R. Silvestri, *Information spreading in stationary Markovian evolving graphs*, in Proceedings of the 2009 IEEE International Symposium on Parallel & Distributed Processing, Washington, DC, USA, 2009, IEEE Computer Society, pp. 1–12.

[9] N. Eagle, A. S. Pentland, and D. Lazer, *Inferring friendship network structure by using mobile phone data*, Proceedings of the National Academy of Sciences, 106 (2009), pp. 15274–15278.

[10] P. Esfandiar, F. Bonchi, D. Gleich, C. Greif, L. Lakshmanan, and B.-W. On, *Fast Katz and commuters: Efficient estimation of social relatedness in large networks*, in Algorithms and Models for the Web-Graph, R. Kumar and D. Sivakumar, eds., vol. 6516 of Lecture Notes in Computer Science, Springer Berlin/Heidelberg, 2010, pp. 132–145.

[11] S. Goodreau, J. A. Kitts, and M. Morris, *Birds of a feather or friend of a friend? Using exponential random graph models to investigate adolescent friendship networks*, Demography, 46 (2009), pp. 103–126.

[12] P. Grindrod and D. J. Higham, *Evolving graphs: Dynamical models, inverse problems and propagation*, Proceedings of the Royal Society, Series A, 466 (2010), pp. 753–770.

[13] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, *Communicability across evolving networks*, Phys. Rev. E, 83 (2011), p. 046120.

[14] T. Gross and B. Blasius, *Adaptive coevolutionary networks: a review*, J. Royal Society Interface, 5 (2008), pp. 259–71.

[15] Y. He, Z. J. Chen, and A. C. Evans, *Small-world anatomical networks in the human brain revealed by cortical thickness from MRI*, Cerebral Cortex, 17 (2007), pp. 2407–2419.

[16] P. Holme, *Network reachability of real-world contact sequences*, Physical Review E, 71 (2005).

[17] P. Holme and J. Saramäki, *Temporal Networks*, ArXiv e-prints, (2011).

[18] G. Kossinets, J. Kleinberg, and D. Watts, *The structure of information pathways in a social communication network*, in Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, KDD '08, New York, NY, USA, 2008, ACM, pp. 435–443.

[19] G. Kossinets and D. J. Watts, *Empirical analysis of an evolving social network*, Science, 311 (2006), pp. 88–90.

[20] D. Liben-Nowell and J. Kleinberg, *The link-prediction problem for social networks*, Journal of the American Society for Information Science and Technology, 58 (2007), pp. 1019–1031.

[21] Y. Liu, M. Liang, Y. Zhou, Y. He, Y. Hao, M. Song, C. Yu, H. Liu, Z. Liu, and T. Jiang, *Disrupted small-world networks in schizophrenia*, Brain, 131 (2008), pp. 945–961.

[22] Z. Lu, B. Savas, W. Tang, and I. Dhillon, *Supervised link prediction using multiple sources*, in Data Mining (ICDM), 2010 IEEE 10th International Conference on, Dec. 2010, pp. 923 –928.

[23] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, *Community structure in time-dependent, multiscale, and multiplex networks*, Science, 328 (2010), pp. 876–878.

[24] M. E. J. Newman, *Networks: An Introduction*, Oxford Univerity Press, Oxford, 2010.

[25] M. E. J. Newman, C. Moore, and D. J. Watts, *Mean-field solution of the small-world network model*, Physical Review Letters, 84 (2000), pp. 3201–3204.

[26] G. Strang, *Computational Science and Engineering*, Wellesley-Cambridge Press, 2008.

[27] M. Szell and S. Thurner, *Measuring social dynamics in a massive multiplayer online game*, Social Networks, 39 (2010), pp. 313–329.

[28] J. Tang, M. Musolesi, C. Mascolo, V. Latora, and V. Nicosia, *Analysing information flows and key mediators through temporal centrality metrics*, in SNS '10: Proceedings of the 3rd Workshop on Social Network Systems, New York, NY, USA, 2010, ACM, pp. 1–6.

[29] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora, *Small-world behavior in time-varying graphs*, Physical Review E, 81 (2010), p. 05510.

[30] A. Wagner, *How the global structure of protein interaction networks evolves*, Proceedings of The Royal Society of London. Series B, Biological Sciences, 270 (2003), pp. 457–466.

[31] D. J. Watts and S. H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature, 393 (1998), pp. 440–442.

[32] K. Zhao, J. Stehlé, G. Bianconi, and A. Barrat, *Social network dynamics of face-to-face interactions*, Physical Review E, 83 (2011), pp. 056109+.