




Phases of Small Worlds: A Mean Field Formulation

Andrew D. Jackson¹ · Subodh P. Patil² 

Received: 2 December 2021 / Accepted: 21 September 2022
© The Author(s) 2022

Abstract

A network is said to have the properties of a small world if a suitably defined average distance between any two nodes is proportional to the logarithm of the number of nodes, N . In this paper, we present a novel derivation of the small-world property for Gilbert–Erdős–Rényi random networks. We employ a mean field approximation that permits the analytic derivation of the distribution of shortest paths that exhibits logarithmic scaling away from the phase transition, inferable via a suitably interpreted order parameter. We begin by framing the problem in generality with a formal generating functional for undirected weighted random graphs with arbitrary disorder, recovering the result that the free energy associated with an ensemble of Gilbert graphs corresponds to a system of non-interacting fermions identified with the edge states. We then present a mean field solution for this model and extend it to more general realizations of network randomness. For a two family class of stochastic block models that we refer to as dimorphic networks, which allow for links within the different families to be drawn from two independent discrete probability distributions, we find the mean field approximation maps onto a spin chain combinatorial problem and again yields useful approximate analytic expressions for mean path lengths. Dimorphic networks exhibit a richer phase structure, where distinct small world regimes separate in analogy to the spinodal decomposition of a fluid. We find that it is possible to induce small world behavior in sub-networks that by themselves would not be in the small-world regime.

Keywords Random networks · Mean field · Small-world properties

Communicated by Pierpaolo Vivo.

✉ Subodh P. Patil
patil@lorentz.leidenuniv.nl
Andrew D. Jackson
jackson@nbi.ku.dk

¹ Niels Bohr International Academy, The Niels Bohr Institute, Blegdamsvej 17, 2100 Copenhagen, Denmark

² Instituut-Lorentz for Theoretical Physics, Leiden University, 2333 CA Leiden, The Netherlands

1 Preliminaries

Although small-worlds networks have proved to be fertile ground for theoretical musing among mathematicians and scientists, the initial recognition and exploitation of the small-worlds phenomenon has far more empirical origins. The existence of such networks was first suggested by the Hungarian journalist and author Frigyes Karinthy in a short story published in 1929¹. In this story Karinthy asserts that, “using no more than five individuals”, an unbroken link of acquaintances can be established to connect any two people in the world. This is the source of the now-familiar notion of “six degrees of separation”. Although Karinthy offered anecdotal evidence for his assertion, forty years passed before Travers and Milgram provided empirical evidence for this assertion and coined the term “small worlds” [1]. The small-worlds property has since become a widely-used empirical tool for exploring the relational structures of a variety of social, biological, and physical systems. Such empirical studies have benefited greatly from the availability of massive data sets, among the largest of which is a snapshot of the approximately 700 million monthly active users of Facebook in 2011 (a figure that has since ballooned to 2.7 billion users in 2020) [2]. Not only does this social network exhibit the properties of a small world, it turns out that Facebook users on average have only four degrees of separation between one another [3]. Small world networks, and their associated properties manifest in a multitude of other contexts as well, with recent experience proving its relevance to epidemiology [4–6]. Similarly, many professional networks from Broadway [7] and Wall Street [8] to scientific collaborations [9] exhibit small world properties, encapsulated for mathematicians and theoretical physicists by their Erdős number, for example.

The first mathematical contribution to the small-worlds problem was provided in 1959 in the form of Erdős–Rényi random graphs [10], a class that was proposed independently by Gilbert [11]. Such graphs correspond to a set on N nodes where a link between any two chosen nodes will exist with probability p . Erdős–Rényi–Gilbert random graphs represent a limiting case of randomness in contrast to regular lattices, where links to nearest neighbors are to some extent guaranteed. On a regular lattice-like network, the average distance between any two nodes scales with some power the number of nodes, N , in the network. Along with other small-worlds networks,² Erdős–Rényi–Gilbert random graphs have the property that the distance between nodes is proportional to $\log N$. Interpolating between these two extremes is a class of random networks introduced by Watts and Strogatz [14], for which a regular ring lattice with links to k nearest neighbors can have any given link reassigned to a random node with probability p . A variant of such networks was studied by Newman and Watts [15]. Instead of reassigning links, random links can be inserted between pairs of nodes. Both classes of models [14, 15] exhibit small-world scaling above a critical probability $p \sim 1/N$. The regularity inherited from the original lattice before link reconnections or insertions are performed makes it possible to demonstrate small-world scaling analytically in the large N limit.

There is a qualitative distinction however between the small-world scaling seen in Erdős–Rényi–Gilbert and Watts–Strogatz classes of networks. The former exhibit a smaller clustering coefficient than the latter, whereas many real world examples of small-world networks exhibit large clustering coefficients [14]. Realizations of random networks that model a wider range of real systems can be obtained by specifying different degree distributions, by which one

¹ The story, “Chain-Link”, appeared in a collection entitled *Everything is Different*. An English translation of this story is available online at http://vadeker.net/articles/Karinthy-Chain-Links_1929.pdf.

² Scale-free networks, i.e. networks whose degree distributions exhibit an asymptotic power law behavior $p(k) \propto k^{-\gamma}$ with k denoting the degree (see e.g. [12]) have furthermore been shown to be *ultra-small* where average distances scale as $\log(\log N)$ provided $\gamma < 3$ [13].

can analytically study the distribution of shortest path lengths [16–20], as well as the shell (i.e. differential) [21–23] and cycle structure [24]. In many contexts of interest, networks are not static entities but rather grow over time. Analytic studies of the distribution of shortest paths have also been made in the context of network growth via node duplication, as detailed in [25, 26].

Erdős–Rényi graphs are nevertheless useful in understanding the properties of real world networks through the variety of analytical results that can be proved for them³. Perhaps the most celebrated of these is the critical probability found in the semi-regular network structure of [15] also exists in the purely random context [10, 11]. There exists a threshold value of $p_c \sim 1/N$ above which statements about any representative graph are almost always true. It can be demonstrated, for example, that for $p > (\log N)/N$ there will be a ‘giant component’ containing all nodes in the large N limit [10, 28, 29]⁴. The emergence of such a hub structure is central to the small-world scaling of large networks⁵. However, although one can demonstrate small-world scaling of the average distance between nodes, ℓ , in an Erdős–Rényi–Gilbert network in limiting cases where $p \gg (\log N)/N$ (or vice versa), an analytic understanding of the approach to the transition point for ℓ remains elusive.

Furthermore, although the semi-regularity of the Watts–Strogatz class of networks permits a continuum generalization for which a variant of the mean field approximation has been proposed [32], direct application of such methods fails for Erdős–Rényi–Gilbert graphs because of the lack of an underlying lattice sub-structure⁶. In this paper, we present another variant of the mean field approximation, not predicated on a continuum generalization of the underlying graph structure that allows for an analytic demonstration of small world scaling in Erdős–Rényi–Gilbert random networks. In the case of Watts–Strogatz graphs [14, 15], the underlying lattice structure results in a phase transition at $p = 1/N$ from an $\ell \propto N$ scaling to an $\ell \propto \log N$ scaling. For Erdős–Rényi–Gilbert graphs, the (percolation) transition occurs at a critical $p_c = 1/N$ [34–36] and is of a somewhat different nature due to the absence of an underlying lattice. Here, the transition is from a regime characterized by the existence of multiple disconnected components of similar size (naturally identifiable as ‘domains’) to a regime dominated by a single giant connected component. Moreover, we can show that the mean field approximation fails close to the critical point between these two regimes even in the large N limit due to its inability to cope with the inherent inhomogeneities when the network exists in a phase composed of multiple disconnected components. However, the utility of the approximation described here, characteristic of any mean field approximation, is that it offers analytic understanding of a restricted class of phenomena. Fortunately, the emergence of small-world scaling away from the critical point is among them.

The assumption that all nodes in a network have the same probability of forming links is evidently an idealization of real networks in which some nodes can have an increased probability of forming links. This is the case, e.g., when considering co-author linkages within a given academic field. Paul Erdős had some 240 co-authors, which is unusually

³ Remarkably, the simple abstraction represented by an Erdős–Rényi network may even offer a model for certain macroscopic neuronal oscillations [27].

⁴ As we will see below, when $p = \kappa/N$ the number of particles *not* in the giant component is $\mathcal{O}(N)$. When $\kappa \rightarrow \log N$, the number of nodes not in the giant component is $\mathcal{O}(N^0)$. In other words, past the critical edge probability $p_* = (\log N)/N$, there are no longer any finite clusters in the large N limit.

⁵ However, many interesting properties can also be inferred for sub-critical Erdős–Rényi networks, such as the degree distribution of the giant component [30], and the distribution of shortest paths across all components [31].

⁶ See however [33] for a network model with link randomness that interpolates between the Watts–Strogatz and Erdős–Rényi–Gilbert models.

large in mathematics. The ability of such special nodes to reduce the path lengths was noted by Karinthy in 1929: “It’s always easier to find someone who knows a famous or popular figure than some insignificant person”. Fortunately, the mean field approximation is readily generalized to describe graphs containing such inequivalent nodes, and it will be shown that qualitatively new network behavior emerges from this extension. We consider a simple two family stochastic block variant of the Erdős–Renyi–Gilbert network, where links within a given family appear with fixed but distinct probabilities and links between the families appear with some finite probability and allowing for the relative sizes of these families to be arbitrary. This is not quite the widely studied *planted partition* stochastic block model, as we allow for the links within members of each family to appear with different probabilities, and so we will refer to this variant in what follows as a *dimorphic network*.

The dimorphic variant of the Erdős–Renyi–Gilbert network exhibits a richer phase structure than the standard case and is also accurately modelled by the mean field approximation provided that the network does not consist of “too many” disconnected components. One phase corresponds to the situation where the network factorizes into two strongly connected sub-networks with a limited number of ‘bridge’ links between them. The transition to this phase from a more homogeneous, but still disordered phase is roughly analogous to the spinodal decomposition of fluids. We also find that it is possible to induce small world behavior in a sub-network even if it does not meet the criteria to be in the small-world regime by itself, provided it connects to a sub-network that does. One can also envision a variant of the dimorphic network where the links within and between sub-networks have different weights. The formalism described here can readily be extended to describe such, and we shall pursue this possibility in subsequent investigations.

The remainder of this paper is organized as follows. In Sect. 2, we construct a formal generating functional for an arbitrary undirected weighted network, recovering the result that the associated free energy for an ensemble of Gilbert graphs corresponds to a system of non-interacting fermions residing at the edges. We also show how one can obtain formal expressions for ensemble averages of various network observables such as the average degree of the network, the degree distribution as well as the distance matrix. We note, however, that actually evaluating these quantities will be unduly cumbersome in practice. In Sect. 3 we introduce a mean field approximation that facilitates the evaluation of the ensemble average of the distance matrix for Erdős–Renyi–Gilbert networks and demonstrate that it accurately reproduces the small worlds scaling away from the critical edge probability $p_* = (\log N)/N$. We show that this approximation fails as one approaches the latter critical point, and that this is to be expected since the network fragments into multiple domains. In Sects. 4 and 5, we generalize the mean field approximation to dimorphic networks and consider its different phases. The various technical details will be relegated to appendices. We offer concluding remarks in Sect. 6, where we discuss generalizations of the mean field approximation presented here to more general varieties of network disorder.

2 Random Networks from Random Matrices

Various properties of a network can be obtained from its associated adjacency matrix, \mathbf{A} . All diagonal elements are 0. For unweighted networks, the off-diagonal element \mathbf{A}_{ij} is 1 if nodes i and j are connected by an edge and otherwise 0. For weighted networks the

non-vanishing entries of A_{ij} can be any real positive number⁷. Undirected networks have symmetric adjacency matrices; directed networks do not. It is clear that random networks can be described by the associated ensemble of random adjacency matrices defined by a particular probability measure.

Given a probability measure $P(\mathbf{A})$ over the space of $N \times N$ real matrices

$$\int \mathcal{D}\mathbf{A} P[\mathbf{A}] = 1 \tag{1}$$

where

$$\mathcal{D}\mathbf{A} P[\mathbf{A}] := \prod_{i,j} dA_{ij} P_{ij}, \tag{2}$$

we can define the generating functional

$$Z[J] = \int \mathcal{D}\mathbf{A} P[\mathbf{A}] e^{-\text{Tr} \mathbf{A} J^T} \tag{3}$$

such that expectation values of arbitrary powers of the adjacency matrix can be obtained from successive derivatives of $Z[J]$

$$\langle \mathbf{A}_{ij}^m \rangle = (-1)^m \frac{\delta^m Z[J]}{\delta J_{i'1} \delta J_{i'k} \dots \delta J_{l'l} \delta J_{l'j}} \Big|_{J=0}. \tag{4}$$

The Gilbert–Erdős–Renyi model [11] corresponds to an unweighted, undirected network of N nodes, where a link between any two nodes occurs with a fixed probability p . The associated probability measure is given by⁸

$$\begin{aligned} (i < j) \quad P_{ij} &= p \delta(A_{ij} - 1) + (1 - p) \delta(A_{ij}) \\ (i = j) \quad P_{ij} &= \delta(A_{ii}) \\ (i > j) \quad P_{ij} &= \delta(A_{ij} - A_{ji}) \end{aligned} \tag{5}$$

Computation of the generating functional $Z[J]$ for any given probability measure is straightforward since the measure and the source term factorize as

$$Z[J] = \prod_{i,j} \int dA_{ij} P_{ij} e^{-A_{ij} J_{ij}}. \tag{6}$$

It follows directly from Eq. (5) that

$$Z[J] = (1 - p)^{\frac{N(N-1)}{2}} \prod_{i < j} \left[1 + e^{-(\lambda + J_{(ij)})} \right]; \quad \lambda := -\log\left(\frac{p}{1-p}\right), \quad J_{(ij)} := J_{ij} + J_{ji}. \tag{7}$$

We can also define the generating function, $W[J]$, for all connected correlation functions as

$$e^{-W[J]} := Z[J] \tag{8}$$

which from Eq. (7) is given by

$$W[J] = \frac{N(N-1)}{2} \log\left(\frac{1}{1-p}\right) - \sum_{i < j} \log\left[1 + e^{-(\lambda + J_{(ij)})}\right]. \tag{9}$$

⁷ Here, we will not consider networks where nodes can have edges that begin and end on the same node (i.e. loops) and will thus have adjacency matrices with diagonal entries.

⁸ We also illustrate a variant of the Watts-Strogatz model [14] in Appendix A.

Since the pre-factor in Eq. (7) serves only to normalize $Z[0] = 1$, it can be discarded provided that all moments are obtained by dividing Eq. (4) by $Z[0]$. The generating function for Gilbert random graphs then becomes

$$Z[J] = \prod_{i < j} \left[1 + e^{-(\lambda + J_{(ij)})} \right] \text{ with } \lambda := -\log \left(\frac{p}{1-p} \right) \text{ and } J_{(ij)} := J_{ij} + J_{ji} \quad (10)$$

with a corresponding free energy $W[J]$ given by

$$W[J] = - \sum_{i < j} \log \left[1 + e^{-(\lambda + J_{(ij)})} \right]. \quad (11)$$

We note that the above corresponds to the free energy for an ‘exponential random graph’ obtained from a probability measure that maximizes the Gibbs entropy subject to a particular set of constraints [37]. As noted in [37], the partition function Eq. (10) and free energies Eq. (11) correspond to a system of non-interacting fermions localized on the links. This is made even more apparent when one calculates the one particle occupation numbers:

$$-\frac{1}{Z} \frac{\delta Z}{\delta J_{ij}} \Big|_{J=0} = n_{ij} = \frac{1}{e^\lambda + 1} = p. \quad (12)$$

Taking the appropriate derivatives of Eq. (10) with respect to the source components J_{ij} , one can obtain ensemble averages of any network observable that can be expressed in terms of the adjacency matrix. In the case of unweighted networks for instance, the average number of paths of length k connecting nodes i and j , defined as $\#_{ij}(k)$ is given precisely by Eq. (4):

$$\#_{ij}(k) = \langle \mathbf{A}_{ij}^k \rangle. \quad (13)$$

The degree of the i th node in any given network realization is given by the bilinear form $(e_i, \mathbf{A}u)$ where e_i is the i th unit vector, and where $u = \sum e_j$ is the vector with 1 in all of its entries. Therefore the ensemble averaged degree of the i th node, defined as k_i , is given by

$$k_i = \langle (e_i, \mathbf{A}u) \rangle = \sum_{j=1}^N \langle \mathbf{A}_{ij} \rangle. \quad (14)$$

Note that this need not be the same for all i as the P_{ij} in Eq. (2) need not be the same for all links connecting nodes i, j . The average degree of the entire network, $\langle k \rangle$, can be obtained as

$$\langle k \rangle = \frac{1}{N} \langle (u, \mathbf{A}u) \rangle = \frac{1}{N} \sum_{i,j=1}^N \langle \mathbf{A}_{ij} \rangle. \quad (15)$$

The analogs of the above expressions for general heterogeneous networks can also be found, for instance in Refs. [38, 39].

The average value of the shortest path length between any two nodes is of particular interest in this investigation. This first requires us to specify the notion of a distance on a network. For a general weighted network, a natural (although by no means unique) choice is provided by the analog of effective conductance in an electrical circuit [40]:

$$D_{ij} = \min \left\{ \left(\frac{1}{A_{ih}} + \dots + \frac{1}{A_{hj}} \right) \right\} \quad (16)$$

where the sum includes all possible paths connecting nodes i and j . For the unweighted networks to be considered here, this reduces to the usual notion of the minimum number of

links connecting nodes i and j . In this case, the diagonal elements of the distance matrix \mathbf{D} are defined as 0, and the off-diagonal elements are smallest number of links connecting any pair of nodes (which can be infinite if no such path exists). The average distance between any two nodes is thus given by

$$\ell = \binom{N}{2}^{-1} \sum_{i < j} \langle \mathbf{D}_{ij} \rangle = \binom{N}{2}^{-1} \langle (u, \mathbf{D}u) \rangle. \tag{17}$$

Proving the small worlds property for a given class of random networks thus consists in proving the limiting behavior

$$\lim_{N \rightarrow \infty} \ell = \lim_{N \rightarrow \infty} \binom{N}{2}^{-1} \sum_{i < j} \langle \mathbf{D}_{ij} \rangle \rightarrow \log N. \tag{18}$$

Depending on the particular realization of network randomness, the resulting distance matrix can prove to be disconnected. This means that \mathbf{D} can be brought into block diagonal form by a similarity transform involving a permutation matrix. The individual blocks will correspond to connected component graphs, and the off-diagonal block entries will be infinite. In such cases, it seems natural to define an average distance exclusively on the basis of sums performed within the distinct connected components.

There are a variety of algorithms for obtaining the distance matrix of a particular network given its adjacency matrix, for instance Dijkstra’s algorithm and its variants [41]. In the special case of an unweighted network, however, one can also obtain a formal expression for the distance matrix in terms of the adjacency matrix. Given that the k th power of the adjacency matrix has entries \mathbf{A}_{ij}^k that count the number of paths of length k between nodes i and j , the shortest distance between any two nodes i, j is given by the first power k for which $\mathbf{A}_{ij}^k \neq 0$. Therefore,

$$\mathbf{D}_{ij} = \min\{k \mid \mathbf{A}_{ij}^k > 0\} \tag{19}$$

An analytic approximation to the minimum function (exact in the limit $\Lambda \rightarrow \infty$) is given by

$$\mathbf{D}_{ij} = \lim_{\Lambda \rightarrow \infty} -\frac{1}{\Lambda} \log \sum_{q=1}^{\infty} \Theta(\mathbf{A}_{ij}^q - 1/2) e^{-\Lambda q} \tag{20}$$

where the offset of the Heaviside step function could be any real number between 0 and 1. In particular, we wish to calculate the ensemble expectation value of the above expression, that is

$$\langle \mathbf{D}_{ij} \rangle = \lim_{\Lambda \rightarrow \infty} -\frac{1}{\Lambda} \langle \log \sum_{q=1}^{\infty} \Theta(\mathbf{A}_{ij}^q - 1/2) e^{-\Lambda q} \rangle. \tag{21}$$

As it stands, Eq. (21) is calculationally cumbersome. It is tempting to bring the expectation value in Eq. (21) inside the logarithm by perhaps assuming that network disorder is “annealed”, or (by some abuse of terminology) invoking a variant of the replica trick [42, 43]. It is unclear if either assumption is justified. Moreover, with the exception of Edwards-Anderson type (Gaussian) disorder for weighted networks [43] [requiring the more general metric Eq. (16)], it is unclear if any calculational advantage is to be gained. Rather, in the next section we will suggest a way to approximate the above for the special case of unweighted networks using a particular variant of the mean field approximation.

3 Small Worlds—A Mean Field Description

In this section we consider a “mean field” description of Gilbert–Erdős–Renyi random networks. Although this description is not exact, it will be seen to provide a semi-quantitative description of their properties. For the case of unweighted networks, we see from Eq. (19) that the elements of the distance matrix \mathbf{D}_{ij} give the length of the shortest path between nodes i and j . This is simply the smallest integer k such that \mathbf{A}_{ij}^k is non-zero. It should be clear that no such shortest path can involve any repeated indices in the product of the k adjacency matrices since this would correspond to backtracking over a previous step. The determination of the distribution of shortest path lengths involves two steps. First, we must determine the probability that \mathbf{A}_{ij}^k is non-zero for a given value of k . This is given as one minus the probability that \mathbf{A}_{ij}^k vanishes subject to the constraint that there are no repeated indices in the intermediate summations within the matrix product. One simply computes the probability that the string $A_{i_1 i_1} A_{i_1 i_2} \dots A_{i_{k-1} i_k}$ vanishes, where $i_1 \dots i_{k-1}$ are summed over subject to the constraint that no index is repeated. From this, one sees immediately that the probability, p_k , that there will be a shortest, non-backtracking path⁹ of length k between the nodes i, j is given by

$$p_k = 1 - (1 - p^k)^{\gamma_k} \tag{22}$$

where

$$\gamma_k = (N - 2)! / (N - k - 1)! . \tag{23}$$

The probabilities p_k represent a true ensemble average for the unweighted network. The second part of the calculation requires us to modify the p_k by imposing the constraint that there is no path of length less than k that connects the nodes in question. Since it is difficult to impose this constraint exactly, we will proceed as follows. The ensemble average probability, q_k , that there does not exist a shortest path of length k is $1 - p_k$. Using this ensemble averaged quantity, the probability that the shortest path will be exactly of length k (and no shorter), denoted as $P(k)$ can be estimated as

$$P(1) = p, \text{ and } P(k) = p_k \prod_{n=1}^{k-1} q_n, \quad 1 < k < N. \tag{24}$$

We note that there is also some probability, P_∞ , that there will be no path of any length connecting two nodes. This probability is readily seen to be

$$P_\infty = \prod_{n=1}^{N-1} q_n. \tag{25}$$

As we shall see shortly, this quantity is directly relates to the size of the largest connected component (c.f. the quantity F_∞ in [23]). With the inclusion of this additional term (which can be 0), the probability distribution function (PDF) will have the correct normalization

$$\sum_{k=1}^{N-1} P(k) + P_\infty = 1. \tag{26}$$

⁹ In what follows, the adjective *shortest* should be taken to refer to shortest, non-backtracking paths. The subsequent expressions rely on the approximation that different shortest paths of length k represent independent events—as discussed further (cf. footnote 10), this relates to the assumptions that underlie our mean field approximation.

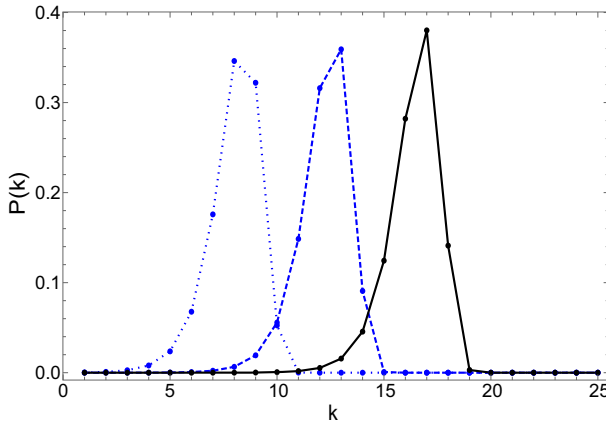


Fig. 1 Mean field values of $P(k)$ with $p = 3/(N - 1)$ for $N = 10^4$ (blue dotted), 10^6 (blue dashed), and 10^8 (black). Using Eq. (31) we find that $x_{\text{mode}} = 7.93, 12.12$ and 16.32 respectively, consistent with k_{mode} being identified by $\lfloor x_{\text{mode}} \rfloor_{\pm}$ as discussed below Eq. (31)

These results would be exact if the matrices \mathbf{A}^k for various values of k were uncorrelated. Unfortunately, this is not the case. The operation of ensemble averaging does not commute with other aspects of the computation and should be performed last. It is thus an approximation to make use of ensemble averaged values of q_k in the calculation of the $P(k)$, and it remains to be seen whether this approximation is reliable¹⁰. We note, however, the strong similarity of this approximation to the mean field approximation familiar from statistical mechanics and many-body physics. Mean field results can be useful provided that the system is homogeneous with correlations/fluctuations that are sufficiently small. Determination of the degree of reliability of mean field results in the present case will require comparison with simulations.

In the following examples of mean field results and their comparison with the results of simulations, we will consider the choice

$$p = \frac{\kappa}{N - 1}. \tag{27}$$

This is motivated by our wish to model systems for which the properties of the individual nodes (e.g., the average number of links per node, κ) are independent of the size of the system. Although more elaborate choices of p might be appropriate for specific systems, their adoption would require physical justifications that are beyond the scope of our investigation. The simulations were done over ensembles consisting of 10^2 samples, and converge rapidly away from the critical point.

Figure 1 shows the values of $P(k)$ determined by Eqs. (22)–(24) for several values of N . Consistent with small-worlds expectations, the peak of the distribution increases logarithmically with N and its width is roughly independent of N . Thus, relevant features of this distribution can be determined from a number of terms proportional to $\log N$. Although we have verified the validity of the mean field approximation for different values of N up to $N = 10^3$, note that it is very challenging to perform simulations for values of N as large as those shown in Fig. 1. This figure provides a strong indication of small world behavior for this choice of κ , although we shall see below that this result is not true in general.

¹⁰ It is in this sense that the treatment in this section amounts to a mean field approximation.

An analytic expression of how the mean field description captures these essential features of the distribution of path lengths can be obtained by considering the maximum value (i.e., the mode) of the $P(k)$ distribution for large N . In fact, this expression will also tell us when we are in the small-worlds domain. From Eq. (24) we see that

$$\frac{P(k+1)}{P(k)} = p_{k+1} \left(\frac{1}{p_k} - 1 \right). \tag{28}$$

The distribution will have an extremum whenever this ratio is equal to one. Imposing this condition implies that

$$1 = \left[2 - (1 - pp^k)^{\gamma_k(N-k-1)} \right] (1 - p^k)^{\gamma_k}. \tag{29}$$

Replacing k with the continuous variable x , using the analytic approximation $\Gamma(N + \alpha)/\Gamma(N) \rightarrow N^\alpha$ as $N \rightarrow \infty$, and using the fact that, for example,

$$(1 - p^{x+1})^{\gamma_{x+1}} \approx e^{-\gamma_{x+1}p^{x+1}} \text{ for } p < 1,$$

we see that Eq. (29) is equivalent to the condition

$$1 = \frac{1}{2} \left(e^{\frac{\tau}{k}} + e^{-\tau} \right) \text{ with } \tau = p\kappa^x = \frac{\kappa^{x+1}}{N-1}. \tag{30}$$

Numerically evaluating the root of the above for τ as a function of κ , and solving for x which we identify as x_{mode} results in¹¹

$$x_{\text{mode}}(\kappa, N) = \frac{1}{\log \kappa} \log(N-1) + \frac{1}{\log \kappa} \log r(\kappa) - 1, \quad k_{\text{mode}} = [x_{\text{mode}}]_{\pm} \tag{31}$$

where $r(\kappa)$ is the root of Eq. (30). Given that k is a discrete variable, whose continuous generalization x attains a maximum at some value, the most we can conclude is that k_{mode} is either of the nearest integers rounded up or down around the value obtained from Eq. (31), denoted $[x_{\text{mode}}]_{\pm}$. As seen in Fig. 1, this provides a reasonably accurate determination of the mode of the PDF given by Eq. (24). Note, however, the singular behaviour of k_{mode} when the average number of links per node is $\kappa \leq 1$. This is related to the fact that, for sufficiently small κ , the $P(k)$ decrease monotonically with k and no local maximum exists, due to the fragmentation of the network into disconnected components. Moreover, we note that Eq. (31) diverges as $1/(\kappa - 1)$ as $\kappa \rightarrow 1$ from above, in agreement with the divergence found for the mean path lengths in [31].

To see this more clearly, Fig. 2 shows the dependence of P_∞ on the average number of links per node, κ , as determined by the mean field result Eq. (25). As one passes through the critical edge probability $p_* = (\log N)/N$ (which we rederive shortly), mean field results suggest that the composition of the network undergoes a significant change – an expected result from the perspective of percolation theory [34–36]. For $p > p_*$, essentially all nodes are linked to one another in a genuine small world state. For $p < p_*$ the network exists in a phase composed of a multitude of clusters of small size. In such a phase, there will be a significant number of nodes that are not linked. In other words, the mean field approximation predicts that $P_\infty \rightarrow 1$ for $\kappa \ll 1$ and $P_\infty \rightarrow 0$ for $\kappa \gg 1$. The mean field values of P_∞ are shown in Fig. 2 as a function of κ for several values of N . We see from the figure that this transition becomes increasingly sharp with increasing N . It is tempting to regard P_∞ , or rather $(1 - P_\infty)^{1/2}$ as

¹¹ We note the formal similarity of Eq. (31) with Eqs. 1 and 3 in [44] which consider pairwise mean path lengths conditional on the pair of nodes in question having a particular degree, the function $r(\kappa)$ evidently quantifying the dependence under the identification $\kappa = \langle k \rangle$.

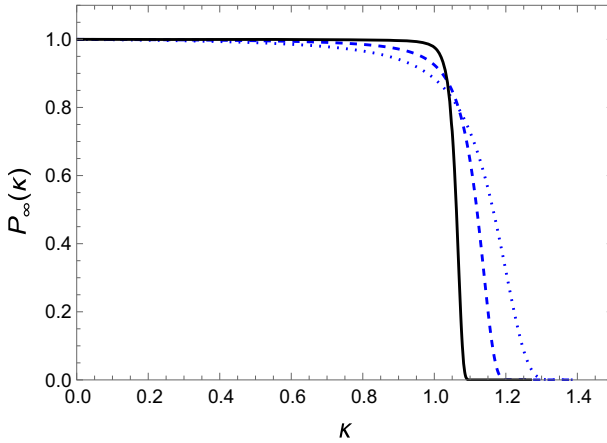


Fig. 2 P_∞ as a function of κ for $N = 100$ (dotted), 250 (dashed) and 2500 (solid)

an order parameter and to view this as a mean field indication of a phase transition at $\kappa = 1$. However, mean field descriptions of phase transitions are known to be incorrect for many systems (e.g., the Ising model in one and two dimensions) as a consequence of its failure to account for significant fluctuations near a critical point, rendering meaningless any simple averaged descriptions of the system. In fact, we are obliged to consider an entirely different quantity not obtainable from our mean field approximation (although well understood in the context of percolation theory [35, 36]) as the order parameter pertaining to the phase transition in Erdős–Renyi–Gilbert random networks. We identify this order parameter in the next subsection, and turn to the question of how to interpret the mean field approximation in addition to comparing with simulations.

3.1 Mean Fields and Simulations

Before comparing the mean field results with simulations, we first consider the probability that a given node is a “hermit” with no links to any of the remaining $(N - 1)$ nodes, which is one of the few results that can be determined exactly without the necessity of simulations. This probability is given by $(1 - p)^{N-1}$, or

$$\left(1 - \frac{\kappa}{N - 1}\right)^{N-1}, \tag{32}$$

which approaches $\exp[-\kappa]$ in the limit of large N . This is an exact statement valid for all κ and N . Evidently, the number of nodes which have no connection with the largest small-worlds cluster must be greater than or equal to the number of such hermits. The latter can be similarly estimated via the assumption that for large enough N , the fraction of nodes contained in the largest cluster does not depend on N (although it does depend on κ). Denoting this fraction as f_{cluster} , we realize that if we were to add one more node and if f_{cluster} is to remain unchanged, the probability that this new node not belong to the largest cluster must be $1 - f_{\text{cluster}}$. That is

$$\left(1 - \frac{\kappa}{N - 1}\right)^{Nf_{\text{cluster}}} = 1 - f_{\text{cluster}}, \tag{33}$$

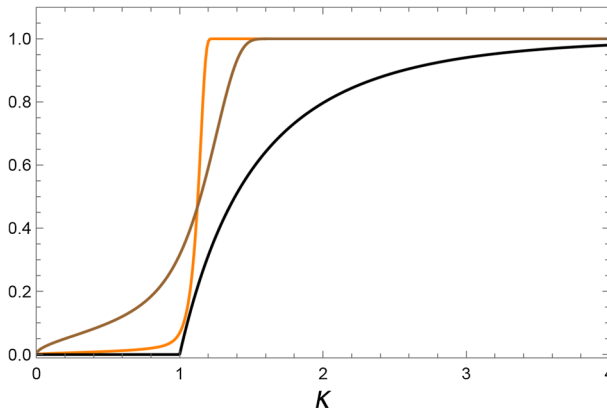


Fig. 3 Comparison of f_{cluster} (black curve) with the commensurate mean field order parameter $(1 - P_\infty)^{1/2}$ for $N = 10^2$ (brown) and $N = 10^4$ (orange)

which in the large N limit implies

$$f_{\text{cluster}} = 1 - \exp[-\kappa f_{\text{cluster}}]. \tag{34}$$

This is in fact the order parameter relevant to the percolation phase transition in Erdős–Rényi–Gilbert random networks [34–36]. Unfortunately, this is not obtainable from our mean field approximation, which focuses on the links of the network as opposed to the number of nodes in the largest cluster. We note that Eq. (34) has only the trivial solution $f_{\text{cluster}} = 0$ for $\kappa < 1$, whereas it also admits a non-trivial positive solution for $\kappa \geq 1$, which also depicts simulations of the largest cluster size. Eq. (34) has the formal solution [30]

$$f_{\text{cluster}} = 1 + \frac{1}{\kappa} W(-\kappa e^{-\kappa}), \tag{35}$$

where $W(z)$ is the Lambert W function¹². However, by considering the quantity $\tilde{f}_{\text{cluster}} = 1 - f_{\text{cluster}}$ and defining $\kappa = \log N$, one can rewrite Eq. (34) as

$$\tilde{f}_{\text{cluster}} = \frac{1}{N} N^{\tilde{f}_{\text{cluster}}}. \tag{36}$$

Starting with $\tilde{f}_{\text{cluster}} = 0$ and iterating the equation repeatedly, one obtains a power series expansion in $(\log N)/N$ and read off the general series coefficients to obtain the exact solution

$$\tilde{f}_{\text{cluster}} = \frac{1}{N} + \frac{1}{N} \sum_{k=1}^{\infty} \frac{(k+1)^{k-1}}{k!} \left(\frac{\log N}{N}\right)^k, \tag{37}$$

from which it is clear that $\tilde{f}_{\text{cluster}} \rightarrow 0$, or $f_{\text{cluster}} \rightarrow 1$ as $N \rightarrow \infty$.

In order to compare the percolation order parameter f_{cluster} with the mean field order parameter P_∞ , we realize that in the large N limit, the fraction of connected pairs of nodes is approximated to leading order by f_{cluster}^2 . Hence, the fraction that are not connected is given by $1 - f_{\text{cluster}}^2$. Thus, by comparing $(1 - P_\infty)^{1/2}$ to f_{cluster} , one can make a fairer determination as to whether the mean field order parameter accurately captures the nature of

¹² Defined as the root of the equation $w e^w - z = 0$, i.e. $w = W(z)$.

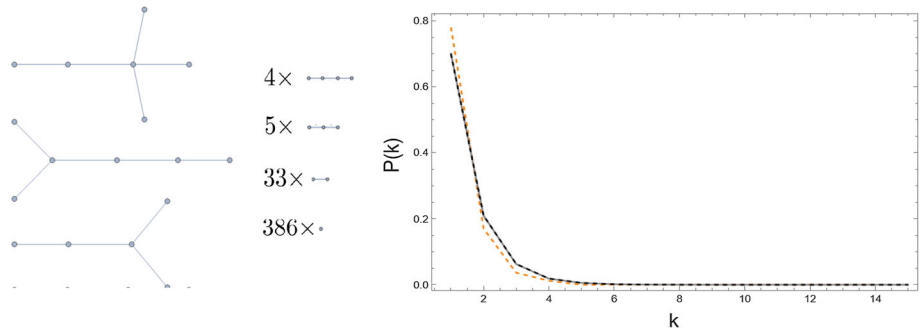


Fig. 4 Comparison of simulated network (orange dashed) with the normalized mean field approximation (gray) and the exact answer given by Eq. (38) (black dashed, indiscernible from mean field result) for $p = 0.3/(N - 1)$; $N = 500$. Graph on left is a representative draw from the ensemble

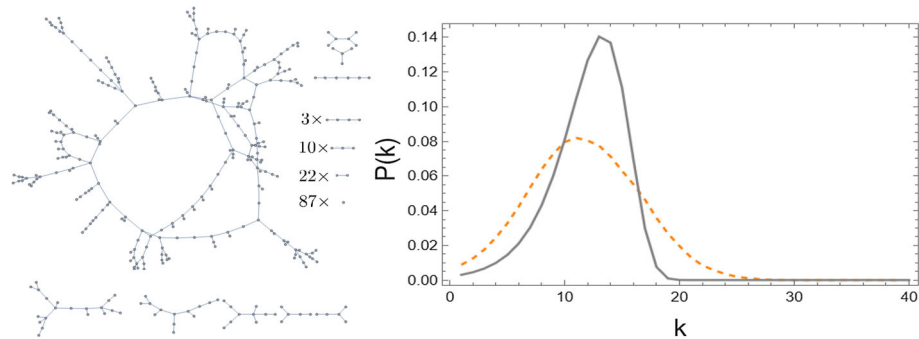


Fig. 5 Comparison of simulated network (orange dashed) with the mean field approximation (gray) for $p = 1.5/(N - 1)$; $N = 500$. Graph on left is a representative draw from the ensemble

the phase transition¹³. As can be seen from Fig. 3 we see that the mean field results appear to get the nature of the phase transition wrong by suggesting a first-order transition (given the evident discontinuity developing in the thermodynamic limit) rather than the second-order transition suggested by the analytic approximation for f_{cluster} . This failure of the mean field approximation is neither surprising nor alarming, and merely highlights its expected failure whenever fluctuations invalidate the mean field treatment. Nevertheless, the approximation ought to improve rapidly as one moves away from the critical point, which we now investigate numerically. It should come as no surprise that when the network consists of a giant connected component, the mean field approximation works well given the high dimensionality of system set by N .

We performed a series of simulations of the PDF for the quantities $P(k)$ that describe the probability that the shortest path between two nodes will be of length k . The results of such a check are shown in Figs. 4, 5 and 6 for an ensemble size of a 100 samples of a system of $N = 500$ nodes with κ of 0.3, 1.5, and 3.0, respectively. Here, we have chosen to neglect the single term P_∞ . Thus, the final term in Eq. (26) has been neglected and the remaining terms renormalized so that their sum is 1. For the cases $\kappa = 0.3$ and 3.0, the system is expected to be relatively homogeneous. In the former case, κ is simply too small for large clusters

¹³ The main shortcoming of our mean field approximation is that it fails to condition on the giant component. When one does so, it has been shown exactly that $(1 - P_\infty)^{1/2} = f_{\text{cluster}}$ (cf. Eq. 119 in [30]).

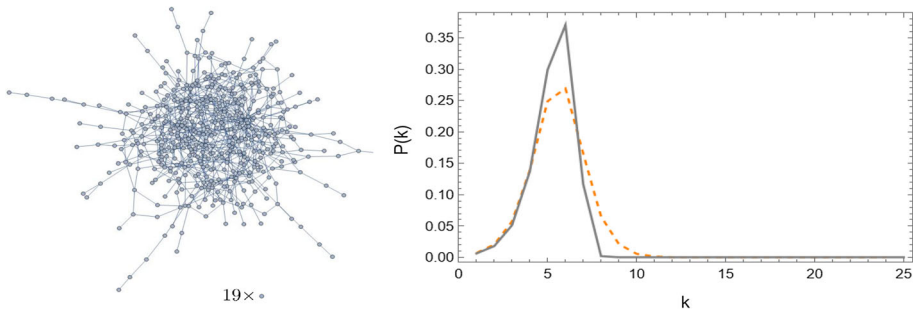


Fig. 6 Comparison of simulated network (orange dashed) with the mean field approximation (gray) for $p = 3.0/(N - 1)$; $N = 500$. Graph on left is a representative draw from the ensemble

and thus large distances, k , to form. Although P_∞ is not negligible, it has no effect other than the renormalization of the $P(k)$ shown in the figure. In the latter case, we have seen that the largest cluster is formed predominantly by the merger of smaller clusters. As can be seen from Eq. (32), all that remains is a small (i.e., $\exp[-\kappa]$) fraction of isolated nodes in the large N limit. In both cases, the resultant homogeneity of the networks is relatively high, and the agreement between mean field and simulated results is reasonably good. In fact, for the sub-critical value of $\kappa = 0.3$, the exact distribution of shortest paths is known to be exponential, and given by [31]

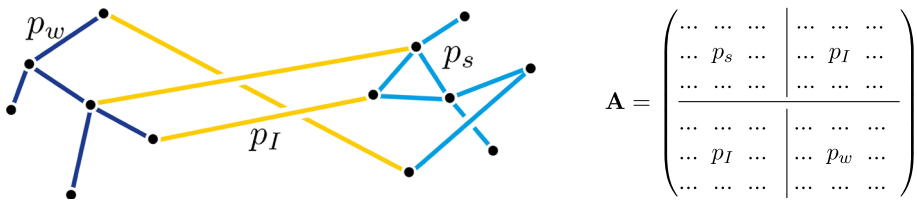
$$P(k) = (1 - \kappa)\kappa^{k-1}, \tag{38}$$

where we see from Fig. 4 that the mean field is in excellent agreement with the former. That the simulation differs is an artifact of the finite size of the simulated system.

The case shown in Fig. 5 is rather different since the value of $\kappa = 1.5$ has been chosen so that the largest cluster in the simulation contains half of the nodes. The ensemble of adjacency matrices is expected to be maximally heterogeneous with considerable differences between its various elements. In short, the mean field description should be at its worst. Specifically, we should expect that the root-mean-square (rms) deviation of the $P(k)$ should be relatively large and that the agreement with mean field results should be rather poor.

To conclude this section, we have presented a mean field description of the properties of an ensemble of randomly drawn adjacency matrices. This description is sufficiently simple to be applied to networks of size 10^7 or larger. It correctly describes the qualitative transition of the system from an set of unconnected nodes to one dominated by a single small-worlds connected network containing an increasing fraction of nodes as the number of links per node grows. This result was obtained by making the conservative assumption that the average number of links per node, κ , is independent of the size of the system, N . Any monotonic growth in κ as a function of N would necessarily result in a genuine small world involving the participation of all nodes in the thermodynamic limit $N \rightarrow \infty$. Since many popular models build in small worlds behaviour by construction, we consider it appealing that random networks show similar behaviour generically. As expected, the mean field description fails to capture the true nature of the small world phase transition. Nevertheless, its description of the distribution of the distances between nodes appears to be quite reliable for all values of κ considered here. Thus, it is with some confidence that we can now turn to the more challenging and potentially more interesting topic of small worlds and dimorphic networks.

4 Dimorphic Networks



We now consider a two family stochastic block model in which m out of a total of N nodes form links with each other with probability p_s while the remaining $N - m$ nodes form links with each other with probability p_w . A *planted partition* model would correspond to the situation where $p_s = p_w$, and the *balanced planted partition* model would correspond to the situation where furthermore, $m = N/2$ [45]. The probability that a link is formed between nodes in different families is p_I . That is, the network factorizes into two independent Erdős–Renyi–Gilbert sub-networks with associated probabilities p_s and p_w respectively, with bridge links between the sub-networks appearing with probability p_I . The subscripts s and w (strong and weak) are meant to suggest our eventual interest in the regime where one family is much more strongly connected than the other, although in what follows we allow them to be arbitrary¹⁴.

It is convenient to partition the index $i \in \{1, \dots, N\}$ into lower case Latin $a \in \{1, \dots, m\}$ for nodes in the ‘strong’ family and Greek $\alpha \in \{m + 1, \dots, N\}$ for nodes in the ‘weak’ family, so that the corresponding random adjacency matrix is given by

$$\begin{aligned}
 p(A_{ab} = 1) &= p_s \\
 p(A_{\alpha\alpha} = 1) &= p_I \\
 p(A_{\alpha\beta} = 1) &= p_w .
 \end{aligned}
 \tag{39}$$

The generalization of the mean field approximation to the dimorphic network consists of calculating the probability that the shortest paths between two nodes i, j will be of length k . For $k = 2$ this can be done straightforwardly:

$$\begin{aligned}
 p(A_{ab}^2 = 0) &= (1 - p_s^2)^{m-2} (1 - p_I^2)^{N-m} \\
 p(A_{a\beta}^2 = 0) &= (1 - p_s p_I)^{m-1} (1 - p_I p_w)^{N-m-1} \\
 p(A_{\alpha\beta} = 0) &= (1 - p_I^2)^m (1 - p_w^2)^{N-m-2} .
 \end{aligned}
 \tag{40}$$

In order to proceed for $k > 2$, it is useful to note that, for a string of adjacency matrix elements with the summation over the indices yet to be performed, the probability that this string vanishes is given by

$$p(A_{i_1 j_1} A_{j_1 j_2} \dots A_{j_{k-2} j_{k-1}} A_{j_{k-1} j} = 0) = 1 - p_s^{\alpha_0} p_w^{\alpha_1} p_I^{k - \alpha_0 - \alpha_1} \quad (\text{no summation over indices}) \tag{41}$$

where α_0 counts the number of times a pair of indices are both drawn from the set $\{a\}$ and α_1 counts the number of times a pair of indices are both drawn from the set $\{\alpha\}$. Therefore, the

¹⁴ Another variant of the dimorphic network with interesting applications would be to allow the links within and between the different families to have different weights. The formalism developed here can readily be generalized to this case and will be the subject of a follow up investigation.

probability that there is a shortest path of length k between any two nodes i, j is given by

$$p_{k,\{s,I,w\}} = 1 - \prod_{\alpha_0=0}^k \prod_{\alpha_1=0}^{k-\alpha_0} \left(1 - p_s^{\alpha_0} p_w^{\alpha_1} p_I^{k-\alpha_0-\alpha_1}\right)^{\Gamma_{\alpha_0\alpha_1}^{k,\{s,I,w\}}} \tag{42}$$

where the subscripts $\{s, I, w\}$ denote whether the pair of nodes i, j are both in set a , have mixed indices a, β or are both in α respectively, and $\Gamma_{\alpha_0\alpha_1}^{k,\{s,I,w\}}$ counts the number of distinct ways in which the combination α_0, α_1 appears in the sum without any repeated indices. The challenge is to calculate the combinatorial factors $\Gamma_{\alpha_0\alpha_1}^{k,\{s,I,w\}}$ (see Appendix C), to determine the quantities $q_{k,s} := 1 - p_{k,s}$ etc. and to construct the analogues of Eq. (24)

$$P_s(k) := p_{k,s} \prod_{j=1}^{k-1} q_{j,s} \quad 1 < k < N, \tag{43}$$

with similar expressions for $P_I(k)$ and $P_w(k)$. These correspond to the probability that there will be a shortest path of length k beginning and ending in the strong linked sub-network (P_s), beginning and ending in the weak linked sub-network (P_w) or beginning in one and ending in the other (P_I). (The construction of these three quantities—the dimorphic analogues of the mean field approximation to the $P(k)$ given by Eq. (24)—is presented in Appendix B.) As we note in the appendix, the calculation of the requisite combinatorial factors above maps onto the problem of computing the numbers of consecutively aligned spins on a finite spin chain with fixed boundary conditions. As we show in the next section, dimorphic networks evidently exhibit a rich phase structure. This includes a phase where small worlds behavior is induced in the weakly connected sub-network even if it is not itself within the small-worlds regime.

5 Phases of Small Worlds in Dimorphic Networks

In the following, we consider the dimorphic network to consist of a ‘strong’ sub-network with p_s well within what would be the small-world regime if it were an independent network and vary p_w and p_I . We find three distinct phases for the network which can be classified as disordered small worlds, spinodal small worlds, and induced small worlds. For each of these, we compare the results of simulations with our mean-field approximation, which as we shall see, remains accurate in these distinct phases provided the network as a whole does not decompose into too many disconnected components. In what follows, we consider the natural parameterization for dimorphic network probabilities

$$p_s := \frac{\kappa_s}{m}, \quad p_w := \frac{\kappa_w}{N - m}, \quad p_I := \frac{\kappa_I}{N - m}, \tag{44}$$

where, as above, the full network consists of N nodes, the strong sub-network contains m nodes, and the weak sub-network contains $N - m$ nodes.

5.1 Disordered Small Worlds

In this regime, we consider $\kappa_s \neq \kappa_w \neq \kappa_I$ such that individually, each $\kappa_i \gtrsim 2$, chosen to keep the network fragmentation to a minimum to ensure the validity of the mean-field approximation. In Fig. 7, we see that the mean-field approximation accurately compares to simulations. By varying N , one also finds the small world scaling of paths that begin and

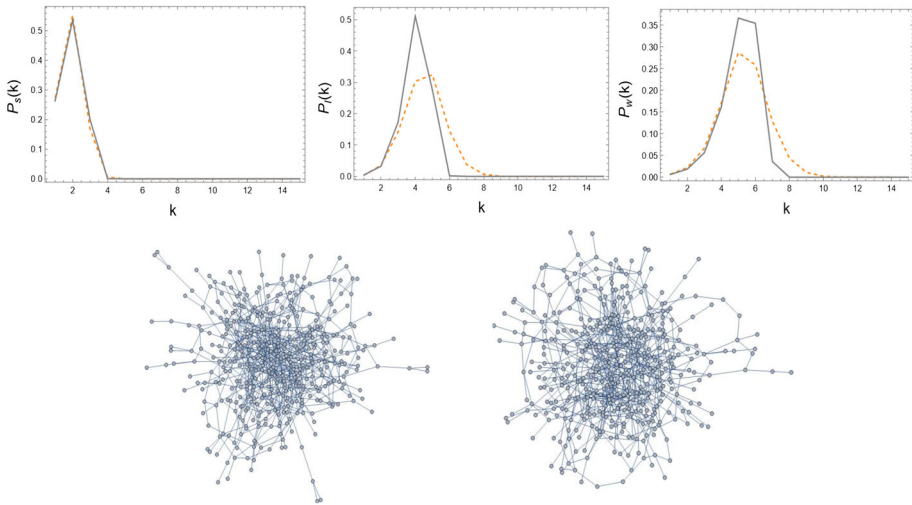


Fig. 7 Comparison of simulations (orange dashed) with the mean field approximation (gray): $P_S(k)$, $P_I(k)$ and $P_W(k)$ for $\kappa_S = 5$, $\kappa_W = 2$ and $\kappa_I = 3$ for $m = 20$, $N = 500$. Depicted below are two representative largest connected component draws from the simulation ensemble

end with a given sub-network, or beginning in one and ending in another. The distinction between the sub-networks is not particularly relevant in this regime, and one can consider this network globally as a disordered small world, as can readily be verified by increasing N and m .

5.2 Spinodal Small Worlds

In this regime, we consider $\kappa_S, \kappa_W \gtrsim 2$, with $\kappa_S, \kappa_W \gg \kappa_I > \kappa_{I,crit}$, where we have defined

$$\kappa_{I,crit} = \frac{1}{m}. \tag{45}$$

With these parameters, the network decomposes into two strongly connected sub-networks with a small number of bridge links between them. The critical value for κ_I can be understood as the value much below which the strong and weak sub-networks fail to form any links between them. Precisely at the critical value, any given realization of the network will be as likely as not to form bridge links between the strong and weak sub-networks. Were one to begin in the disordered small-worlds regime and decrease κ_I until it is much less than κ_S and κ_W , but still greater than $\kappa_{I,crit}$, the network would undergo a transition roughly analogous to spinodal decomposition in a fluid, wherein two phases co-exist and the transition occurs instantaneously due to the absence of any nucleation or latent energy difference between the two phases.

We again compare the results of our mean-field approximation to simulations in Fig. 8, where we find good agreement for the mean and modes of the distributions. By virtue of the bridge links, one finds small world scaling in the spinodal regime for each type of path – whether beginning or ending in a given sub-network, or beginning in one and ending in another, as depicted in Fig. 9. Here, the ratio m/N is kept fixed as N is increased.

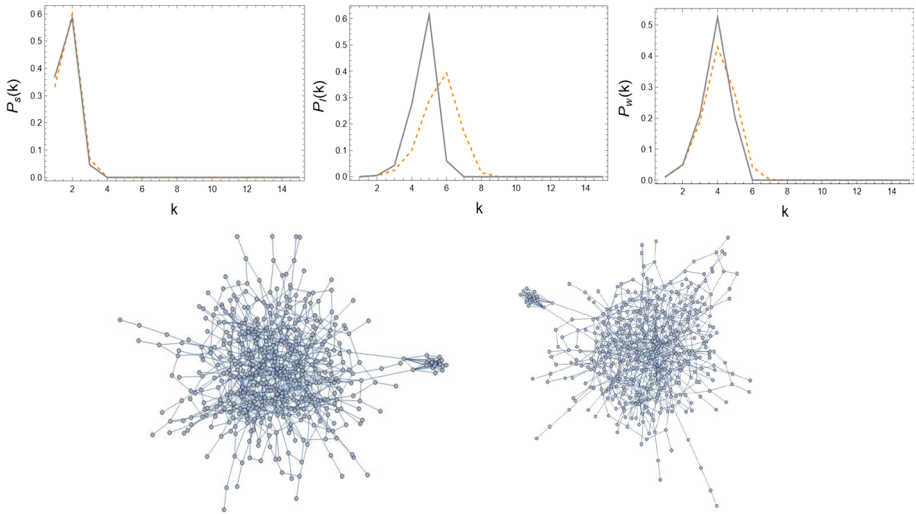


Fig. 8 Comparison of simulations (orange dashed) with the mean field approximation (gray): $P_S(k)$, $P_I(k)$ and $P_W(k)$ for $\kappa_S = 7$, $\kappa_W = 5$ and $\kappa_I = 0.2$ for $m = 20$, $N = 500$. Depicted below are two representative largest connected component draws from the simulation ensemble

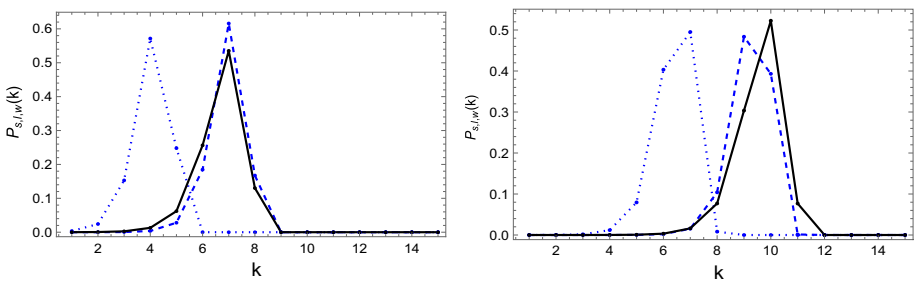


Fig. 9 Scaling of $P_S(k)$ (blue dotted), $P_I(k)$ (blue dashed) and $P_W(k)$ (black) for $\kappa_S = 7$, $\kappa_W = 5$ and $\kappa_I = 0.2$ for $m = 2 \times 10^3$, $N = 5 \times 10^4$ (left) and $m = 2 \times 10^5$, $N = 5 \times 10^6$ (right)

5.3 Induced Small Worlds

In this regime, we consider $\kappa_S \gtrsim 2$, $\kappa_W < 1$ $\kappa_I \gg 1$ and once more compare the results of our mean-field approximation to simulations. In Fig. 10, we see that even when the weak sub-network is not expected to be in the small world regime, the presence of bridge links with sufficient probability ($\kappa_I = 20$ in this example) connects the individual nodes of the weak sub-network to the strong sub-network such that small world scaling is recovered for nodes that begin and end in the weak sub-network. That is, small world behavior can be *induced* in sub-networks that, by themselves, are not in the small world regime. This is straightforward to understand intuitively – in the limit where the weak sub-network is completely disconnected, any bridge links will induce a connected sub-network out of the disconnected network. The average length for paths that begin and end in the connected part of the weak sub-network will thus be on average simply the average distance within the strong network + 2 for the hop back and forth. Although we require that the network not have too many disconnected components for the validity of the mean-field approximation, this simple reasoning does

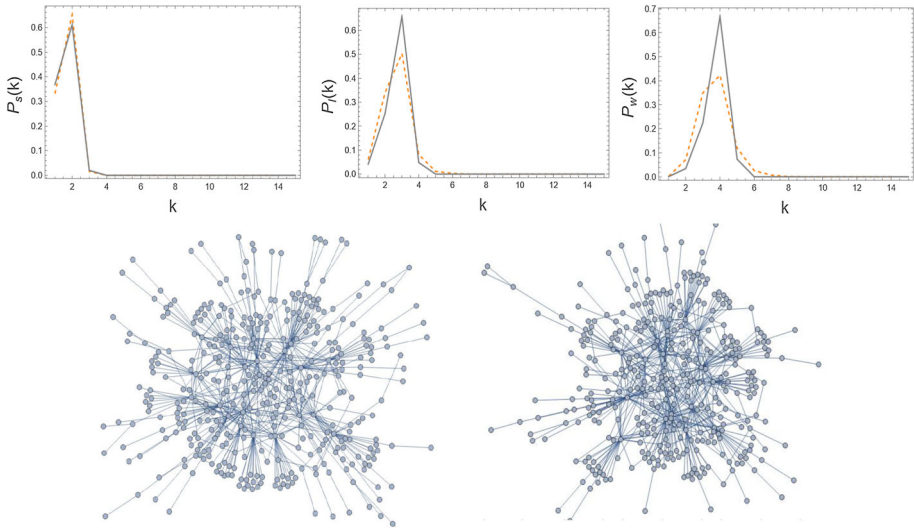


Fig. 10 Comparison of simulations (orange dashed) with mean field approximation (gray): $P_S(k)$, $P_I(k)$ and $P_W(k)$ for $\kappa_S = 7$, $\kappa_W = 0.3$ and $\kappa_I = 20$ for $m = 20$, $N = 500$. Depicted below are two representative largest connected component draws from the simulation ensemble

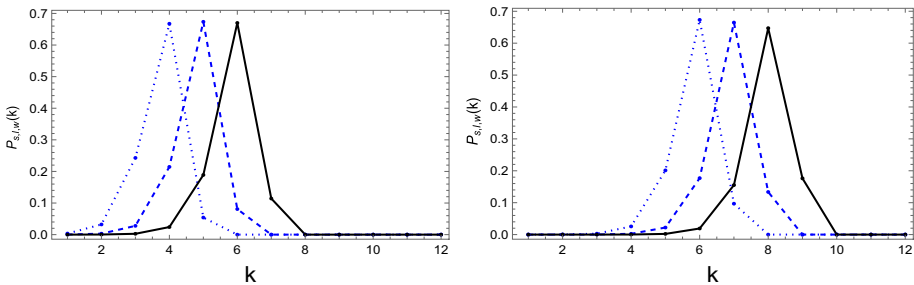


Fig. 11 Small world scaling of $P_S(k)$ (blue dotted), $P_I(k)$ (blue dashed) and $P_W(k)$ (black) for $\kappa_S = 7$, $\kappa_W = 0.3$ and $\kappa_I = 20$ for $m = 2 \times 10^3$, $N = 5 \times 10^4$ (left) and $m = 2 \times 10^5$, $N = 5 \times 10^6$ (right). N.B. the difference of +2 between the mode of the distributions for $P_W(k)$ and $P_S(k)$

seem to hold for parameter values consistent with the validity of the mean-field treatment, as seen in Fig. 11.

6 Discussion

In this paper, we introduced a novel mean-field approximation for random networks. The motivations for such an approximation is many-fold. A priori, one sees the immediate need for some sort of systematic approximation by first constructing the generating functional for arbitrary realizations of network randomness, from which ensemble averages of quantities of interest can be derived in principle. Even simple quantities of interest, such as the average of the elements of the distance matrix prove to be far too cumbersome to evaluate even in idealized examples. The motivation for the particular mean-field approximation introduced

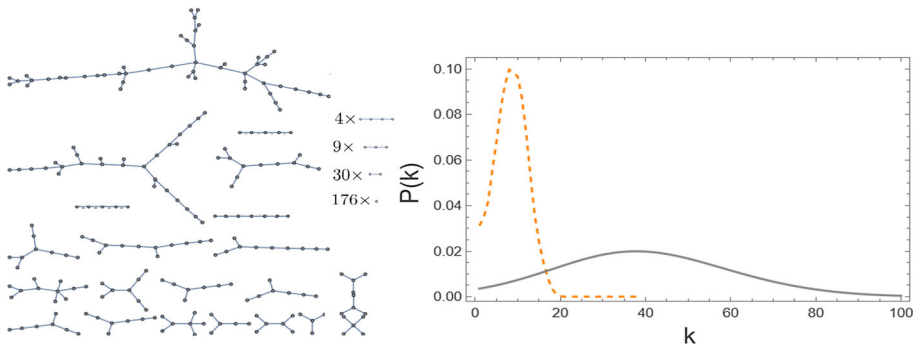


Fig. 12 Approaching maximal failure of the mean field approximation close to the critical point at $\kappa = 1$: Comparison of simulated network (orange dashed) with mean field (gray) for $p = 1.1/(N - 1)$; $N = 500$, along with representative draw from the graph ensemble

in Sect. 3 was to explore the scaling behavior of average network distances, and that it should also be applicable to a wider variety of contexts. That is, even in cases where an underlying lattice sub-structure can't be appealed to, or when we introduce disorder in varying forms.

The mean field approximation immediately yields a network order parameter $(1 - P_\infty)^{1/2}$ (Eq. 25), requiring suitable interpretation. The existence of a critical point at $p = 1/(N - 1)$ is manifest, however, as discussed in Sect. 3. One of the more predictable features of mean-field approximations its failure when ensemble fluctuations can no longer be neglected. As reasoned in Sect. 3, this is precisely the case as one approaches the phase transition, as depicted in Fig. 12,¹⁵ and in Figs. 4, 5 and 6.

The utility of the mean-field approximation is manifest in its generalizability to various forms of network disorder. The dimorphic variant of the stochastic block model studied in Sect. 4 exhibits a rich phase structure. It is intended as a model of network disorder where not all nodes are the same but can connect to other nodes with varying probabilities. The machinery introduced here suggests an immediate generalization to weighted networks where the dimorphism is extended to the weights and will be the subject of a followup study [46]. Furthermore, these results can in principle extend to stochastic block models with more than two families. As discussed in the appendix, this requires the calculation of combinatorial factors obtainable by constructing the generating function for auto-correlations over an n -symbol alphabet, where n is the number of families. Our ultimate aim, is to arrive at an analytic understanding of various features of a prototype for disordered Gilbert–Erdős–Renyi networks and to apply the insights gained to various biological, sociological and epidemiological systems.

Acknowledgements We extend our warmest thanks to Diego Garlaschelli for his detailed comments on the manuscript and his many useful suggestions for its improvement. We'd also like to thank our anonymous referees for helping us substantially improve our discussion and referencing.

Data Availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

¹⁵ As remarked in Sect. 3, the mode of our mean field approximation for the shortest path lengths diverges as $1/(\kappa - 1)$ as $\kappa \rightarrow 1$ from above, as found for the average path lengths in [31].

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: The Watts–Strogatz Model

The Newman–Watts variant of the Watts–Strogatz random network [14, 15] corresponds to a ring of nodes with links between every node and its *m*th nearest neighbors, with random shortcuts inserted between any two non-adjacent nodes with probability *p*. For the case *m* = 1, the corresponding probability measure is given by

$$\begin{aligned}
 (i + 2 \leq j) \quad P_{ij} &= p \delta(A_{ij} - 1) + (1 - p) \delta(A_{ij}) \\
 (i, j = i + 1) \quad P_{ij} &= \delta(A_{ij} - 1) \\
 (i = j) \quad P_{ij} &= \delta(A_{ii}) \\
 (i > j) \quad P_{ij} &= \delta(A_{ij} - A_{ji}) \quad (\text{Newman – Watts } m = 1) \quad (A1)
 \end{aligned}$$

where the indices are taken to be periodic. The above straightforwardly generalizes to Newman–Watts models with *m* nearest neighbors. The generating function for Newman–Watts random graphs with *k* = 1 follows from Eqs. (A1) and (6):

$$\begin{aligned}
 Z[J] &= \prod_i e^{-J_{(i+1)}} \prod_{i+1 < j} [pe^{-J_{(ij)}} + (1 - p)] \quad (\text{Newman – Watts } k = 1) \\
 &= e^{-\text{Tr}J(V+V^T)} (1 - p)^{\frac{(N-1)(N-2)}{2}} \prod_{i+1 < j} [1 + e^{-(\lambda+J_{(ij)})}] \quad (A2)
 \end{aligned}$$

where *V* is the *N*-dimensional permutation (or clock) matrix [47]. From this, we find that

$$W[J] = \frac{(N - 2)(N - 1)}{2} \log \left(\frac{1}{1 - p} \right) - \sum_{i+1 < j} \log [1 + e^{-(\lambda+J_{(ij)})}] + \text{Tr}J(V + V^T). \quad (A3)$$

Appendix B: Mean Field Approximation for Dimorphic Networks

The dimorphic network can be viewed as two separate component Gilbert networks with respective probabilities *p_s* and *p_w* for links within the component networks, with inter-component ‘bridge links’ with probability *p_l*. The adjacency matrix thus decomposes into blocks, with the probability that any element within the diagonal blocks being unity given by *p_s* and *p_w* respectively, with probability *p_l* for the off-diagonal blocks. We are thus interested in calculating

$$p \left(\sum_{j_1 \neq i, j} \sum_{j_2 \neq j_1, i, j} \dots \sum_{j_{k-1} \neq j_{k-2} \dots j_1, i, j} A_{ij_1} A_{j_1 j_2} \dots A_{j_{k-1} j} = 0 \right), \quad (B1)$$

with different choices for *i, j* yielding *p_{k,s}*, *p_{k,l}* and *p_{k,w}*. Given the block structure, it helps to rewrite the matrix products within the summations as matrix products over blocks, i.e. we

write

$$A_{ij} = A_{i_A j_B}^{AB} \tag{B2}$$

where A, B run over the binary indices $0, 1$ with $i_0 \subset \{1, \dots, m\}$ and $i_1 \subset \{m + 1, \dots, N\}$. Hence the summation in Eq. (B1) can be expressed as e.g. for $p_{k,s}$ as

$$A_{a_{j_1}}^{0A_{j_1}} A_{j_1 j_2}^{A_{j_1} A_{j_2}} \dots A_{j_{k-1} b}^{A_{j_{k-1}} 0} \tag{B3}$$

where the sum over the A_{j_i} indices is unrestricted whereas those of the corresponding lower indices do not have any of the j_i repeated. For $p_{k,I}$ and $p_{k,w}$ the first and last (unsummed) indices in the above are assigned the values $0, 1$ and $1, 1$ respectively. For any given values of the lower indices, we have $p(A_{ab}^{00} = 1) = p_s$, $p(A_{a\alpha}^{01} = 1) = p_I$, and $p(A_{\alpha\beta}^{11} = 1) = p_w$. Clearly, as one sums over blocks in Eq. (B3), the A_{j_i} indices run over values that when viewed as a consecutive string, scan over all the integers between 0 and 2^{k-1} represented as binary numbers. For $p_{k,s}$, we then affix 0 to the beginning and end of each binary string, from which it should be clear from Eq. (41) that α_0 in Eq. (42) corresponds to the number of consecutive 0 's, and α_1 corresponds to the number of consecutive 1 's in any given resulting binary string. Let's take $k = 3$ as an illustrative example—the summation over the block indices for $p_{k,s}$ results in

$$A_{a_{j_1}}^{0A_{j_1}} A_{j_1 j_2}^{A_{j_1} A_{j_2}} A_{j_2 b}^{A_{j_2} 0} = A_{aa_1}^{00} A_{a_1 a_2}^{00} A_{a_2 b}^{00} \rightarrow (1 - p_s^3)^{(m-2)(m-3)} \tag{B4}$$

$$+ A_{aa_1}^{00} A_{a_1 \alpha_2}^{01} A_{\alpha_2 b}^{10} \rightarrow (1 - p_s p_I^2)^{(m-2)(N-m)} \tag{B5}$$

$$+ A_{\alpha_1}^{01} A_{\alpha_1 a_2}^{10} A_{a_2 b}^{00} \rightarrow (1 - p_s p_I^2)^{(m-2)(N-m)} \tag{B6}$$

$$+ A_{\alpha_1}^{01} A_{\alpha_1 \alpha_2}^{11} A_{\alpha_2 b}^{10} \rightarrow (1 - p_I^2 p_w)^{(N-m)(N-m-1)} \tag{B7}$$

where we color coded A_{j_1} red and A_{j_2} as blue. As we can see, the summation over the block indices when viewed as a string run over $00, 01, 10$ and 11 , i.e. from 0 to 2^2 in binary. Affixing 0 's at the beginning and end of the string results in 0000 (three consecutive pairs of 0 's, so $\alpha_0 = 3, \alpha_1 = 0, k - \alpha_0 - \alpha_1 = 0$, cf. Eq. (B4)), 0010 and 0100 (one consecutive pairs of 0 's, so $\alpha_0 = 1, \alpha_1 = 0, k - \alpha_0 - \alpha_1 = 2$, cf. Eqs. B5 and B6) and 0110 (one consecutive pairs of 1 's, so $\alpha_0 = 0, \alpha_1 = 1, k - \alpha_0 - \alpha_1 = 2$, cf. Eqs. B7). The remaining summation over distinct combinations of the lower indices yields the exponents in Eqs. (B4)–(B7). Repeating the exercise affixing $0, 1$ and $1, 1$ to the beginning and end of the binary strings that run from $00 \rightarrow 11$ yields

$$q_{3,s} = (1 - p_s^3)^{\binom{m-2}{2}} (1 - p_s p_I^2)^{2 \binom{m-2}{1} \binom{N-m}{1}} (1 - p_s^2 p_w)^{\binom{N-m}{2}} \tag{B8}$$

$$q_{3,I} = (1 - p_s^2 p_I)^{\binom{m-1}{2}} (1 - p_s p_I p_w)^{\binom{m-1}{1} \binom{N-m-1}{1}} \\ \times (1 - p_I^3)^{\binom{m-1}{1} \binom{N-m-1}{1}} (1 - p_I p_w^2)^{\binom{N-m-1}{2}} \tag{B9}$$

$$q_{3,w} = (1 - p_I^2 p_s)^{\binom{m}{2}} (1 - p_I^2 p_w)^{2 \binom{m}{1} \binom{N-m-2}{1}} (1 - p_w^3)^{\binom{N-m-2}{2}} \tag{B10}$$

where $q_{k,s} := 1 - p_{k,s}$ etc., and where we've introduced the notation $\binom{a}{b} = a!/(a-b)!$. We can thus state the formal answer:

$$q_{k,s} = \prod_{\alpha_0=0}^k \prod_{\alpha_1=0}^{k-\alpha_0} \left(1 - p_s^{\alpha_0} p_w^{\alpha_1} p_I^{k-\alpha_0-\alpha_1} \right)^{\Gamma_{\alpha_0 \alpha_1}^{k,s} \binom{m-2}{n_s^0} \binom{N-m}{n_1^s}} \tag{B11}$$

$$q_{k,I} = \prod_{\alpha_0=0}^k \prod_{\alpha_1=0}^{k-\alpha_0} \left(1 - p_s^{\alpha_0} p_w^{\alpha_1} p_I^{k-\alpha_0-\alpha_1}\right)^{\Gamma_{\alpha_0\alpha_1}^{k,I} \left| \begin{matrix} m-1 \\ n_0^I \end{matrix} \right| \left| \begin{matrix} N-m-1 \\ n_1^I \end{matrix} \right|} \tag{B12}$$

$$q_{k,w} = \prod_{\alpha_0=0}^k \prod_{\alpha_1=0}^{k-\alpha_0} \left(1 - p_s^{\alpha_0} p_w^{\alpha_1} p_I^{k-\alpha_0-\alpha_1}\right)^{\Gamma_{\alpha_0\alpha_1}^{k,w} \left| \begin{matrix} m \\ n_0^w \end{matrix} \right| \left| \begin{matrix} N-m-2 \\ n_1^w \end{matrix} \right|} \tag{B13}$$

where the combinatorial problem has been reduced to computing the different $\Gamma_{\alpha_0,\alpha_1}^{k,\{s,I,w\}}$, which count the number times α_0 consecutive 0's and α_1 consecutive 1's appear in the set of 2^{k-1} strings formed by writing $0 \rightarrow 2^{k-1}$ in binary (with each number being a string of n_0 0's and n_1 1's) and affixing 0's to the beginning and end of this representation for $p_{k,s}$, 0 and 1, and 1's for $p_{k,I}$ and $p_{k,w}$ respectively. Note that when we set $N = m$ and $p_s = p_I = p_w$ we recover the answer for a single component network found in the previous section.

One can straightforwardly compute all the Γ 's via a recursion relationship derived in the next section of this appendix. Furthermore, the number of 0,1 indices summed over, given by n_0, n_1 are given by

$$n_0^s = \frac{1}{2} (k - 2 + \alpha_0 - \alpha_1) \tag{B14}$$

$$n_0^I = \frac{1}{2} (k - 1 + \alpha_0 - \alpha_1) \tag{B15}$$

$$n_1^w = \frac{1}{2} (k - 2 + \alpha_1 - \alpha_0) \tag{B16}$$

where in all cases, we have $n_0^{\{s,I,w\}} + n_1^{\{s,I,w\}} = k - 1$. As in Eq. (24), we are interested in computing

$$P_s(k) := p_{k,s} \prod_{j=1}^{k-1} q_{j,s} \quad 1 < k < N, \tag{B17}$$

and similarly for $P_I(k)$ and $P_w(k)$. That is, $P_s(k)$ corresponds to the probability that there will be a shortest path of exactly length k that begins and ends in the strong linked sub-network, $P_w(k)$ corresponds to the same but for a path that begins and ends in the weak linked sub-network, and $P_I(k)$ for paths that begin in one and ends in the other sub-network.

Appendix C: Multiplicity Factors for Dimorphic Networks and Spin Chains

The $\Gamma_{\alpha_0,\alpha_1}^{k,\{s,I,w\}}$ relate to an interesting question in the context of spin chains¹⁶—it counts the number of ways one can have α_0 pairs of consecutive spin pointing up and α_1 pairs of consecutive spins pointing down on a chain of k spins, with the boundary spins with specific orientations. In the present context, it relates to the quantities $z_s(n, l, m)$ defined as the number of ways in which l pairs of consecutive zeroes and m pairs of consecutive ones appear in a string of n bits that begins and ends with 0, and $z_l(n, l, m)$ defined as the number of ways in which l pairs of consecutive zeroes and m pairs of consecutive ones appear in a string of n bits that begins with 1 and ends with 0. One could have defined a similar quantity for $z_w(n, l, m)$, where the string begins and ends with 1, but we notice that since flipping

¹⁶ What follows is a straightforward extension of the results contained in [48].

all the bits is an automorphism, it must be the case that $z_w(n, l, m) = z_s(n, m, l)$. Defining $z(n, l, m) = 0$ for $l + m \geq n$ and for all $n, l, m < 0$, we note that the following recursion relations hold:

$$z_s(n + 1, l, m) = z_s(n, l - 1, m) + z_s(n - 1, l, m) + z_l(n - 1, l, m - 1) \quad (\text{C1})$$

$$z_l(n + 1, l, m) = z_s(n, l, m) + z_s(n - 1, l, m - 1) + z_l(n - 1, l, m - 2) \quad (\text{C2})$$

For Eq. (C1), consider a string of length $n + 1$ with l pairs of consecutive 0's and m pairs of consecutive 1's. The total number of these define the l.h.s. above. Consider now a given string that satisfies this criteria. Either it begins with 00 or 01. In the first case, it consists of a 0, followed by an n string beginning and ending with 0, with $l - 1$ pairs of consecutive 0's and m pairs of consecutive 1's (the number of these is the first term on the rhs of Eq. (C1)). In the second case, it either begins with a 010 or a 011. If it begins with 010, it consists of 01, followed by an $n - 1$ string beginning and ending with 0 with l pairs of consecutive 0's and m pairs of consecutive 1's, the number of which is the second term on the rhs of Eq. (C1). If it begins with a 011, it consists of 01, followed by an $n - 1$ string beginning with 1 and ending with 0 with l pairs of consecutive 0's and $m - 1$ remaining pairs of consecutive 1's, the number of which is the third term on the rhs of Eq. (C1). Since this exhausts the possibilities, the equality follows. A similar argument results in Eq. (C2). It follows that

$$\Gamma_{\alpha_0, \alpha_1}^{k, s} = z_s(k + 1, \alpha_0, \alpha_1) \quad (\text{C3})$$

$$\Gamma_{\alpha_0, \alpha_1}^{k, l} = z_l(k + 1, \alpha_0, \alpha_1) \quad (\text{C4})$$

$$\Gamma_{\alpha_0, \alpha_1}^{k, w} = z_s(k + 1, \alpha_1, \alpha_0) \quad (\text{C5})$$

One can generate all the appropriate degeneracy factors¹⁷ by straightforwardly iterating from the case for $k = 2, 3$.

References

1. Travers, J., Milgram, S.: An experimental study of the small world problem. *Sociometry* **32**, 425 (1969)
2. Ugander, J., Karrer, B., Backstrom, L., Marlow, C.: The anatomy of the Facebook social graph, arXiv e-prints, [arXiv:1111.4503](https://arxiv.org/abs/1111.4503) (2011)
3. Backstrom, L., Boldi, P., Rosa, M., Ugander, J., Vigna, S.: Four degrees of separation, arXiv e-prints, [arXiv:1111.4570](https://arxiv.org/abs/1111.4570) (2011)
4. Newman, M.E.: Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002)
5. Keeling, M.J., Eames, K.T.: Networks and epidemic models. *J. R. Soc. Interface* **2**, 295 (2005)
6. Miller, J.C.: Percolation and epidemics in random clustered networks. *Phys. Rev. E* **80**, 020901 (2009)
7. Uzzi, B., Spiro, J.: Collaboration and creativity: the small world problem. *Am. J. Sociol.* **111**, 447 (2005)
8. Davis, G.F., Yoo, M., Baker, W.E.: The small world of the American Corporate Elite, 1982–2001. *Strategic Organization* **1**, 301 (2003)
9. Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64**, 026118 (2001)
10. Erdős, P., Rényi, A.: On random graphs I. *Publicationes Mathematicae Debrecen* **6**, 290 (1959)
11. Gilbert, E.N.: Random graphs. *Ann. Math. Stat.* **30**, 1141 (1959)
12. Barabási, A.-L., Albert, R.: Emergence of scaling in random networks. *Science* **286**, 509 (1999)
13. Cohen, R., Havlin, S.: Scale-free networks are ultrasmall. *Phys. Rev. Lett.* **90**, 058701 (2003)
14. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. *Nature* **393**, 440 (1998)

¹⁷ In principle, there should also be a generating function for these symmetry factors by considering the problem of word auto-correlations over a two-symbol alphabet. This suggests that the generalization of the dimorphic network to a multimorphic setup, i.e. to a stochastic block model with more than two families can be obtained by determining these generating functions.

15. Newman, M.E.J., Watts, D.J.: Renormalization group analysis of the small-world network model. *Phys. Lett. A* **263**, 341 (1999)
16. Dorogovtsev, S., Mendes, J., Samukhin, A.: Metric structure of random networks. *Nucl. Phys. B* **653**, 307 (2003)
17. van der Hofstad, R., Hooghiemstra, G., Van Mieghem, P.: Distances in random graphs with finite variance degrees, arXiv Mathematics e-prints, math/0407092 (2004)
18. van den Esker, H., van der Hofstad, R., Hooghiemstra, G.: Universality for the distance in finite variance random graphs. *J. Stat. Phys.* **133**, 169 (2008)
19. Nitzan, M., Katzav, E., Kühn, R., Biham, O.: Distance distribution in configuration-model networks. *Phys. Rev. E* **93**, 062309 (2016)
20. Melnik, S., Gleeson, J.P.: Simple and accurate analytical calculation of shortest path lengths, arXiv preprint [arXiv:1604.05521](https://arxiv.org/abs/1604.05521) (2016)
21. Shao, J., Buldyrev, S.V., Cohen, R., Kitsak, M., Havlin, S., Stanley, H.E.: Fractal boundaries of complex networks. *Europhys. Lett.* **84**, 48004 (2008)
22. Shao, J., Buldyrev, S.V., Braunstein, L.A., Havlin, S., Stanley, H.E.: Structure of shells in complex networks. *Phys. Rev. E* **80**, 036105 (2009)
23. Katzav, E., Nitzan, M., Ben-Avraham, D., Krapivsky, P., Kühn, R., Ross, N., Biham, O.: Analytical results for the distribution of shortest path lengths in random networks. *Europhys. Lett.* **111**, 26006 (2015)
24. Bonneau, H., Hassid, A., Biham, O., Kühn, R., Katzav, E.: Distribution of shortest cycle lengths in random networks. *Phys. Rev. E* **96**, 062307 (2017)
25. Steinbock, C., Biham, O., Katzav, E.: Distribution of shortest path lengths in a class of node duplication network models. *Phys. Rev. E* **96**, 032301 (2017)
26. Steinbock, C., Biham, O., Katzav, E.: Analytical results for the distribution of shortest path lengths in directed random networks that grow by node duplication. *Eur. Phys. J. B* **92**, 1 (2019)
27. Goldental, A., Vardi, R., Sardi, S., Sabo, P., Kanter, I.: Broadband macroscopic cortical oscillations emerge from intrinsic neuronal response failures. *Front. Neural Circ.* **9**, 65 (2015). <https://doi.org/10.3389/fncir.2015.00065>
28. Erdos, P., Renyi, A.: On the evolution of random graphs. *Publ. Math. Inst. Hungary. Acad. Sci.* **5**, 17 (1960)
29. Erdős, P., Rényi, A.: On the strength of connectedness of a random graph (1964)
30. Tishby, I., Biham, O., Katzav, E., Kühn, R.: Revealing the microstructure of the giant component in random graph ensembles. *Phys. Rev. E* **97**, 042318 (2018)
31. Katzav, E., Biham, O., Hartmann, A.K.: Distribution of shortest path lengths in subcritical Erdős-Rényi networks. *Phys. Rev. E* **98**, 012301 (2018)
32. Newman, M.E.J., Moore, C., Watts, D.J.: Mean-field solution of the small-world network model. *Phys. Rev. Lett.* **84**, 3201–3204 (2000)
33. Song, H.F., Wang, X.-J.: Simple, distance-dependent formulation of the Watts-Strogatz model for directed and undirected small-world networks. *Phys. Rev. E* **90**, 062801 (2014)
34. Molloy, M., Reed, B.: A critical point for random graphs with a given degree sequence. *Random Struct. Algorithms* **6**, 161 (1995). <https://doi.org/10.1002/rsa.3240060204>
35. Molloy, M., Reed, B.: The size of the giant component of a random graph with a given degree sequence. *Comb. Probab. Comput.* **7**, 295 (1998)
36. Newman, M.: *Networks*. Oxford University Press, Oxford (2018)
37. Park, J., Newman, M.E.J.: Statistical mechanics of networks. *Phys. Rev. E* **70**, 066117 (2004)
38. Squartini, T., Mastrandrea, R., Garlaschelli, D.: Unbiased sampling of network ensembles. *New J. Phys.* **17**, 023052 (2015)
39. Cimini, G., Squartini, T., Saracco, F., Garlaschelli, D., Gabrielli, A., Caldarelli, G.: The statistical physics of real-world networks. *Nat. Rev. Phys.* **1**, 58 (2019)
40. Opsahl, T., Agneessens, F., Skvoretz, J.: Node centrality in weighted networks: generalizing degree and shortest paths. *Soc. Netw.* **32**, 245 (2010)
41. Dijkstra, E.W.: A note on two problems in connexion with graphs. *Numerische Mathematik* **1**, 269 (1959)
42. Mezard, M., Parisi, G., Virasoro, M.A.: *Spin Glass Theory and Beyond: An Introduction to the Replica Method and its Applications*. World Scientific Publishing Company, Singapore (1987)
43. Edwards, S.F., Anderson, P.W.: Theory of spin glasses. *J. Phys. F* **5**, 965 (1975)
44. Holyst, J.A., Sienkiewicz, J., Fronczak, A., Fronczak, P., Suchecki, K.: Universal scaling of distances in complex networks. *Phys. Rev. E* **72**, 026108 (2005)
45. Amini, A.A., Levina, E., et al.: On semidefinite relaxations for the block model. *Ann. Stat.* **46**, 149 (2018)
46. Jackson, A.D., Patil, S.P.: In preparation (2022)
47. Brualdi, R.A.: *Combinatorial Matrix Classes*, vol. 13. Cambridge University Press, Cambridge (2006)
48. Dover, J.M.: Some notes on pairs in binary strings, arXiv e-prints, [arXiv:1609.00980](https://arxiv.org/abs/1609.00980) (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.