Simple graph models of information spread in finite populations

Burton Voorhees¹ and Ryder Bergerud²

¹Center for Science, Athabasca University, 1 University Drive, Athabasca, Alberta, Canada T9S 3A3
²Department of Mathematics, University of Victoria, Victoria, British Columbia, Canada

1. Summary

We consider several classes of simple graphs as potential models for information diffusion in a structured population. These include biased cycles, dual circular flows, partial bipartite graphs and what we call ‘single-link’ graphs. In addition to fixation probabilities, we study structure parameters for these graphs, including eigenvalues of the Laplacian, conductances, communicability and expected hitting times. In several cases, values of these parameters are related, most strongly so for partial bipartite graphs. A measure of directional bias in cycles and circular flows arises from the non-zero eigenvalues of the antisymmetric part of the Laplacian and another measure is found for cycles as the value of the transition probability for which hitting times going in either direction of the cycle are equal. A generalization of circular flow graphs is used to illustrate the possibility of tuning edge weights to match pre-specified values for graph parameters; in particular, we show that generalizations of circular flows can be tuned to have fixation probabilities equal to the Moran probability for a complete graph by tuning vertex temperature profiles. Finally, single-link graphs are introduced as an example of a graph involving a bottleneck in the connection between two components and these are compared to the partial bipartite graphs.

2. Introduction

Using graphs to represent structured populations has become an important tool in approaching a variety of questions relating to the way that information spreads in non-homogeneous populations. Questions that have been addressed include the extinction or fixation of mutant genes [1–4]; epidemics [5–7]; spread of computer viruses, rumours and gossip [8–11]; development of rumour spreading algorithms for synchronization processes in parallel computation [12–14]; uptake of innovations and new ideas [15]; locating the source of rumours or viruses (natural and computer) [16,17]; and tracking terrorists [18,19].
A population is represented as a directed graph, each vertex corresponding to an individual population member, while edges are labelled with information describing the interaction between population members. In many cases, the edge labels are probabilities so that the weight \( w_{ij} \) of the edge that connects vertex \( i \) to vertex \( j \) describes the probability of an effect being propagated from vertex \( i \) to vertex \( j \).

Following this, an update procedure is defined describing temporal dynamics. Here attention is restricted to discrete processes but continuous models have also been explored [20]. Some of the useful update procedures are birth–death, death–birth, voter models and probabilistic voter models [21]. In this paper, birth–death updating is used. In cases of rumour spread, there are competing narratives and birth–death processes provide a good model.

The state space approach developed in Voorhees [22] and Voorhees & Murray [23] allows determination of all fixation probabilities for arbitrary initial conditions. The population is partitioned into two disjoint classes, depending on the presence or absence of a defining characteristic (e.g. a mutant gene, believing a rumour, being infected with a virus, etc.). Vertices are labelled 0 if this characteristic is not present and 1 if it is present. Thus, for a population of fixed size \( N \) the state space is the set of \( 2^N \) binary vectors of length \( N \) with all zero vectors corresponding to extinction and all one vector to fixation. A state vector \( v = (v_1, \ldots, v_N) \) is also a binary number. Taking \( v \) as the denary form of this number, \( v = \sum_{k=1}^{N} v_k 2^{N-k} \), and denoting the probability that the state denoted \( v \) will go to fixation by \( x_{v} \), a set of master equations were derived [22], yielding fixation probabilities for all initial states.

\[
[N + (r - 1)m - ra(v) \cdot v - b(v) \cdot v']x_v - r \sum_{i=1}^{N} a_i(v) v_i v_{x+v+2^{N-1}} - r \sum_{i=1}^{N} b_i(v) v_i v_{x-2^{N-1}} = 0, \tag{2.1}
\]

where \( W \) is the edge weight matrix of the population graph, \( v'_j = 1 - v_j \) and \( a(v) = v \cdot W, b(v) = v' \cdot W \).

This yields interesting results [22,23] but is limited in that equations (2.1) are a set of \( 2^N - 2 \) linear equations in an equal number of unknowns so that exact solution for populations having more than a few members is not possible unless strong symmetry constraints are imposed.

As far as general solutions for fixation probabilities are concerned, Moran [24] derived the single vertex fixation probability for a complete graph on \( N \) vertices as

\[
\rho_M = \frac{1 - 1/r}{1 - 1/r^N} = \frac{r^{N-1}}{1 + r + \ldots + r^{N-1}}, \tag{2.2}
\]

and the \( k \)-vertex fixation probability for the complete graph is

\[
\rho_{M,k} = \frac{r^{N-1} + \ldots + r^{N-k}}{1 + r + \ldots + r^{N-1}}. \tag{2.3}
\]

A graph is fixation equivalent to a complete graph if its single vertex fixation probability is given by the Moran expression of equation (2.2). Lieberman et al. [4] proved that a graph with a probabilistic edge weight matrix is fixation equivalent to a complete graph if and only if the edge weight matrix is doubly stochastic. Broom & Rychtář [25] derived the single vertex fixation probability for the \( n \)-star graph and also provided a means to compute this probability for a line graph. Zhang et al. [26] derived the \( k \)-vertex fixation probability for star graphs, and Voorhees & Murray [23] give the single vertex fixation probability for the complete bipartite graph \( K_{N,N} \) together with summation formulae allowing computation of \( k \)-vertex fixation probabilities for these graphs. The star and bipartite results are reproduced in simple and compact form in Monk et al. [27] using a Martingale formalism, while results reported in Adlam & Nowak [28] indicate that the Moran formula of equation (2.1) may be, at least asymptotically, universal in the family of random graphs. In this paper, several classes of simple graphs are considered with respect to fixation probability and characteristic measures of graph structure.

Because applied interest is in large populations, a good deal of research is focused on random graphs, complete graphs and small world graphs with vertex degrees specified in terms of global averages and treated statistically [29–32]. Our goal is more modest, to provide results in some very simple cases that may point to more general applications, and to examine the extent to which structural parameters of graphs may provide information related to fixation probabilities and other questions of interest.

Graph parameters considered are eigenvalues of the graph Laplacian, subset conductances, graph communicability (the Estrada parameter) and measures of vertex centralities, and random walk hitting
times. Choice of graphs for study has been both pragmatic and strategic: we have chosen graphs that are
simple to analyse, yet exhibit behaviours suggestive for more general cases.

The graphs considered are biased cycles, dual circular flows, partial bipartite graphs and single-
link coupled graphs. While cycles are perhaps the simplest graphs to consider, biased cycles display
interesting properties, e.g. providing a clear case of asymmetry in the graph Laplacian that shows
an analogy to rotation/angular momentum. In addition, results obtained for cycles may be useful for
analysis of more complex graphs in which interaction cycles can be identified, the simplest case being
dual circular flows, which can be thought of as composed of an ensemble of cycles. Dual circular flows,
in turn, constitute a large family of graphs that include star graphs \[4,22,25\], funnel graphs \[4,22,23\],
cascades \[23\] and layered networks \[33\]. In addition, we use them to provide an example of how the edge
weights of a graph can be tuned to match specified global properties such as temperature profile. Finally,
the partial bipartite and single-link graphs form the two extremes of a class of graphs with vertex set the
union of two components \(V_s\) and \(V_n\) with linkages between these components through subsets \(S_k \subseteq V_s\)
and \(S_m \subseteq V_n\) having, respectively, \(k\) and \(m\) elements. In terms of application, this family of graphs can
provide models for polarized populations in which communication between different subpopulations
takes place through subsets of representatives.

While illustrated here for only a small numbers of vertices, many of our results hold for graphs of
arbitrary size. In particular, theorems (4.1), (4.3), (5.1) and (6.1) giving Laplacian eigenvalues for cycles,
dual circular flows and partial bipartite graphs; equation 7.1, giving the characteristic polynomial for
the Laplacian of a single-link graph; conjectures (6.2) and (6.3) on hitting times and communicability for
partial bipartite graphs; and formulae derived for conductances are valid for any number of vertices.

Partial bipartite graphs are of particular interest. These graphs have only three non-zero Laplacian
eigenvalues and which of these is smallest or largest depends on the values of the connection parameters
\(1 - p\) and \(1 - q\) between the two subsets of vertices. Equality of all three eigenvalues occurs at the value of \(p\)
and \(q\) for which the conductances of each graph component are equal and, although no proof is
available, this appears to correspond to the condition for equality of single vertex fixation probabilities.
Further, given conjecture (6.3), these parameter values give the minimum value of the communicability
and are the condition for equal probability measures on paths going in either direction between graph
components.

The notation used is that vertices are labelled with lower case indices ranging over the number of
vertices. Thus, \(v_k\) labels the \(k\)th vertex of a graph. The single vertex fixation probability for the
\(k\)th vertex is labelled by \(x_{2^{k-1}}\) with \(2^{k-1}\) the base 10 form of the binary number associated with a single one at the
\(k\)th vertex.

3. Measures of graph characteristics

This section briefly describes the structural measures for graphs that are considered in this paper.

3.1. Eigenvalues of the Laplacian

The graph Laplacian of a graph \(G\) with edge weight matrix \(W\) is defined in Barbosa et al. [34] as
\[ \Delta = \Phi^{1/2}(I - W)\Phi^{-1/2}, \]
where \(\Phi\) is a diagonal matrix with entries given by the solution of \(\phi \cdot W = \phi\). The
smallest eigenvalue of \(\Delta\) is always 0 and the eigen-spectrum satisfies \(\lambda_0 = 0 \leq \lambda_1 \leq \cdots \leq \lambda_{N-1} \leq 2\). The
number of zero eigenvalues gives the number of disconnected components of the graph. Thus, if \(\lambda_1\) is
non-zero the graph is connected and the value of \(\lambda_1\) gives a measure of the difficulty involved in cutting
the graph into disconnected parts.

3.2. Subset conductances

For a graph \((G, W)\) with vertex set \(V\) the conductance, also known as the Cheeger constant [35], is defined as
\[ C_{G,W} = \min_{S \subseteq V, S \neq \emptyset} \left( \frac{N}{|S||G\setminus S|} \sum_{\substack{i \in S \atop j \in G \setminus S}} w_{ij} \right). \]
For a subset of vertices S, the conductance of S is

\[ C_{G,W}(S) = \frac{N}{|S||G\setminus S|} \sum_{i \in S} \sum_{j \in G\setminus S} w_{ij}. \] (3.2)

The conductance of S is inversely related to the mixing time for a Markov process starting at a vertex in S.

### 3.3. Communicability

This concept was introduced in 2008 by Estrada & Hatano [36] in order to describe situations ‘in which a perturbation on one node of [a] network is ‘felt’ by the rest of the nodes with different intensities.’ [37, p. 92]. The communicability between vertices i and j of a graph is defined as a weighted sum of all paths starting at vertex i and ending at vertex j, with weighting that favours shorter paths. This allows definition of a variety of communicability functions with the general form 

\[ P = \sum_{k=0}^{\infty} c_k A^k, \]

where A is the graph adjacency matrix and the coefficients \( c_k \) are chosen such that the series converges, gives greater weight to shorter paths (smaller k values), and yields real, non-negative values for the matrix elements of P. As sums over loops of all lengths starting and ending at a vertex, the diagonal elements of P represent a measure of vertex centrality [37].

One choice for coefficients is \( c_k = 1/k! \), yielding \( P = e^A \), and the communicability is defined as \( \text{Tr}(e^A/N) \), i.e. the average vertex centrality. Since cases considered in this paper involve probabilistically weighted edges, the matrix A is replaced by the edge weight matrix W. Hence the matrix function used is \( V = e^W \) and the i,j element of this matrix is the weighted sum of probabilities for paths of length k starting at i and ending at j, for \( 0 \leq k \leq \infty \). We still refer to \( \text{Tr}(e^W/N) \) as communicability, however. In most cases, \( e^W \) cannot be computed directly and the sixth-order approximation \( e^W \approx I + W + 1/2W^2 + 1/6W^3 + 1/24W^4 + 1/120W^5 + 1/720W^6 \) is used.

### 3.4. Expected hitting times

Following Kemeny & Snell [38], the matrix M of expected hitting times for a graph (G, W) is computed using the fundamental matrix of the graph, defined by \( Z = (I - (W - A))^{-1} \) where W is the edge weight matrix and A is the matrix with columns equal to the stationary probability of a random walk on G. If J is the matrix of all ones and \( Z^* \) is the diagonal matrix with \( Z_{ii}^* = Z_{ii} \), then [38]

\[ M = (I - Z + JZ^*)D. \] (3.3)

The \( h_{ij} \) element of M is the expected value of the random variable describing the number of iterations for a random walk that starts at vertex i to first reach vertex j.

### 4. Biased cycles

This section begins with consideration of biased cycles, that is, cycles in which there is a probabilistically preferred direction.

Figure 1 shows the general form of a biased cycle. If the parameter p equals one-half it becomes a balanced cycle, while \( p = 1 \) yields a clockwise cycle and \( p = 0 \) yields a counter-clockwise cycle. The weight matrix W is doubly stochastic; hence, by the Isothermal theorem, the single vertex fixation probability equals the Moran probability, although time to fixation will differ [39].

From equation (3.2), it is easy to see that the conductance of any single connected block of k vertices in a length N biased cycle equals \( N/k(N-k) \). Also, the matrix \( e^W \) is circulant: \( e^W = \text{circ}(V_{11}, V_{12}, V_{13}, V_{14}, V_{15}) \). Since the diagonal elements of \( W^k \) are zero if k is odd and symmetric in \( p \) and \( 1-p \) if k is even, \( \text{Tr}(e^W) \) will be symmetric in \( p \) and \( 1-p \) with maximum value at \( p = 1/2 \). Figure 2 shows entries of \( e^W / 5 \) for \( N = 5 \).

This figure is symmetric about \( p = 1/2 \). \( V_{ii} = V_{jj} \) for all \( i,j \), hence \( \text{Tr}(e^W/5) = V_{11} \) and \( V_{12} = V_{15} \), \( V_{13} = V_{14} \) at 1/2. The diagonal elements \( V_{11} \) are substantially greater than the off diagonal elements with the exceptions \( V_{15} = V_{11} \) at \( p = 0 \) and \( V_{12} = V_{11} \) at \( p = 1 \).
Figure 1. Counter-clockwise probabilities are $p$ and clockwise probabilities are $1 - p$.

Figure 2. Entries in $eW = \text{Circ}(V_{11}, V_{12}, V_{13}, V_{14}, V_{15})$.

Since the steady state for the matrix $W$ is homogeneous, the Laplacian matrix for a biased cycle is $\Delta = I - W$. Surprisingly, the characteristic polynomial for $\Delta$ is rather complicated. For the biased cycle on $N$ vertices,

$$w_{ij} = \begin{cases} 
  p & j = i + 1 \mod (N) \\
  1 - p & j = i - 1 \mod (N) \\
  0 & \text{otherwise.}
\end{cases} \quad (4.1)$$

Proof of theorems (4.1) and (4.3) is tedious but straight forward.

Theorem 4.1. Let $C(N, p)$ be a biased cycle on $N$ vertices. The characteristic polynomial for $\Delta(C(N, p))$ is

$$\sum_{k=0}^{(N-1)/2} (-1)^k \binom{N}{k} p^k (1-p)^k (\lambda - 1)^{N-2k} + p^N + (1-p)^N \quad N \text{ odd}$$

$$\sum_{k=0}^{(N-2)/2} (-1)^k \binom{N}{k} p^k (1-p)^k (\lambda - 1)^{N-2k} - K_N \quad N \text{ even}$$
Circular flows, that is, circular flows in which directed edges exist in both directions. The fitness parameter vertex fixation probability is enhanced with respect to the Moran probability only for limited values of computations and demonstration that this family of graphs contains many cases in which the single illustrated schematically in figure 4.

In Voorhees and Murray a number of results concerning circular flows are given, including entropy fixation probability for the state arising from 5. Dual circular flows

The number of mutants in the fixation probability of the state arising from

Theorem 4.3. Let \(\Delta_A\) be the antisymmetric part of \(\Delta\). For a biased cycle, all eigenvalues of \(\Delta_A\) are either zero or are imaginary and contain a factor of \(1 - 2p\).

If \(1 - 2p\) is negative then there is a clockwise bias in the cycle, whereas if it is positive there is a counterclockwise bias. This bias shows up clearly in the hitting times. Figure 3a – c shows hitting times for the biased cycle as a function of \(p\) for one-, two- and three-step transitions in the counter-clockwise direction. The shape of these curves is related to the possibility of a clockwise transition of \(k\) steps or a clockwise transition of \(N - k\) steps. Figure 3d shows the one-, two-, three- and four-step hitting times in the counter-clockwise direction for the \(N = 9\) biased cycle.

The value of \(p\) at which the peak hitting time occurs gives a measure of the bias of the cycle. For example, for \(N = 5\) the single-step transition in the counter-clockwise direction is favoured for \(p \lesssim 0.68806\), whereas for \(p > 0.68806\) the four-step transition in the clockwise direction is favoured. Because of the symmetry between \(p\) and \(1 - p\), the peak value of \(p\) starting from the symmetrically placed vertex will be one minus the peak value starting from a given vertex. Referring to figure 1, if vertex \(v_1\) is located at the top of a five cycle then the \(p\) value of 0.68806 characterizes vertex \(v_2\) while vertex \(v_5\) will be characterized by a peak \(p\)-value of 0.31194. Likewise, the two- and three-step graphs for \(N = 5\) are mirror images reflected about \(p = 1/2\).

5. Dual circular flows

In Voorhees and Murray, a class of graphs called circular flows is defined and in Voorhees and Murray a number of results concerning circular flows are given, including entropy computations and demonstration that this family of graphs contains many cases in which the single vertex fixation probability is enhanced with respect to the Moran probability only for limited values of the fitness parameter \(r \geq 1\). In these cases, graph edges were directed in a single direction and the graphs were viewed in terms of probability flow through a series of layers. Here this is generalized to dual circular flows, that is, circular flows in which directed edges exist in both directions.

For simplicity, only the case of uniform right- and left-directed edge weights is considered although this condition will be dropped when we consider tuning a graph. A length \(k + 1\) dual circular flow is illustrated schematically in figure 4.

The state space for this system is the set of vectors \((m_0, m_1, \ldots, m_k)\) where \(m_i\) indicates the number of mutants in the \(i\)th subset of the graph, which contains \(n_i\) vertices. Using the notation of Voorhees, \(x_i(m)\) indicates the fixation probability of the state \(m\), whereas \(x_i(m_0, m_i)\) represents the fixation probability for the state arising from \(m_0\) in which the \(i\)th population of mutants has increased (+) or decreased (–) by 1. For a dual circular flow, equation (1.1) becomes

\[
\begin{aligned}
&\sum_{s=0}^{k} \left( n_{s+1} - m_{s+1} \right) p + \left( n_{s} - m_{s} \right) (1 - p) \right) x_i(m) + r \sum_{s=0}^{k} \left( n_{s} - m_{s} \right) \left( \frac{m_{s+1} + m_{s} (1 - p)}{n_s} \right) x_i(m, m_{s+1}) \\
&- \sum_{s=0}^{k} \left( n_{s+1} - m_{s+1} \right) p + \left( n_{s} - m_{s} \right) (1 - p) \right) x_i(m, m_{s-1}) \\
&- r \sum_{s=0}^{k} \left( n_{s} - m_{s} \right) \left( \frac{m_{s+1} + m_{s} (1 - p)}{n_s} \right) x_i(m, m_{s+1}) = 0.
\end{aligned}
\]  

(All sums taken \(\mod(k + 1)\).)

Figure 5 shows single vertex fixation probabilities for the \((n_0, n_1, n_2) = (1, 2, 3)\) dual circular flow. The \(p = 0\) case corresponds to the \((1, 2, 3)\) cascade graph and the \(p = 1\) case corresponds to the \((1, 2, 3)\) funnel graph. The fixation probability for the single vertex at level 0 is \(x_1\), for each of the two vertices at level 1
it is $x_2$, and for each of the three vertices at level 2 it is $x_8$. The overall single vertex fixation probability is $\rho = (x_1 + 2x_2 + 3x_3)/6$ and the Moran probability is denoted $\rho_m$.

Figure 5a shows the transition between fixation probabilities for cascade and funnel flows as $p$ varies from 0 to 1. Figure 5b–d shows the differences between single vertex fixation probabilities from different levels (0, 1, 2). Figure 5e shows cross sections of figure 5a for specified values of $r$, whereas figure 5f–h shows plots of $x_1 - \rho_m$, $x_2 - \rho_m$ and $x_8 - \rho_m$, respectively, as functions of $r$ for specified $p$-values.

Examination of these figures indicates that starting at the single, level zero vertex always suppresses selection, starting at one of the three level-two vertices always enhances selection, relative to the Moran value, while starting at one of the two level-one vertices will enhance or suppress selection depending on the values of $p$ and $r$ (for example, for $p = 0.4$ selection is enhanced for $r > 1$). At $p = 0.25$, $x_2 > x_8$ for $r < 4.0471$ while for $r$ values greater than this $x_8 > x_2$. Likewise, at $p = 0.9$, $x_2 < x_1$ for $r < 2.9803$ and the inequality is reversed for $r$ greater than this value. Thus, the two vertices at level 1 exhibit the least stability in terms of fixation behaviour.

This example illustrates the complex behaviour that appears even in simple dual circular flow graphs. In Voorhees & Murray [23], a number of cases are considered showing that the single vertex fixation probability for such graphs can display both enhancement and suppression of selection, depending on the particular value of the fitness $r$. For example, in the (1, 2, 3)-funnel graph ($p = 1$), selection is only enhanced for $1 < r < 5.3695$, whereas for the (1, 2, 3)-cascade ($p = 0$), it is only enhanced for $1 < r < 2.0304$.

The following theorem relates the characteristic polynomials for the Laplacian of all dual circular flow graphs to that of a biased cycle of the same length.

Figure 3. Hitting times for single vertex transition in counter-clockwise direction as function of $p$ for $N = 5–9$ (from bottom). (a) single-step, (b) double-step, (c) three-step, (d) $N = 9$ case for one through four-step transitions (bottom to top).
Figure 4. Schematic of a $k+1$ level dual circular flow. If $p = 0$ or $1$ this reduces to the simple homogeneous circular flow defined in Voorhees [22] and Voorhees & Murray [23].

**Theorem 5.1.** Let $(n_0, n_1, \ldots, n_k)$ represent a dual circular flow with $N = \sum_{i=0}^{k} n_i$. Then the characteristic polynomials for $\Delta$ and $\Delta_A$ are just $(\lambda - 1)^{N-k-1}$ multiplying the corresponding characteristic polynomials of a length $k$ biased cycle with bias $p$.

**Outline of proof**

Beginning with the flow $(1, 1, 2)$, the claim is demonstrated for this case. This is extended by induction to the flows $(1, 1, n)$ and $(1, \ldots, 1, n)$. With this established a further induction demonstrates the result for flows $(1, \ldots, 1, n_{k-1}, n_k)$. Proceeding in this way, an eventual proof is obtained.

One point to note is that all eigenvalues of $\Delta_A$ again involve a factor of $1 - 2p$, providing a measure of directional bias in the circular flow.

By equation (3.2) and inspection of figure 4, it is obvious that the conductance of level $s$ in a dual circular flow equals $N/(N-n_s)$. Computation of the communicability and the associated vertex centralities yield results like those indicated in figure 6 which shows these plotted as functions of $p$ for the dual flows 123, 235 and 1234, with $V_i$ indicating the centrality of a vertex at level $i$.

Inspection of this figure indicates that the centrality of a vertex in a given layer is inversely related to the number of vertices in that layer—multiplying vertex centrality by the number of vertices in a layer shows that the centrality of a layer is proportional to the number of vertices—layers with more vertices have greater centrality. This can be understood in terms of the topological structure of these graphs—there are no edges connecting vertices within a layer. This also implies that the hitting times between layers will be the same as for a biased cycle of the same length, although hitting times between specific vertices will be substantially longer.

Dual circular flows also provide an example of the way that edge coefficients in a graph can be tuned to match desired global features. To address this, the graph of figure 4 is generalized, allowing different weight coefficients between levels. This is indicated in figure 7. The target parameter chosen for tuning is vertex temperature [4] at each level of the graph and the inter-level edge weights $p_i$ will be selected to produce specified temperature profiles. In particular, conditions will be determined for which the single vertex fixation probability matches that of a complete graph.

Referring to figure 7, the temperature of a vertex at level $i$ is

$$t_i = \frac{n_{i+1}p_{i+1} + n_{i-1}(1 - p_{i-1})}{n_i}. \quad (5.2)$$

Hence, if a temperature profile $t = (t_0, \ldots, t_k)$ is to be matched the following set of equations must be satisfied:

$$n_{i+1}p_{i+1} - n_{i-1}p_{i-1} = t_i n_i - t_{i-1}. \quad (5.3)$$

Two examples will illustrate the factors involved in doing this.
Figure 5. Fixation probabilities for (1, 2, 3) dual circular flow. (a) Single vertex fixation probability, (b) fixation probabilities $x_1$ minus $x_2$, (c) fixation probabilities $x_1$ minus $x_8$, (d) fixation probabilities $x_2$ minus $x_8$, (e) single vertex fixation probability as a function of $p$ for specified $r$ values and (f–h) fixation probabilities $x_1$, $x_2$ and $x_8$ for specified $p$-values.
Figure 6. Communicability ($V$, in red) and associated vertex centralities (black) for dual circular flows: (a) (1, 2, 3) flow (note the communicability equals the $v_1$ vertex centrality), (b) (2, 3, 5) flow and (c) (1, 2, 3, 4) flow.

Figure 7. Schematic of a general $k + 1$ level dual circular flow.

Example 5.2. A three-level flow $(n_0, n_1, n_2)$ with $t_i = 1$ for all $i$.

In this case, the goal is to adjust the connection coefficients $p_i$ so as to produce a graph with single vertex fixation probability equal to the Moran probability. Equation (5.3) becomes

\[
\begin{align*}
  n_1p_1 - n_2p_2 &= n_0 - n_2 \\
  n_2p_2 - n_0p_0 &= n_1 - n_0 \\
  n_0p_0 - n_1p_1 &= n_2 - n_1.
\end{align*}
\]

(5.4)

Setting $p_2 = p$ and solving for $p_0$ and $p_1$ yields

\[
\begin{align*}
  p_0 &= \frac{n_0 - n_1 + n_2p}{n_0} \quad \text{and} \quad p_1 = \frac{n_0 - n_2 + n_2p}{n_1}. \\
\end{align*}
\]

(5.5)

Since $p_0$ and $p_1$ are probabilities they must lie in the range $[0, 1]$, and this requires

\[
\begin{align*}
  \frac{n_1 - n_0}{n_2} &\leq p \leq \frac{n_1}{n_0} \quad \text{and} \quad \frac{n_2 - n_0}{n_2} &\leq p \leq \frac{n_1}{n_2}. \\
\end{align*}
\]

(5.6)

For example, if $(n_0, n_1, n_2) = (1, 2, 3)$, equation (5.6) requires $1/3 \leq p \leq 2$ and $2/3 \leq p \leq 2/3$. From the second of these, the only value allowed for $p = p_2$ is $2/3$ and this yields $p_0 = 1$, $p_1 = 0$. Thus, the (1, 2, 3) flow with edge weights $(p_0, p_1, p_2) = (1, 0, 2/3)$ has the Moran single vertex fixation probability.
If \((n_0, n_1, n_2) = (2, 3, 4)\), on the other hand, there is a larger range of possible edge weights: the conditions that must be satisfied are \(1/2 \leq p_2 \leq 3/4\) with \(p_0 = 2p_2 - 1/2\) and \(p_1 = 2(2p_2 - 1)\).

If the number of levels in the graph is even (i.e. \(k\) is odd), an additional constraint appears. Equation (5.3) separate into two sets of equations with the left sides of each set summing separately to zero. This yields the conditions

\[
\sum_{i=0}^{(k-1)/2} t_{2i}n_{2i} = \sum_{i=0}^{(k-1)/2} n_{2i+1} \quad \text{and} \quad \sum_{i=0}^{(k-1)/2} t_{2i+1}n_{2i+1} = \sum_{i=0}^{(k-1)/2} n_{2i}. \tag{5.7}
\]

For example, if \(k = 3\) and \(t_i = 1\) for all \(i\), this requires \(n_0 + n_2 = n_1 + n_3\). In addition, each set of equations is solved in terms of a single parameter so there are two parameters in the solution. For example, in the general \(k = 5\) case, the equations are

\[
\begin{align*}
n_2p_2 - n_0p_0 &= t_1n_1 - n_0 \\
n_4p_4 - n_2p_2 &= t_3n_3 - n_2 \\
n_0p_0 - n_4p_4 &= t_5n_5 - n_4
\end{align*}
\tag{5.8a}
\]

and

\[
\begin{align*}
n_1p_1 - n_5p_5 &= t_0n_0 - n_5 \\
n_3p_3 - n_1p_1 &= t_2n_2 - n_1 \\
n_5p_5 - n_3p_3 &= t_4n_4 - n_3
\end{align*}
\tag{5.8b}
\]

with constraints

\[
\begin{align*}
n_0 + n_2 + n_4 &= t_1n_1 + t_3n_3 + t_5n_5 \\
n_1 + n_3 + n_5 &= t_0n_0 + t_2n_2 + t_4n_4
\end{align*}
\tag{5.8c}
\]

Equations (5.8a, b) have solutions in terms of \(p_4\) and \(p_5\), respectively:

\[
\begin{align*}
p_0 &= \frac{t_5n_5 - n_4(1 - p_4)}{n_0}, & p_2 &= \frac{n_4p_4 + n_2 - t_3n_3}{n_2} \\
p_1 &= \frac{t_0n_0 - n_5(1 - p_5)}{n_1}, & p_3 &= \frac{n_5p_5 + n_3 - t_4n_4}{n_3}
\end{align*}
\tag{5.9}
\]

For a solution to be valid, all transition probabilities must be in \([0,1]\) imposing the additional constraints

\[
\begin{align*}
1 - \frac{t_5n_5}{n_4} &\leq p_4 \leq 1 + \frac{n_0 - t_5n_5}{n_4}, & \frac{t_3n_3 - n_2}{n_4} &\leq p_4 \leq \frac{t_3n_3}{n_4} \\
1 - \frac{t_0n_0}{n_5} &\leq p_5 \leq 1 + \frac{n_1 - t_0n_0}{n_5}, & \frac{t_4n_4 - n_3}{n_5} &\leq p_5 \leq \frac{t_4n_4}{n_5}
\end{align*}
\tag{5.10}
\]

while the values of \(p_4, p_5\) are also constrained to lie in \([0,1]\).

If edge weights exist satisfying the temperature profile \(t_i = 1, \, 0 \leq i \leq k\), then the graph with these edge weights will have the Moran fixation probability. The next example, however, demonstrates that matching a general temperature profile does not always lead to matching fixation probabilities.

**Example 5.3.** Consider the circular flow \((n_0, n_1, n_2)\) with temperature profile \((2, 3/2, 1/3)\). Any edge weights satisfying \(p_0 = 3p_2 - 2\), \(p_1 = (3p_2 - 1)/2\) will match this profile. If \(p_2 = 1\) then \(p_0 = p_1 = 1\) as well—this case is considered in Voorhees & Murray [23]. But all values of \(p_2\) in the interval \([2/3, 1]\) yield valid solutions matching the given profile. Figure 8 shows a plot of the single vertex fixation probability minus the Moran probability for this range of \(p\), illustrating that very different fixation probabilities result.
6. Partial bipartite graphs

The vertex set for a complete bipartite graph $K_{s,n}$ consists of two subsets $V_s$ and $V_n$ containing $s$ and $n$ vertices, respectively, with edge set $\{(u_i, v_j) | u_i \in V_s, v_j \in V_n\}$. That is, every vertex in $V_s$ has a link to every vertex in $V_n$ and vice versa but no pairs of edges in $V_s$ are linked nor are any pairs of edges in $V_n$. Furthermore, if edges are directed and weighted, then all weights for edges from $V_s$ to $V_n$ are $1/n$ and all weights for edges from $V_n$ to $V_s$ are $1/s$. This family of graphs includes the star graphs ($s = 1$).

In Voorhees & Murray [23], the birth–death single vertex fixation probability for the complete bipartite graph $K_{s,n}$ is given as

$$\rho_{K_{s,n}}(s,n) = \frac{r^s + s - 1}{s} \cdot \frac{(snr + n^2 - sn + s^2)(tr + s)^{n-s-1}}{P(s,n)}.$$  (6.1)

Here we consider the more general case of partial bipartite graphs, as illustrated in figure 9. This figure shows two subsets $V_s$ and $V_n$ of vertices with $|V_s| = s$ and $|V_n| = n$. Within each subset, every vertex connects equally to every other vertex with probabilities $p/(s-1)$ and $q/(n-1)$, respectively. In addition, every vertex in $V_s$ connects to every vertex of $V_n$ with the probability $(1-p)/n$ and every vertex of $V_n$ connects to every vertex of $V_s$ with probability $(1-q)/s$.

The edge weight matrix for this graph has block form:

$$W = \begin{pmatrix} 
\frac{p}{s-1}(J_{s,s} - I_s) & 1 - p & 0 \\
1 - q & s^{-1}I_{n,s} & \frac{q}{n-1}(J_{n,n} - I_n) \\
\end{pmatrix},$$  (6.2)

where $I_{i,j}$ is the $i$ by $j$ matrix of all ones and $I_j$ is the $j$ by $j$ identity matrix. The normalized steady-state solution for $\mathbf{v} \cdot W = \mathbf{v}$ is

$$\mathbf{v} = \frac{1}{sn(2-p-q)} (n(1-q), \ldots, n(1-q), s(1-p), \ldots, s(1-p)),$$  (6.3)

with $s$ entries of $n(1-q)$ and $n$ entries of $s(1-p)$. The Laplacian matrix and its antisymmetric part have block form

$$\Delta = I_N - \begin{pmatrix}
J_{s,s} - \frac{p}{s-1}(J_{s,s} - I_s) & \frac{(1-p)(1-q)}{n} J_{s,n} \\
\sqrt{(1-p)(1-q)} J_{n,s} & I_n - \frac{q}{n-1}(J_{n,n} - I_n) \\
\end{pmatrix}.$$
and $C$ is enhanced for $0 < q_n < 1$.

In the bipartite case, the subsets of interest are thus defined as $S = \{V_s, V_n\}$. The conductance of a set of vertices $S$ will be inversely related to the mixing time of a Markov process starting from a vertex in $S$. Thus, $C(V_s) \geq C(V_n)$ implies that the mixing time starting from a vertex in $V_s$ will be less than that starting from a vertex in $V_n$. The condition for $C(V_s) \geq C(V_n)$ is

$$1 - p \geq \frac{n}{s}(1 - q).$$

**Figure 9.** Schematic of a partial bipartite graph.

$$\Delta_A = \frac{(n-s)(1-p)(1-q)}{sn}\begin{pmatrix} 0 & \bar{s}_r \bar{n}_s \\ \bar{n}_s & 0 \end{pmatrix}. \quad (6.4)$$

Proof of the next theorem follows directly from the form of these matrices.

**Theorem 6.1.** The characteristic polynomials for $\Delta$ and $\Delta_A$ of the $N$ vertex partial bipartite graph with edge weight matrix $W$ given in equation (6.2) are

$$\lambda \left( \lambda - \frac{s - 1 + p}{s - 1} \right)^{s-1} \left( \lambda - \frac{n - 1 + q}{n - 1} \right)^{n-1} \left( \lambda - 2 + p + q \right)^{-n/2} \left[ \lambda^2 + \frac{(n-s)}{sn} (1-p)(1-q) \right]. \quad (6.5)$$

Thus, $\Delta$ has eigenvalues of $0, (s-1+p)/(s-1)$ with multiplicity $s-1$, $(n-1+q)/(n-1)$ with multiplicity $n-1$, and $2-p-q$, whereas $\Delta_A$ has $N-2$ zero eigenvalues with the remaining two eigenvalues imaginary and equal to $\pm \sqrt{(1-p)(1-q)/sn}$.

The total probability flow from $V_s$ to $V_n$ equals $s(1-p)$ while the total flow from $V_n$ to $V_s$ is $n(1-q)$. In the bipartite case, the subsets of interest are $V_s$ and $V_n$. The right and left conductance of the graph are thus defined as

$$C(V_s) = \frac{\bar{s} + n}{sn} \sum_{i \in V_s, j \in V_n} w_{ij} = \frac{n}{s}(1-p)$$

and

$$C(V_n) = \frac{\bar{s} + n}{sn} \sum_{i \in V_n, j \in V_s} w_{ij} = \frac{n}{s}(1-q). \quad (6.6)$$

The conductance of a set of vertices $S$ is inversely related to the mixing time of a Markov process starting from a vertex in $S$. Thus, $C(V_s) \geq C(V_n)$ implies that the mixing time starting from a vertex in $V_s$ will be less than that starting from a vertex in $V_n$. The condition for $C(V_s) \geq C(V_n)$ is

$$1 - p \geq \frac{n}{s}(1-q). \quad (6.7)$$

**Figure 10** shows the single vertex probability for the $(2,5)$ partial bipartite graph for fixed values of $q$ with $p$ and $r$ as variables and for $r = 2$ with $p$ and $q$ as variables.

Although figure 10c and 10d appear quite similar there is a significant difference. In figure 10d the $p = 0$ line is at zero, whereas in figure 10c this is not quite the case, as indicated in the detail of figure 10e.

**Figure 11** shows cross sections of figure 10e. While the difference from the Moran probability is extremely small, these figures show an unusual effect in which selection is enhanced for fitness values less than one, suppressed for a limited range of values greater than one and then enhanced again. The graphs of figure 11 illustrate a transition between three distinct regimes of behaviour. In the first, selection is suppressed for $0 < r < 1$ and enhanced for $r > 1$ (pattern A, table 1). In the second, selection is enhanced for $0 < r < 1$, and there is a value $r_{max}$ dependent on $p$, such that selection is suppressed for fitness in the range $[1, r_{max}]$ and enhanced for fitness greater than $r_{max}$ (pattern C, table 1). In the third, selection is enhanced for $0 < r < 1$ and suppressed for $r > 1$ (pattern D, table 1).
Figure 10. Fixation probabilities minus Moran probabilities for (2, 5) partial bipartite graph for varying \( q \) values (a–d), (e) detail of (c) \( 0 \leq p \leq 1/3 \) and (f) the \( x_1 \) (vertex in \( V_2 \) component) and \( x_4 \) (vertex in \( V_5 \) component) probabilities for \( r = 2 \).

for \( 0 < r < 1 \), enhanced for fitness in the range \([1, \, r_{\text{max}}]\) and suppressed for \( r > r_{\text{max}} \) (pattern B, table 1). Table 1 shows comparative \( q \)- and \( p \)-values illustrating this effect, which will be discussed further in §7.

The \( q = 0.6667, \, p = 0.118 \) case is an anomaly in that selection is enhanced for the two regions \( 0 < r < 1 \) and \( 42.2788 < r < 175.4007 \) and suppressed for larger \( r \) (the enhancement is of the order of \( 10^{-11} \) in the second region).
Figure 11. Cross sections of figure 10e.

Table 1. Patterns of single vertex fixation probability in (2, 5) partial bipartite graph.

| q          | pattern A                              | pattern B \((p, r_{\text{max}})\)                              | pattern C \((p, r_{\text{max}})\)                              | pattern D                |
|------------|----------------------------------------|---------------------------------------------------------------|---------------------------------------------------------------|--------------------------|
| 0          | \(p < 0.5\)                            | \((0.5,257.0019),(0.55,23.5695),(0.6,12.5060), (0.75,5.3575), (0.875,3.7144), (1,2.8845)\) |                                                             |                          |
| 0.125      | \(p < 0.445\)                          | \((0.445,261.1673),(0.45,91.6745), (0.55,6.2814), (0.6,4.1189), (0.75,1.3160), (0.85,1.0000)\) |                                                             | \(p > 0.85\)             |
| 0.3333     | \(p < 0.3375\)                         | \((0.3375,109.1445),(0.3402,28.3141), (0.35,11.2907), (0.3667,4.9994), (0.4,2.1249), (0.5,1.0000)\) |                                                             | \(p > 0.5\)              |
| 0.5        | \(p < 0.2\)                            | \((0.2075,1.0260), (0.2081,1.1176), (0.2085,1.2023), (0.21,1.4381), (0.215,2.2926), (0.228,404.7369)\) |                                                             | \(p > 0.228\)            |
| 0.625      | \(p < 1/16\)                           |                                                             |                                                             | \(p > 1/16\)             |
| 0.6667     | \(p < 0.0833\)                         | \((0.0833,1.0000),(0.0999,1.7239), (0.1,2.7687), (0.1111,5.9870), (0.117,247.4142)\) |                                                             | \(p > 0.125\)            |
| 0.875      | \(p > 0.585\)                          | \((0.48,821.1031),(0.49,13.0399), (0.5,7.0984), (0.51,5.0663), (0.55,2.3417), (0.585,1.2061)\) |                                                             | \(p < 0.48\)             |

Figure 12 shows probabilities for the (3,4) partial bipartite graph as functions of \(p\) and \(q\) for \(r\) taking values 5/4, 3/2 and 2. In all cases, the ranges of \(p\) and \(q\) are \([0, 7/8]\) in order to avoid a singular that arises if \(p\) or \(q\) is 1.

Figure 12b–d plots \(x_1\) and \(x_8\), the single vertex probabilities for a vertex starting in the three vertex subset and the four vertex subset, respectively. The equation for \(x_1 - x_8 = 0\) contains over 350 terms in
Figure 12. Single vertex fixation probabilities for (3, 4) partial bipartite graph with $r = 5/4, 3/2$ and 2. (a) Fixation probability minus Moran probability with $r = 2$, (b–d) fixation probabilities starting on three sides (sheet that is lower at $q = 0$) compared to that starting on fourth side for $r = 5/4, 3/2, 2$.

products of $p$ and $q$ with coefficients of up to 20 digits. Nevertheless, a simple solution exists and is the same for all three values of $r$ considered: $q = (1 + 3p)/4$. This compares exactly with the condition based on conductances given in equation (6.7). Likewise, for the $(2, 5)$ case the $x_1 - x_4 = 0$ line is $q = (3 + 2p)/5$, which again is the same condition as given by equation (6.7). This suggests that equation (6.7) provides a sharp differentiation between fixation probabilities for vertices starting in one or the other of the two graph components $V_s$ and $V_n$.

If the condition of equation (6.7) is satisfied the subset $V_s$ will be said to be evangelical relative to $V_n$ while if this condition is not satisfied $V_s$ is isolated relative to $V_n$. In the extreme case $p = 1$, we will say that $V_s$ is quarantined.

Estimates of hitting times were computed for the $(s, n)$ partial bipartite graph for a number of examples and the form of these cases suggests the following conjecture.

**Conjecture 6.2.** Let $G$ be an $(s, n)$ partial bipartite graph with $h(s, s)$ and $h(n, n)$ the respective hitting times between pairs of vertices in $V_s$ and $V_n$, and $h(s, n)$ the hitting time starting from a vertex in $V_s$ and ending at one in $V_n$ while $h(n, s)$ is the hitting time starting from a vertex in $V_n$ and ending at one in $V_s$. Then

\[
\begin{align*}
    h(s, s) &= \frac{s(s - 1)(2 - p - q)}{(1 - q)(s - 1 + p)}, \\
    h(n, n) &= \frac{n(n - 1)(2 - p - q)}{(1 - p)(n - 1 + q)}, \\
    h(s, n) &= \frac{(n - 1)(2n - 1) - (n - 1)^2p - n(n - 2)q}{(1 - p)(s - 1 + q)}, \\
    h(n, s) &= \frac{(s - 1)(2s - 1) - (s - 1)^2q - s(s - 2)p}{(1 - q)(n - 1 + p)}.
\end{align*}
\]

and

\[
\begin{align*}
    h(s, s) &= \frac{s(s - 1)(2 - p - q)}{(1 - q)(s - 1 + p)}, \\
    h(n, n) &= \frac{n(n - 1)(2 - p - q)}{(1 - p)(n - 1 + q)}, \\
    h(s, n) &= \frac{(n - 1)(2n - 1) - (n - 1)^2p - n(n - 2)q}{(1 - p)(s - 1 + q)}, \\
    h(n, s) &= \frac{(s - 1)(2s - 1) - (s - 1)^2q - s(s - 2)p}{(1 - q)(n - 1 + p)}.
\end{align*}
\]
Given this conjecture, the condition for \( h(s, s) \geq h(n, n) \) is

\[
q \geq \frac{(s - 1)(n - 1)(n - s) + (n - 1)[n + s(s - 1)]}{sn(n - 1) - [n(n - 1) - s(s - 1)][1 - p]}.
\]

(6.9)

If \((s, n) = (3, 4)\) then equation (6.9) becomes \( q \geq (1 + p)/(5 + p) \).

Note that the hitting times \( h(s, s) \) and \( h(n, n) \) both involve a factor of \( 2 - p - q \) which is an eigenvalue of the Laplacian. If \( p \) and \( q \) are sufficiently small, \( 2 - p - q \) is the largest eigenvalue. From equation (6.5), ‘sufficiently small’ requires

\[
q \leq \frac{(1 - p)s - 1}{s - 1} \quad \text{and} \quad p \leq \frac{(1 - q)n - 1}{n - 1},
\]

(6.10)

with equality yielding \( p = (s - 1)/(N - 1) \), \( q = (n - 1)/(N - 1) \). More generally, eigenvalues of the Laplacian can be plotted as functions of \( p \) and \( q \), as illustrated in figure 13a for \((s, n) = (3, 4)\), and the curves for equality of pairs of the non-zero eigenvalues can be computed, showing \( q \) as a function of \( p \) as in figure 13b.

The point of equality for all three eigenvalues is \((p, q) = (1/3, 1/2)\).

Ranking eigenvalues according to size for the six different regions in this graph is shown in table 2.

In the limiting case of \( p = q = 1 \), the eigenvalue \( 2 - p - q \) becomes zero and the graph separates into two disconnected components. Care is required, however, to determine how \( h(s, s) \) and \( h(n, n) \) behave in this limit since the factors \((2 - p - q)/(1 - p) \) and \((2 - p - q)/(1 - q) \) give results that depend on how the limit \((p, q) \to (1, 1) \) is approached. For example, suppose that \( p = q^2 \). Then \( 2 - p - q = 2 - q - q^2 = (1 - q)(2 + q + q^2 + \cdots + q^{k-1}) \) and at \( q = 1h(s, s) = (1 + k)(k - 1) \) and likewise \( h(n, n) = (1 + k)(k - 1) \). These differ from the hitting times for size \( s \) and size \( n \) complete graphs, which are \( s - 1 \) and \( n - 1 \), respectively. The way to deal with this in computing \( h(s, s) \) is to set \( p = 1 \) going to the limit (or, in computing \( h(n, n) \) setting \( q = 1 \)). Otherwise there is always the possibility of transitions between \( V_s \) and \( V_n \) giving additional contributions to hitting times.

Exact computation of \( e^W \) is possible for partial bipartite graphs and examination of the cases (2, 3), (3, 4), (2, 5), (3, 5) and (4, 5) suggests the following.

**Conjecture 6.3.** Let \( G \) be an \((s, n)\) partial bipartite graph, \( V_{ss} \) and \( V_{nn} \) the diagonal elements of \( e^W \) corresponding to vertices in \( V_s \) and \( V_n \), with \( V_{sn} \) and \( V_{ns} \) refer to off block diagonal elements. Then

\[
V_{ss} = \frac{se}{2 - p - q} \left[ 1 - q + (1 - p)e^{p+q-2} \right] \quad V_{sn} = \frac{s(1 - p)e}{2 - p - q} \left[ 1 - e^{p+q-2} \right]
\]

\[
V_{ns} = \frac{n(1 - q)e}{2 - p - q} \left[ 1 - e^{p+q-2} \right] \quad \text{and} \quad V_{nn} = \frac{ne}{2 - p - q} \left[ 1 - p + (1 - q)e^{p+q-2} \right].
\]

(6.11)

In addition,

\[
\text{Tr}(e^W) = e \left[ 1 + e^{p+q-2} + (n - 1)e^{-(n-1)/(n-1)} + (s - 1)e^{-(s-1)/(s-1)} \right].
\]

(6.12)

Setting partial derivatives of \( \text{Tr}(e^W) \) with respect to \( p \) and \( q \) to zero and solving for \( q \) yields

\[
q = \frac{s - 1 - sp}{s - 1} \quad \text{and} \quad q = \frac{(n - 1)(1 - p)}{n}.
\]

(6.13)

Setting these equal indicates that the minimum value of \( \text{Tr}(e^W) \) occurs at

\[
(p, q) = \left( \frac{1}{n + s - 1} \right)(s - 1, n - 1),
\]

(6.14)

which is the condition for eigenvalue equality in equation (6.10).

Figure 14 shows plots of \( V_{ss} \) and \( V_{nn} \), \( V_{sn} \) and \( V_{ns} \), and \( \text{Tr}(e^W) \) for \((s, n) = (3, 4)\).

In figure 14a, \( V_{33} \) is lowest at \( p = 0 \), \( q = 1 \) and in figure 14b, \( V_{sn} \) is zero at \( p = 1 \), whereas \( V_{ns} \) is zero at \( q = 1 \). The minimum value in figure 14c occurs at \((p, q) = (1/3, 1/2)\), which coincides with the point at which all non-zero eigenvalues of \( \Delta \) are equal. The line of equality in figure 14a is difficult to express but can be approximated as the straight line \( q = 1.0923872p - 0.3800648 \), thus for \( p > (q + 0.3800648)/1.0923827, V_{ss} \) will be larger than \( V_{nn} \). In figure 14b, the line of equality is given by \( q = (1 + 3p)/4 \). Likewise, if \( q > (1 + 3p)/4 \) then \( V_{ns} < V_{sn} \), coinciding with the condition on conductances given in equation (6.7).
7. Single-link graphs

In this section, we consider graphs consisting of two components with linkage between components occurring only through a single vertex in each component, as illustrated in figure 15.

These graphs and partial bipartite graphs are the extreme cases for graphs having two components with linkages occurring only between subsets of each component. Here the component \( V_3 \) contains \( s \) vertices, one of which is linked to one of the \( n \) vertices in \( V_n \). Each non-linked vertex in \( V_3 \) connect to every other vertex in this component with probability \( 1/(s-1) \) while the linking vertex connects to all other vertices in \( V_3 \) with probability \( p/(s-1) \) and connects to the linking vertex in \( V_n \) with probability \( 1-p \). Likewise, all non-linked vertices in \( V_n \) connect to every other vertex in this component with probability \( 1/(n-1) \) while the linking vertex connects to all other vertices in \( V_n \) with probability \( q/(n-1) \) and connects to the linking vertex in \( V_3 \) with probability \( 1-q \).

Single vertex fixation probabilities are suppressed relative to the Moran probability for cases considered, as shown in figure 16 for \((s,n) = (3,4)\) for \( r = 5/4 \) and \( r = 2 \). This figure also shows comparisons of fixation probabilities for possible initial mutant vertices for \( r = 5/4 \) and 2. Here \( x_1 \) is the fixation probability for a non-linked vertex in \( V_3 \), \( x_4 \) is the fixation probability for the linking vertex in \( V_3 \), \( x_8 \) is the fixation probability for the linking vertex in \( V_4 \), and \( x_{16} \) is the fixation probability for a non-linked vertex in \( V_4 \).

In figure 16c,d, if \( p = 0 \) the \( x_{16} \) probability is zero and the \( x_1 \) probability is the Moran probability for \( s-1 \) vertices. Likewise, if \( q = 0 \) the \( x_1 \) probability is zero and the \( x_{16} \) probability is the Moran probability for \( n-1 \) vertices because fixation can only occur if the initial mutant is introduced in the \( \{v_1, v_2, v_3\} \) or the \( \{v_4, v_5, v_6, v_7\} \) vertex sets, respectively. If \( p = 1 \) (\( q \neq 1 \)), on the other hand, the \( x_1 \) probability is zero and if \( q = 1 \) (\( p \neq 1 \)), the \( x_{16} \) probability is zero (these are not shown in figure 16 which only shows \( p \) and \( q \) between the values of 0 and 7/8). If \( p = q = 1 \) the graph has split into disconnected components, which
Figure 14. Plots of (a) $V_{31}$ and $V_{44}$, (b) $V_{34}$ and $V_{43}$, and (c) $\text{Tr}(e^W)$ for (3, 4) partial bipartite graph.

Figure 15. Single line graph.

shows up in the appearance of a second zero eigenvalue for the Laplacian as shows up in figure 17 below.

Taking $x = \lambda - 1$, the characteristic polynomial for the Laplacian of this graph is

$$
(x - \frac{1}{s - 1})^{s-2} \left( x - \frac{1}{n-1} \right)^{n-2} \left\{ (x^2 - (1-p)(1-q)) \left( x + \frac{s-2}{s-1} \right) \left( x + \frac{n-2}{n-1} \right) \right. \\
\left. - \left[ \frac{q}{n-1} \left( x + \frac{s-2}{s-1} \right) + \frac{p}{n-1} \left( x + \frac{n-2}{n-1} \right) \right] x + \frac{pq}{(s-1)(n-1)} \right\} \quad (7.1)
$$

Thus, there will be $s - 2$ roots $\lambda = s/(s - 1)$, $n - 2$ roots $\lambda = n/(n - 1)$. The last term in brackets factors into $x + 1$ (yielding the 0 eigenvalue) and a cubic. Figure 17 shows plots of the cubic factor for chosen values of $p$ and $q$.

In the lowest curve of figure 17a, $p = q = 0$, corresponding to an eigenvalue of 2 (since $x = \lambda - 1$), whereas in the upper curve of figure 17e, $p = q = 1$, giving a second eigenvalue of 0 ($x = -1$) as the graph is split into two disconnected components.

With the linked vertices denoted $v_s$ and $v_n$, the conductances for the sets $V_s \setminus v_s$, $V_n \setminus v_n$, $v_s$, $v_n$, $V_s$ and $V_n$ are given by

$$
\begin{align*}
C(V_s) &= \frac{n+s}{(s-1)(n+1)} \\
C(V_n) &= \frac{n+s}{(s+1)(n-1)} \\
C(v_s) &= C(v_n) = \frac{n+s}{n+s-1} \\
C(V_s) &= \frac{(n+s)(1-p)}{sn} \quad \text{and} \quad C(V_n) = \frac{(n+s)(1-q)}{sn}.
\end{align*}
$$

(7.2)

All but the two of these are constant and the condition $C(V_s) = C(V_n)$ is that $q = p$.

Approximation of $e^W$ and $\text{Tr}(e^W)$ for the (3, 4) case yields figure 18.
Figure 16. Fixation probabilities for (3, 4) single-link graph for \( r = 5/4 \) (a,c,e) and \( r = 2 \) (b,d,f): (a,b) \( \rho - \rho_m \), (c,d) \( x_1 \) and \( x_{16} \), (e,f) \( x_4 \) and \( x_8 \).

The minimax values \( \text{Tr}(e^{W/7}) \) are at \((p,q) \sim (0.793893, 0.394234)\). The line of equality in figure 17b is approximately given by \( q \sim 2.021513p + 0.0074732 \)—if \( q \) is greater than the above value, vertex \( v_1 \) has the greater centrality. In figure 17e, the line of equality is \( q = p \) which, as in the partial bipartite case, corresponds to the condition for \( C(V_3) = C(V_4) \), and we conjecture that this holds for a general single-link graph. Table 3 shows the hitting times for the (3, 4) single-link graph.

In contrast to hitting times for the partial bipartite graphs, most of the hitting times in table 5 show divergences for cases where \( p \) or \( q \) are 0 or 1. Consideration of the graph topology makes it apparent that the hitting time \( h_{uv} \) will diverge whenever there is either zero probability of a transition from vertex
Figure 17. Roots of cubic factor of the (3, 4) single-link characteristic polynomial (a–e) correspond to $p = 0, 1/4, 1/2, 3/4, 1$, respectively. In each figure, $q$ varies from bottom to top with values 0, 1/4, 1/2, 3/4, 1.

Figure 18. (a) $\text{Tr}(e^{W}/7)$, (b) centrality of $v_1$ and $v_5$ vertices, (c) centrality of $v_3$ and $v_4$ vertices, (d) average centrality of $V_3$ and $V_4$ components, (e) communicability of $\{v_1, v_2\}$ and $\{v_5, v_6, v_7\}$ vertex sets, (f) communicability of $v_3$ and $v_4$ vertices.
Construction of graphic population models involves few conceptual problems, the procedure is well established: locate each population member, characterized by one or more variable conditions, at a vertex of a graph and label edges between vertices with some measure of interaction between linked individuals. Then define an updating process relating individual interactions to the dynamics of characteristic variables. While this sort of model can become mathematically complex, the conceptual framework is simple.

Equation (2.1) treats the easiest case, in which each population member exhibits a binary-valued characteristic and the edge weights are just probabilities of interaction, coupled with birth–death updating. Solutions to this equation provide birth–death fixation probabilities for any given initial state.

The problem is the combinatorial explosion that arises. If the population has size $N$, each population member is characterized by a vector of characteristics $c = (c_1, \ldots, c_k)$, and each of these characteristics $c_i$ can exhibit one of $m_i$ possible values, then the state space contains almost $\left( \prod_{i=1}^k m_i \right)^N$ members.

One way of dealing with this is through methods from statistical mechanics in which graphs are constructed according to algorithms for developing random, small world or other connections between vertices. In this paper, a different direction is taken—we study properties of some very simple graphs in the hope of eventually discovering connections to the statistics of larger graphs. In analogy, a molecular gas can be studied in terms of statistical mechanics by assuming that all molecules are point particles and using concepts such as temperature, entropy, pressure, free energy and so on; but it is also possible to study the structure of the individual molecules with an eye towards eventual discovery of connections between molecular structural properties and large-scale gas behaviour.

Four classes of graphs were considered: biased cycles, dual circular flows, partial bipartite graphs and single-link graphs. For biased cycles, theorem (4.1) gives the characteristic polynomial of the Laplacian matrix and the corollary to this theorem expresses eigenvalues in terms of roots of unity for the special single-link graphs. For biased cycles, theorem (4.1) gives the characteristic polynomial of the Laplacian of a circular flow to that of the biased cycles, with the term $1 - 2p$ as a measure of cycle bias. Computation of hitting times for simple cases shows the characteristic form illustrated in figure 3, with maximum hitting times coming at $p$-values characterizing a ‘tipping’ point between clockwise and counter-clockwise motion for a random walk.

Biased cycles provide the simplest example of the dual circular flow graphs and theorem (5.1) links the characteristic polynomial of the Laplacian of a circular flow to that of the biased cycles, with the term $1 - 2p$ again playing the role of measuring directional bias. Further generalization allowed the possibility of ‘tuning’ a flow to match specified parameters, exemplified by derivation of the conditions required to match any specified vertex temperature profile.

Bipartite graphs are two-level dual circular flows and, as indicated in the Introduction section, the generalization to partial bipartite graphs leads to results of particular interest, showing a tight connection between conductances, Laplacian eigenvalues, vertex centralities, expected hitting times and single vertex fixation probabilities. These graphs also showed the possibility of counterintuitive behaviour in which fitness values less than 1 can result in enhanced selection, illustrated in figure 11 and table 1. While these tables show only the (2, 5) partial bipartite case, similar behaviour has been found for the (3, 4) case, although the ranges of $p$ for which the behaviour appears, at least in preliminary investigation, are narrower. From table 1, for fixed $q$ there appears to be two critical $p$-values, $p_-$ and $p_+$, depending on $q$, such that in most cases with $q < 1/2$ pattern A appears for $p < p_-$, pattern D appears for $p > p_+$ and pattern B for $p_- < p < p_+$; while for $q \geq 1/2$ pattern C appears rather than pattern B when $p_- < p < p_+$. The value of $r_{\text{max}}$, the fitness value that distinguishes between patterns A and B or between patterns C and D depends on both the $q$- and $p$-values. In table 1, pattern C shows up for cases in which $q$ is relatively large while $p$ is relatively small. This corresponds to a situation in which the five vertex is strongly connected internally, with weak linkages to the two vertex subset, while this latter subset is weakly connected internally with strong links to the five vertex subset. Under these conditions, a mutant with $r < 1$ appearing in $V_2$ will have a low probability of extinction and, if chosen for reproduction, a high probability of reproducing in $V_5$ so that in a sense $V_2$ may act as a root. Much further work is required, however, to develop any detailed understanding of this behaviour, as well as extending analysis to dual circular flows in which vertices in each level can influence each other.

The final case consists of a two-component graph in which the linkage between components occurs only through a single vertex in each component. Laplacian eigenvalues, hitting times and communicability were determined, and fixation probabilities were found for the (3, 4) example, allowing comparison to the corresponding partial bipartite case. As indicated, partial bipartite and single-link graphs are the two extremes of a family of graphs that provide models of two populations.
Table 5. Summary of results (roman indicates general results, italic indicates limited exemplary cases).

| Graphs              | Eigenvalues                  | Conductance                         | Communicability, Centrality | Expected Hitting Times | Fixation Probabilities |
|---------------------|------------------------------|-------------------------------------|------------------------------|------------------------|------------------------|
| Biased Cycles       | Theorems (4.1) and (4.2), Corollary (4.2), Equation (4.2) | $N/k(N-k)$ for a length $k$ block | Figure 2                    | Figure 3a–d           | Same as Moran          |
| Dual Circular Flows | Theorem (5.1)                | $N/(N-n_s)$ for level $s$           | Figure 6                    | Same as biased cycle between levels | Figure 5               |
| Partial Bipartite  | Theorem (6.1), Equation (6.5) | Equation (6.6)                      | Conjecture (6.2); Figure 14; Equations (6.11) and (6.12); Conjecture (6.2), Equations (6.8) and (6.9) | Figures 10–12, Tables 1 and 2 |
| Single Link         | Figure 17, Equation (6.15)   | Equation (6.16)                     | Figure 18                   | Tables 3 and 4         | Figure 16              |
communicating across smaller sets of representatives. Adjustment of connection parameters provides models for different levels of polarization in the two populations. Further work is required to analyse both the partial bipartite graphs and single-link graphs, as well as to consider more general cases in which linkages between two populations are through subsets $S_k \subseteq V_s$ and $S_m \subseteq V_n$. Being able to characterize polarization within a population in this way could provide an important tool for studies of social interactions.

Table 5 shows a general summary of results reported in the present paper.

Authors' contributions. R.B. developed programs in Maple 15 to solve the computational problems involved in this paper, including taking an edge weight matrix $W$ from that generating and solving equation (2.1), and finding and interpreting hitting times. B.V. developed the conceptual design of the research, wrote the final draft of the paper and is responsible for revisions. Both authors approve publication.

Competing Interests. The authors have no competing interests.

Acknowledgements. We thank a referee for bringing references [27,28] to our attention and suggesting the further analysis leading to discovery of the behaviour indicated in table 1. Alex Murray contributed to earlier work on this research programme.

References

1. Hauert C. 2008 Evolutionary dynamics. In Proceedings of the NATO Advanced Study Institute: evolution from cellular to social scales (eds AT Skjeltorp, AV Belokurikin), pp. 11–44. Dordrecht, The Netherlands: Springer.

2. Masuda N, Ohtsuki H. 2009 Evolutionary dynamics and fixation probabilities in directed networks. New J. Phys. 11, 033012. (doi:10.1088/1367-2630/11/3/033012)

3. Shaked R, Roos P. 2011 Fast and deterministic computation of fixation probability in evolutionary graphs. In CB’11: The Ninth International Conference on Computational Intelligence and Bioinformatics. November, 2011.

4. Lieberman E, Hauert C, Nowak MA. 2005 Evolutionary dynamics on graphs. Nature 433, 312–316. (doi:10.1038/nature03204)

5. Simon PL, Taylor M, Kiss IZ. 2011 Exact epidemic models of social interactions. (eds) Evolutionary graph theory with application to complex networks, pp. 11–44. Dordrecht, The Netherlands: Springer.

6. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. 2014 Epidemic processes in complex networks. (http://arxiv.org/abs/1309.6323)

7. Kuperman MN. 2013 Invited review: epidemics on social networks. Papers in Physics 5, 050003. (doi:10.4279/PPI050003)

8. Barbour AD, Reinert G. 2013 Asymptotic behavior of gossip processes and small world networks. Adv. Appl. Prob. 45, 861–1000. (doi:10.1239/aap/1386885854)

9. Coletti CF, Rodriguez PM, Schiavi RB. 2012 A spatially stochastic model for rumor transmission. J. Math. Biol. 62, 479–508. (doi:10.1007/s00285-010-0344-x)

10. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. 2014 Epidemic processes in complex networks. (http://arxiv.org/abs/1309.6323)

11. Kuperman MN. 2013 Invited review: epidemics on social networks. Papers in Physics 5, 050003. (doi:10.4279/PPI050003)

12. Pinto PC, Thiran P, Vetterli M. 2012 Locating the source of diffusion in large-scale networks. Phys. Rev. Lett. 109, 068702. (doi:10.1103/PhysRevLett.109.068702)

13. Shah D, Zamir TR. 2011 Rumors in a network: who’s the culprit? IEEE Trans. Inform. Theory 57, 5163–5181. (doi:10.1109/TIT.2011.2158885)

14. Barbour AD, Reinert G. 2013 Asymptotic behavior of gossip and rumour spreading. In ACM-SIAM Symposium on Discrete Algorithms (TAAS), pp. 1500–1511. (doi:10.1007/s00285-010-0344-x)

15. Dong W, Zhang W, Tan CW. 2013 Rooting out the rumor culprit from suspects. (http://arxiv.org/abs/1301.6323)

16. Sageman M. 2004 Understanding terrorism. Philadelphia, PA: University of Pennsylvania Press.

17. Offat-Saber R, Murray SM. 2004 Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans. Automatic Control 49, 1520–1533. (doi:10.1109/TAC.2004.831413)

18. LeBlanc H, Zhang H, Koutsoukos X, Sundaram S. 2013 Resilient asymmetric consensus in robust networks. IEEE J. Selected Areas Commun. 31, 766–781. (doi:10.1109/JSAC.2013.100428)

19. Censor-Hillel K, Shachnai H. 2012 Fast information spreading in graphs with large weak conducance. SIAM J. Comput. 41, 1451–1465. (doi:10.1137/11045380)

20. Houshmandzadeh H, Vailati M. 2011 The fixation probability of a beneficial mutation in a geographically structured population. New J. Phys. 13, 073020. (doi:10.1088/1367-2630/13/7/073020)

21. Shaked R, Roos O, Johnson A. 2012 A review of evolutionary game theory with applications to game theory. Biosystems 107, 66–80. (doi:10.1016/j.biosystems.2011.09.006)

22. Voorhees B. 2013 Birth–death fixation probabilities for simple digraphs. Proc. R. Soc. A 469, 20120248. (doi:10.1098/rspa.2012.0248)

23. Voorhees B, Murray A. 2013 Fixation probabilities for simple digraphs. Proc. R. Soc. A 469, 20120267. (doi:10.1098/rspa.2012.0267)

24. Moran P. 1958 Random processes in genetics. Math. Proc. Camb. Philos. Soc. 54, 65–71. (doi:10.1017/S0305004100031993)

25. Broom M, Rychtár J. 2008 Analysis of the fixation probability of a mutant on a special class of non-directed graphs. Proc. R. Soc. A 464, 2609–2627. (doi:10.1098/rspa.2008.0558)

26. Zhang C, Wu Y, Liu W, Yang X. 2012 Fixation probabilities on complete star and bipartite digraphs. Disc. Dyn. Nat. Soc. 2012, 940465. (doi:10.1155/2012/940465)

27. Monk T, Green P, Paulin M. 2014 Martingales and fixation probabilities in evolutionary graphs. Proc. R. Soc. A 470, 20130730. (doi:10.1098/rspa.2013.0730)

28. Adlam B, Nowak MA. 2014 Universality of fixation probabilities in randomly structured populations. Sci. Rep. 4, 6692. (doi:10.1038/srep06692)

29. Newman MEJ. 2010 Networks: an introduction. New York, NY: Oxford University Press.

30. Durrett R. 2013 Some features of the spread of epidemics and information on a random graph. Proc. Natl.Acad. Sci. USA 107, 4491–4498. (doi:10.1073/pnas.0914402107)

31. Barzel B, Barabasi AL. 2013 Universality in network dynamics. Nat. Phys. 9, 673–681. (doi:10.1038/nphys2741)

32. Cohen R, Havlin S. 2010 Complex networks: structure, robustness and function. Cambridge, UK: Cambridge University Press.

33. Li Y, Zhang Z. 2012 Digraph Laplacian and the degree of asymmetry. Internet Math. 8, 381–401. (doi:10.1080/15427951.2012.708910)

34. Barbosa VC, Donangelo R, Souza SR. 2010 Early appraisal of the fixation probability in directed networks. Phys. Rev. E 82, 046114. (doi:10.1103/PhysRevE.82.046114)

35. Chung F. 2005 Laplacian and the Cheeger inequality for directed graphs. Ann. Comb. 9, 1–19. (doi:10.1007/s00026-005-0237-z)

36. Estrada E, Hatano N. 2008 Communicability in complex networks. Phys. Rev. E 77, 036111. (doi:10.1103/PhysRevE.77.036111)

37. Estrada E, Hatano N, Benzi M. 2012 The physics of communicability in complex networks. Phys. Rev. E 85, 016115. (doi:10.1103/PhysRevE.85.016115)

38. Kemeny JG, Snell JL. 1983 Finite Markov chains. New York, NY: Springer.

39. Freni M, Ranei PB, Traulsen A. 2013 The effect of population structure on the rate of evolution. Proc. R. Soc. B 280, 20130211. (doi:10.1098/rspb.2013.0211)