Abstract

The associations between interacting components of complex systems, in many cases, are influenced by the spatial embedding of the components. Systems like these have an intrinsic dependence on space and, by extension, so must the topology of the corresponding network used to encode the interactions of the system. Despite this, few studies have attempted to attribute a space to empirical or synthetic network data, or investigated if alternative methods to maximum-likelihood estimation can be used to do so. We therefore investigate the problem of reconstructing the spatial structure of synthetic spatial networks, given only their topology, to ascertain a verified method that is applicable to empirical networks. We detail an alternative constrained optimisation technique — Simulated Annealing (adapted for our purposes) and show that we are able to reconstruct the underlying spatial structure of a specific class of range-dependent random networks with a high degree of success over a significant portion of the associated parameter space. Application to an empirical network attributes a space that coincides with results from established community-detection algorithms, and highlights potential reasons for the failure of these algorithms. The results suggest that the attributed spatial structures may be meaningful in the same way community-detection algorithms are meaningful, and could be favourable in that they exhibit, on an individual vertex level, a strength of affiliation to communities. The verification of our method opens the potential to explore other complex systems using simulated annealing and we expect that it may be possible to identify spaces associated with other mesoscale properties of networks using the method proposed herein.

Contents

1 INTRODUCTION

1.1 Network Models

Networks are prevalent throughout the physical world and are comprised of vertices and the connections between them — edges. Vertices represent distinct elements such as people or neurons. The associated edges may be friendships or synapses which, when combined with the corresponding vertices, form networks. The topological structures of certain biological, physical and social systems can become tractable when expressed in this way.

Many inherent properties of networks can be mathematically described explicitly. Properties such as mean vertex degree, defined as the average number of edges connected to a vertex, and network density, defined as the ratio of the number of existing edges to the number of possible edges, are calculable and can be used to quantitatively compare and contrast different networks. Other properties can be more difficult to quantify, and analysis often relies on the careful construction of network models [1].

Network models can be fundamental to understanding empirical observations of both complex real-world networks and theorised networks; models able to generate networks enable the synthesis of many networks that are amenable to analysis. These models can be defined intentionally to replicate properties of complex real-world networks. If we are aware as to the mechanism in which a network was generated, we can use models to approximate certain gaps in in collected data. If we can synthesise networks that emulate connection patterns in empirical systems, then analysis carried out on these synthesised networks is likely relevant to understanding the implications of such patterns [1]. For instance, to investigate the spread of a disease in a human population, we can generate synthetic networks to reflect such a population and simulate contagion across these networks [2]. Network models in this instance can uncover the traits of a network that are significant in influencing the spread or control of a disease.

Random networks are model networks in which certain parameters are fixed but other aspects of the network are subject to chance. The mathematicians Paul Erdős and Alfréd Rényi introduced a simple model for generating random networks known as the Erdős–Rényi (ER) model [3]. The ER model constructs an ER network by connecting a fixed number of vertices with edges at random. Each potential edge exists with probability p independent from every other edge. Strictly, the ER network is defined as an ensemble of networks — that is to say, it is defined in terms of a probability distribution over all possible networks that could exist under the model. ER model networks can be characterised by two parameters, the number of vertices of the network n and edge probability p.

Although random networks can be useful for understanding networks [4, 5], evidence suggests that many real-world networks are more [1, 6, 7] clustered than random networks such as the ER network. Clustering is a measure of the tendency of adjacency (direct connection) to be transitive in a network. Where transitivity implies that if there is an edge connecting

v_i and v_j and an edge connecting v_j and v_k then there exists an edge between v_i and v_k. High clustering is familiar when considering social networks, which tend to contain cliques or sub-communities of friends or acquaintances [1] (where transitivity implies that acquaintances of an acquaintance are also acquaintances to a given individual). High clustering is one of two properties that constitute a widely discussed phenomena in network theory known as the *small-world effect* [7]. In addition to high clustering, small-world networks exhibit small network diameters as a function of the total number of vertices they contain. More specifically, for any two vertices in a small-world network the expectation of path length between them is proportional to the logarithm of the total number of vertices in the network. There are strong implications of the properties of small-world networks. Again, in the context of contagion, the small-world effect can help to explain epidemics [8, 9, 10]. It is clear a disease will have a much greater chance of transmission to any given individual if the path length between any two individuals is low.

Network models have been developed that can replicate small world properties, notably the Watts–Strogatz model and range-dependent random networks (over a specific parameter range [11]), both of which we detail and investigate. Much research has followed the development of such models, yet resulting publications "usually discuss spatial aspects of networks very briefly" [6], despite numerous networks of interest bearing an inherent spatial structure. The associations between interacting components of complex systems, in many cases, are influenced by the spatial embedding of the components [6, 12]. Systems like these have an intrinsic dependence on space and, by extension, so must the topology of the corresponding network used to encode the interactions of the system. The networks therefore can be described as spatial, a group of networks described in depth in a review by Marc Barthélemy [6].

1.2 Spatial Networks

The concept of a spatial network is perhaps most familiar when considering geographic spatial networks, in which vertices, such as cities, power-plants or components on a circuit board, are clearly fixed (embedded) at distinct geographical locations. Barthélemy points out that "there is a cost associated with the length of edges" in spatial networks "which in turn has dramatic effects on the topological structure of these networks". A cost could be monetary, as in the case of a power grid where the price associated with a power-line (edge) between a plant (vertex) and a city (vertex) may increase with distance, since it might be reasonable to assume a longer power line will have greater construction and maintenance costs associated with it.

We would expect that more costly edges in a network may be compensated by some advantage, such as costly connections between well connected vertices (hubs). The existence and proportion of edges such as this ultimately depend on the system in question.

Physical distance is easy to recognise as a cost associated with length of an edge when considering geographic spatial networks like a powergrid. For less tangible networks, a physical distance can be substituted by any notion able to attribute a cost to an edge. Social distance in a network could be parameterised by political affiliation or age for instance. The former example translates well to a one–dimensional space — often regarded as a spectrum. We might expect that individuals can be arranged on a political spectrum according to there values and opinions. Newman states that "people have, it appears, a strong tendency to associate with others whom they perceive as being similar to themselves in some way" a tendency referred to as *assortative mixing* [1, 13]. Thanks to assortative mixing, we might expect more social contact between people who are closer on the political spectrum. Assortative mixing therefore suggests that it might be possible to predict where individuals associated with political parties lie in relation to one another on such a spectrum, based on interaction data alone. Notice that this attributes a spatial structure to a social contact network, despite the fact the individuals are not embedded in a physical space.

Spatially-Embedded Random Network Models

Spatially-embedded random network (SERN) models generate synthetic random networks in which vertices are embedded in some spatial domain. In the interest of subsequent analysis, we detail methods to generate two different classes of SERNs, range-dependent random networks (RDRN) and Watts–Strogatz (WS) networks.

Range–Dependent Random Networks

A range-dependent random network (RDRN) is a simple spatial network in which the probability of an edge existing between any pair of vertices is a function of the distance between the vertices only [11]. Generating an RDRN requires the definition of some probability distribution for the existence of an edge across possible inter-vertex ranges. We define the inter-vertex ranges $r = d(i, j)$ as the distances measured by an arbitrary distance function $d(i, j)$ between every pair of vertices i and j. These vertices could be embedded anywhere in some sort of metric space. Mathematically, the probability of an edge existing between two vertices at the arbitrary indices (i, j) in a RDRN can be specified by some function

$$p_{ij} = p(d(i, j)) = p(r) \tag{1}$$

The research we carry out concerning RDRNs assumes vertices that are located at the points $(1, 2, 3, ..., n)$ in one–dimensional Euclidean space, resulting in the distance function $d(i, j) = |i - j|$. These RDRNs assume a power law [11] for the probability of the existence of an edge between two vertices defined as

$$p_{ij} = p(r) = \alpha\lambda^{r-1}. \tag{2}$$

RDRNs generated according to (2) are a specific subclass of RDRNs, however, the term RDRN hereafter refers to this specific subclass, unless an alternative $p(r)$ is given. Figure 1

shows that (2) decays exponentially with increasing range (in the strict interval $\alpha, \lambda \in (0, 1)$). Here, α linearly scales every edge probability, whilst λ promotes longer range connections.

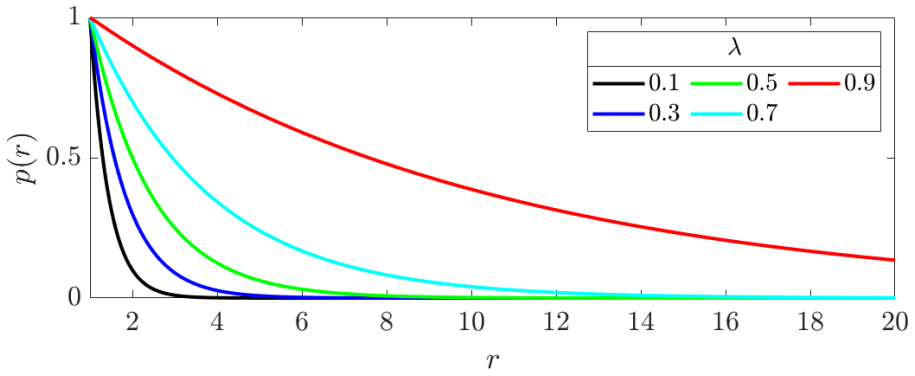


Figure 1: The effect of λ on edge probability $p(r)$ against inter-vertex range r, with $\alpha = 1$.

The Erdős-Rényi Network

Erdős–Rényi (ER) networks are in fact a degenerate case of an RDRN with $\lambda = 1$. This corresponds to a uniform edge probability of α for all possible inter-vertex ranges. For $\lambda = 1$ a RDRN has no range dependence, and is simply an ER network.

The Watts–Strogatz Network

An ensemble of Watts–Strogatz (WS) networks [7] can be specified by three parameters, the number of vertices n the mean degree z (always taken to be a positive even integer) and another parameter called the rewiring probability p, explained here. A WS network is generated by forming a regular ring lattice, with n cyclically embedded vertices, each of which connected with an edge to its $z/2$ nearest neighbours on either side (resulting in a mean vertex degree of z). Following the formation of a ring lattice, we delete each edge with probability p independent from every other edge. Next, for any pair of vertices in the resulting network where there does not exist a direct connection (edge) these vertices are connected with probability q. The parameter q is not essential in the initial specification of the network ensemble, as it depends directly on the other three in order to preserve the mean degree after the rewiring process has taken place. We choose q to ensure that the expected mean degree of the final WS network equals the original mean degree z of the lattice. We achieve this by matching the number of edges we expect to delete with the number of edges we expect to add

$$pzn = q\left[\frac{n(n-1)}{2} - zn + pzn\right]. \tag{3}$$

Here pzn equals the number of edges in the initial lattice we expect to delete, $n(n-1)/2$ is the number of potential edges in the network and zn is the number of edges in the initial

lattice. For networks that are not densely connected (where $n \gg z$)

$$q \approx \frac{2pz}{n}. \tag{4}$$

Instead of constructing a ring lattice, we can also consider a modification of the WS network based on a line of vertices, instead of a ring. Here we begin with a regular one dimensional lattice of n vertices each of which connected with an edge to its $z/2$ nearest neighbours (where these neighbours exist) and follow the same procedure thereafter. It is interesting to note that both the circular and one-dimensional versions of the Watts–Strogatz ensemble of networks are identical in the case of infinite vertices as a line can be defined in terms of a circle of infinite radius.

In the case of finite vertices the one-dimensional case produces 'end' vertices - the outermost vertices of the network, with a different degree to the rest in the initial lattice. These end vertices lead to topologically non-equivalent networks to those constructed cyclically. Figure 2 shows the edge probabilities over inter–vertex ranges for a WS network, showing that a WS network could in fact be considered as a RDRN with a piecewise edge probability function

$$p(r) = \begin{cases} 1-p & \text{if } 2r \leq z, \\ q & \text{otherwise.} \end{cases}$$

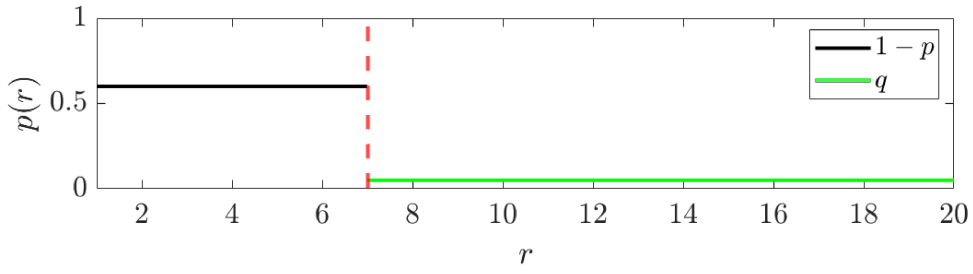


Figure 2: The effect of z on edge probability $p(r)$ against inter-vertex range r for a Watts-Strogatz network with $z = 14$, $p = 0.4$ and $n = 168$.

1.3 Network Analysis

Regardless of how we acquire a network topology, we can calculate or measure numerous quantities and properties in order to uncover particular features of the network to help to better understand it. The term *centrality* [1] embodies a multitude of metrics that aim to quantify the 'significance', 'influence' or 'importance' of vertices or groups thereof. The variety of centrality metrics arise as a result of the ambiguity surrounding what is meant by importance. Whether importance is, for instance, the number of edges connected to a vertex (*degree centrality*) [1] or perhaps the average length of the shortest path between the vertex

and all other vertices in the network (*closeness centrality*) [1]. Each centrality metric provides a different perspective of a network with even simple versions being "very illuminating" [1]. We can also detect *similarity* between vertices in a network which too can be defined in many ways yet again relies on assessing particular quantities using information contained in a networks structure.

We can attempt to determine an appropriate random model for a network we would like to analyse. Decide, for instance, whether it is plausible that the networks topology could have been generated from an ER procedure. We might determine this by assessing whether the level of clustering or the skew of the degree distribution take on likely values assuming generation under an ER procedure.

1.4 Project Outline

Networks can have spatial structure that is not immediately apparent and real networks are often more highly clustered than random networks [1, 6, 7]. There are certain synthetic SERNs that can replicate small-world properties and thereby a high degree of clustering. However, this is not the only way in which clustering can be achieved. One other synthetic network capable of achieving clustering that leads to opportunities for further analysis is a class of power-law RDRN, we detail, that has been shown to identify protein associations in small-world proteome data sets [11]. Despite this, few studies have attempted to fit such SERNs to real networks or if alternative algorithms can be used to detect such structure apart from maximum-likelihood estimation [11] (MLE). In this interest, we review some of the literature concerning these SERNs and methods of generating them, in order to investigate the problem of reconstructing the spatial structure of these synthetic SERNs given only their topology. We generate SERNs that are amenable to analysis and investigate the problem of estimating the original spatial embedding of these SERNs, following the randomisation of their ordered adjacency matrices.

We investigate an alternative constrained optimisation technique to MLE — Simulated Annealing (SA) (adapted for our purposes) to calculate energies associated with orderings (states) of networks embedded in space in well defined ways. These states are entirely determined by the positioning of the vertices in a one-dimensional lattice; however, where possible, we aim to keep the framework of the method general to aid the adaption to different spatial-embeddings.

Additionally, we investigate how to quantify and qualify success for estimations of this nature, based on comparing true (original), disordered (randomised) and reordered (estimated) network embeddings. We investigate the application of the method we propose to empirical network data, in order to detect a hidden spatial structure and aim to determine the nature of any space we attribute to the network.

2 STATE ESTIMATION

Essentially, fitting a spatial network model where there is not an explicit spatial embedding has two problems:

1. Discerning whether a spatial network model is appropriate at all — detecting spatial structure.
2. Estimating the locations of the vertices in space (spatial embedding).

Here we detail an MLE approach to problem 2, under the assumption a spatial network model is appropriate (problem 1), based on work by Peter Grindrod concerning RDRNs[1] [11]. When estimating the spatial structure of empirical networks we cannot make the same assumption, and instead must compare results to ensembles of networks that serve as control groups that are deliberately synthesised, by ER procedures, to not exhibit spatial structures.

2.1 Maximum Likelihood Estimation

Here Grindrod demonstrates maximum likelihood estimation (MLE) to estimate the spatial embedding on the class of RDRN's we identify in Section 1.3.

It is important to note that this could be relaxed to more general systems to incorporate higher dimensions and of a non-uniform spatial embedding. Given a set of vertices ($V = \{1, 2, 3, ..., n\}$) we have $\boldsymbol{s}$ representing a permutation of those vertices, so that it is a bijective function of V to itself. The probability $p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}$ represents the probability (according our RDRN model) of vertices being connected when the distance between them is

$$d(\boldsymbol{s}(i), \boldsymbol{s}(j)) = |\boldsymbol{s}(i) - \boldsymbol{s}(j)|.$$

Essentially, $p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}$ is the probability that the vertices i and j are connected after the permutation $\boldsymbol{s}$ has been applied. Given an RDRN consisting of an arbitrary number of vertices the probability of observing no edges (an empty network) with duplicate and self–connections forbidden is

$$P(\varnothing) = \prod_{ij} [1 - p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}].$$

Consider now the probability of observing a new network, the same as this empty network except that contains a single edge introduced between some pair of vertices a and b

$$p(\text{new}) = \frac{p_{\boldsymbol{s}(a)\boldsymbol{s}(b)}}{1 - p_{\boldsymbol{s}(a)\boldsymbol{s}(b)}} \prod_{ij} [1 - p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}].$$

Notice that we factor out the probability of the edge not existing between a and b from the probability of the empty network and replace it with the probability of it existing. If we

[1] It should be noted that Grindrod refers to "range dependant random *graphs*" as opposed to *networks* and that this has deliberately been changed for consistency.

perform this process for all the edges that exist we arrive at the likelihood function employed by Grindrod

$$\mathcal{L} = \prod_{\text{edges } ij} \left[\frac{p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}}{1 - p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}} \right] P(\varnothing). \tag{5}$$

The likelihood to be maximised in (5) represents the probability of observing the pattern of connections under an RDRN model, given $\boldsymbol{s}$ were the true ordering. We can simplify (5) when we consider the power–law edge probability (2) relevant to the class of RDRNs we investigate.

Calibration

Assuming the RDRN fits our classification, the likelihood of any arrangement depends on the parameters α & λ, these therefore must be estimated prior to any manipulation of the initial random network embedding. Grindrod uses generating functions to relate the mean vertex degree z of an RDRN to the model parameters α and λ, each of which can in turn be related to C — the global clustering coefficient

$$C = \frac{3\alpha\lambda}{(1+\lambda)(1+3\lambda)} = \frac{3z(1-\lambda)\lambda}{2(1+\lambda)(1+3\lambda)}. \tag{6}$$

The global clustering coefficient C is defined as the ratio of closed triads to all triads in a network. Precisely, for all the connected triads of vertices $\{v_i, v_j, v_k\}$ that have two edges connecting v_i and v_j and v_k, C is the fraction of these where there exists an additional edge directly connecting v_i and v_k.

Both the variables C and z can be calculated purely from the topology of a network - and so can be calculated directly from the networks adjacency matrix. The calculation of C and z, in combination with (6), facilitate the estimation of the parameters α and λ necessary to calibrate the MLE process.

Algorithm

With α and λ chosen to match the global parameters C and z we have all the necessary information to calculate the likelihood of an RDRN, assuming it belongs to the class described above. In order to evaluate different orderings of a RDRN we need to define a method to generate new orderings from some random initial ordering.

Grindrod details an intricate "double level" algorithm to generate new orderings of a network embedding, and states that the algorithm combined with MLE achieves "an almost correct ordering (except for the overall sense of direction)". The algorithm updates each likelihood by "trading the odds of the reordered edges", where the odds, by definition, are

$$\text{odds}_{\boldsymbol{s}(i)\boldsymbol{s}(j)} = \frac{p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}}{1 - p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}}. \tag{7}$$

A potential issue with Grindrod's approach is that only re-orderings, by "swaps, block cycles and block flips", that increase the likelihood are accepted. The algorithm will initially

exhaustively test similar arrangements close to the initial "random" arrangement. In a finite search (less than that of possible arrangements) this can result in MLE convergence to local maxima, of which there are many. Yet, in an exhaustive search, there still lies a problem in the number of arrangements that need to be considered. Considering a one-dimensional network of $n = 100$ vertices there exist a possible 100! arrangements of the network embedding and therefore the potential of $100!/2$ ($\approx 4.7 \times 10^{157}$) unique likelihoods (considering that reverse arrangements would have equal likelihoods). In this interest we consider an adaptation to the optimisation technique — Simulated Annealing.

2.2 Simulated Annealing

The Simulated Annealing (SA) algorithm draws on an analogy to metallurgy, specifically, the heating and cooling of metal to mitigate internal stresses, known as annealing. When the metal cools to a "zero" temperature it's structure becomes fixed, and the energy state is at a minimum. The analogy substitutes the metal for the optimisation problem at hand, and the energy state for the problems objective function to be minimised. The key advantage of the SA algorithm is its tendency to converge to global optima [14], by mitigating restrictions when exploring a state space, detailed here. The SA algorithm comprises four main components.

Algorithm

I — Energy Function

The energy (or objective) function $E(\boldsymbol{s})$ is the function that SA aim's to minimise by means of finding the minimum energy state (permutation) $\boldsymbol{s}^*$ of the system. The function is analogous to the likelihood used by MLE. We can encode the existing edges of a network using an adjacency matrix $\boldsymbol{A}_{ij}$ - the square matrix representation of a finite network where

$$\boldsymbol{A}_{ij} = \begin{cases} 1 & \text{if there is an edge between vertices } i \text{ and } j, \\ 0 & \text{otherwise.} \end{cases}$$

We can restate the likelihood in Equation 5 as

$$\mathcal{L} = P(\varnothing) \prod_{\substack{\boldsymbol{A}_{ij}=1 \\ i<j}} \left[\frac{p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}}{1 - p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}} \right]. \tag{8}$$

It is important to note that $P(\varnothing)$ is constant regardless of state for a given number of vertices, so does not need to be calculated. We cannot directly substitute the likelihood from (8) as the problems energy function. The SA algorithm seeks solutions to minimise the energy function, therefore, we take the inverse of the likelihood (with $P(\varnothing)$ omitted). This corresponds to the product of the reciprocal of the odds of observing each edge for a given state. By taking

the logarithm of this product, a standard approach used in MLE, we are able to define the objective function as a summation

$$e = E(s) = -\sum_{i<j} \left[\boldsymbol{A}_{ij} \log \left(\frac{p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}}{1 - p_{\boldsymbol{s}(i)\boldsymbol{s}(j)}} \right) \right]. \tag{9}$$

Similar to before, when considering our specific class of power–law RDRNs the energy function becomes

$$\begin{aligned} e = E(\boldsymbol{s}) &= -\sum_{i<j} \left[\boldsymbol{A}_{ij} \log \left(\frac{p(|\boldsymbol{s}(i) - \boldsymbol{s}(j)|)}{1 - p(|\boldsymbol{s}(i) - \boldsymbol{s}(j)|)} \right) \right] \\ &= -\sum_{i<j} \left[\boldsymbol{A}_{ij} \log \left(\frac{\alpha \lambda^{|\boldsymbol{s}(i) - \boldsymbol{s}(j)| - 1}}{1 - \alpha \lambda^{|\boldsymbol{s}(i) - \boldsymbol{s}(j)| - 1}} \right) \right]. \end{aligned} \tag{10}$$

The SA algorithm relies on the estimation of the parameters α and λ, since these are involved in the specification of the energy function. Both parameters are calculated using (9) to match the global clustering coefficient C and mean vertex degree z. Grindrod cautions, however, that "some care must be taken" when estimating α and λ when the number of vertices in the network is small.

II — Annealing function

We utilise an annealing function to select new states of which to evaluate the energy function. We switch between network arrangements using the annealing function; different states correspond to different permutations of the network embedding. New permutations are reached by permuting current states conservatively, to enable the exploration of the state space in a well defined way.

We explore new states based on 'moves' generated by the annealing function. Following the random selection of two vertices of the network, these moves are selected at random from a possible three. Here we describe each of the three moves based on there application on an initial SERN of 12 vertices, with edges connecting vertices at neighbouring indices, shown in Figure 3.

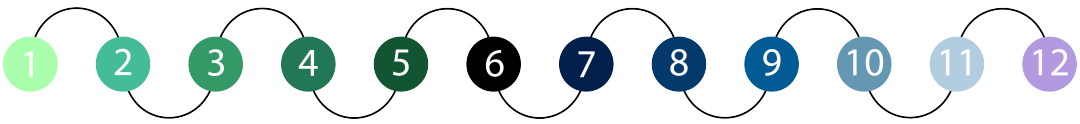


Figure 3: Example initial state of a SERN of 12 vertices.

The three possible moves are:

— ***Pairwise Swaps***: the two selected vertices switch index position in the network. This provides a simple way of permuting a networks state, in order to compare likelihoods on an individual vertex level.

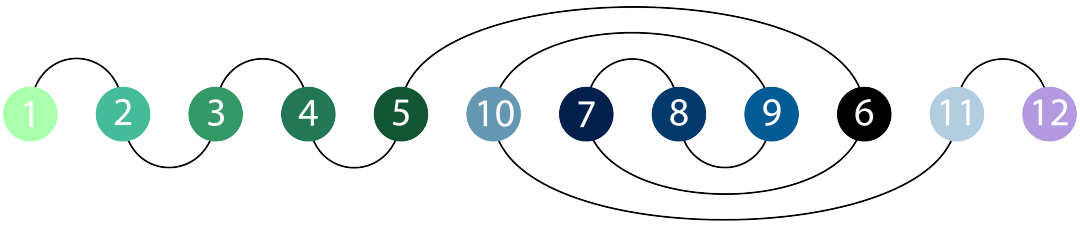


Figure 4: Diagram of an example pairwise swap conditioned on vertices six and ten.

— ***Block Flips***: the two selected vertices define the start and end indices of a block of vertices that's orientation is reversed. This enables larger portions of the network topography to be altered simultaneously, allowing small locally well optimised areas to be retained between moves.

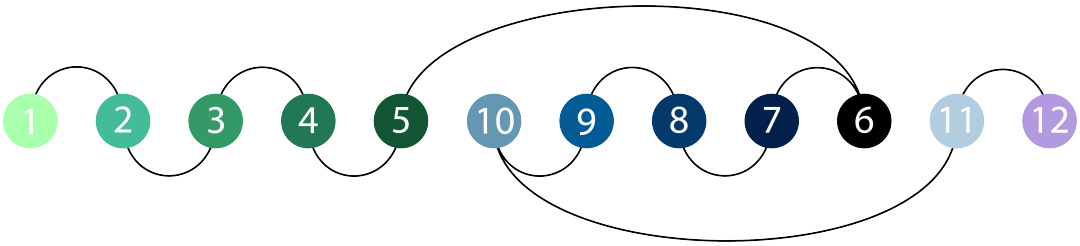


Figure 5: Diagram of an example block flip conditioned on vertices six and ten.

— ***Block Translations***: the two selected vertices define the start and end indices of a block of vertices that's start position is subsequently translated to a random index within the ordering. As with block flips, translations allow for the retention of locally well optimised sections between neighbouring states.

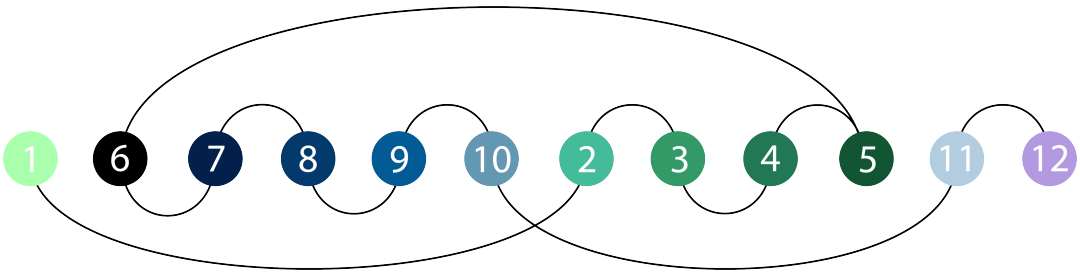


Figure 6: Diagram of an example block translation to the second index position, conditioned on vertices six and ten.

III — Acceptance probability

A state $\boldsymbol{s}$ will have an energy $e = E(\boldsymbol{s})$, associated with it defined by (9). If we permute a state $\boldsymbol{s}$ according to the annealing function we will arrive at a new state $\boldsymbol{s}'$, with corresponding energy e'. The SA algorithm needs to determine whether to either accept or reject each new state generated by the annealing function, in this interest we define an acceptance probability for new states.

The acceptance probability we define is always equal to one for permutations that lead to lower energy states, that is $p(e', e, T) = 1$ if $e' < e$. For new states that raise the system energy we consider a standard acceptance probability associated with SA based on the Boltzmann distribution [15, 16]

$$p(e', e, T) = \exp\left[\frac{-(e' - e)}{T}\right] \quad \text{if } e' > e. \tag{11}$$

Substituting the energy of a state given by (9) into (11) we arrive at

$$\begin{aligned}
p(e', e, T) &= \exp\left\{\frac{\sum_{i<j}\left[\boldsymbol{A}_{ij}\log\left(\frac{p_{s'(i)s'(j)}}{1-p_{s'(i)s'(j)}}\right)\right] - \sum_{i<j}\left[\boldsymbol{A}_{ij}\log\left(\frac{p_{s(i)s(j)}}{1-p_{s(i)s(j)}}\right)\right]}{T}\right\} \\
&= \exp\left\{\frac{\log\left(\prod_{\substack{\boldsymbol{A}_{ij}=1\\ i<j}}\left[\frac{p_{s'(i)s'(j)}}{1-p_{s'(i)s'(j)}}\right]\right) - \log\left(\prod_{\substack{\boldsymbol{A}_{ij}=1\\ i<j}}\left[\frac{p_{s(i)s(j)}}{1-p_{s(i)s(j)}}\right]\right)}{T}\right\} \\
&= \left\{\frac{\prod_{\substack{\boldsymbol{A}_{ij}=1\\ i<j}}\left[\frac{p_{s'(i)s'(j)}}{1-p_{s'(i)s'(j)}}\right]}{\prod_{\substack{\boldsymbol{A}_{ij}=1\\ i<j}}\left[\frac{p_{s(i)s(j)}}{1-p_{s(i)s(j)}}\right]}\right\}^{\left(\frac{1}{T}\right)} \\
&= \left\{\frac{P(\varnothing)\prod_{\substack{\boldsymbol{A}_{ij}=1\\ i<j}}\left[\frac{p_{s'(i)s'(j)}}{1-p_{s'(i)s'(j)}}\right]}{P(\varnothing)\prod_{\substack{\boldsymbol{A}_{ij}=1\\ i<j}}\left[\frac{p_{s(i)s(j)}}{1-p_{s(i)s(j)}}\right]}\right\}^{\left(\frac{1}{T}\right)}.
\end{aligned}$$

Comparing this to (8) we can see that, since $P(\varnothing)$ is independent of state, the base of the exponent is equal to the ratio of the likelihood of the new state to the current state. We are left with an acceptance probability that depends exponentially on the ratio between the two likelihoods. Combining the acceptance probabilities for higher and lower energies arrive at a complete acceptance probability

$$p(e, e', T) = \begin{cases} 1 & \text{if } e' \leq e \\ \left(\frac{\mathcal{L}_{\boldsymbol{s}'}}{\mathcal{L}_{\boldsymbol{s}}}\right)^{\left(\frac{1}{T}\right)} & \text{if } e' > e \end{cases} \tag{12}$$

As only the ratio of the likelihoods needs to be calculated we can simply consider the net change in inter-vertex ranges between neighbouring states.

IV — Temperature Schedule

The temperature schedule is the regime determining the temperature decline. The acceptance probability is a function of the temperature and therefore the cooling schedule needs to be carefully adjusted to suit the acceptance probability; this varies greatly between problems due to the complex dependence of the acceptance function on the annealing function.

The temperature must be decreased to progress the SA algorithm, as this corresponds to more stringent acceptance conditions at later stages of the SA algorithms annealing process. Figure

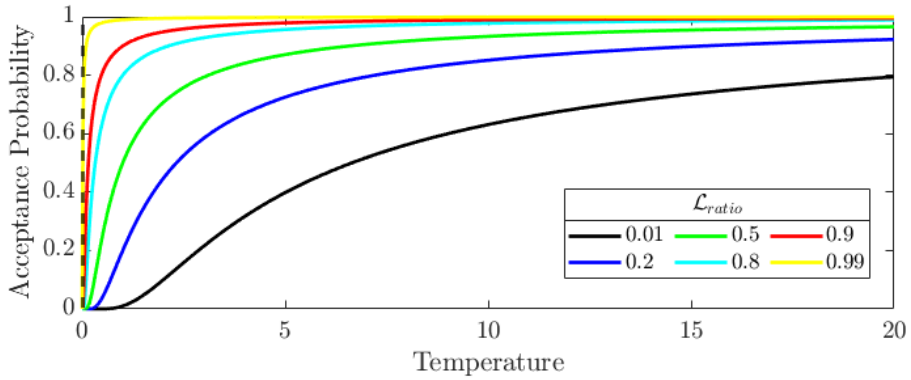


Figure 7: Acceptance probability against temperature for varying $\mathcal{L}_{\text{ratio}} < 1$.

7 shows the dependence of the acceptance probability on temperature for likelihood ratios that are not automatically accepted. We implement an exponential temperature schedule to iteratively decrease the temperature to progress the SA process

$$T(i) = T_{\text{Max}}\gamma^i \tag{13}$$

where $\gamma \in (0, 1)$ is a constant parameter determined by experiment depending on the SERN in question.

2.2.1 Simulated Annealing with Short-Term Memory

We implement an additional capability to the algorithm that explores a set of 'neighbour' states at each discrete temperature level. The neighbour states and the corresponding energies are retained and can be returned to at later stages of the annealing process, to provide a more complete exploration of the state space. The technique has been shown to be "more effective than conventional SA, both in terms of solution quality and time" [17] for travelling salesman (TS) problems. We can justify the appropriateness of this technique for our purposes since TS problems concern finding optimal orderings of vertices.

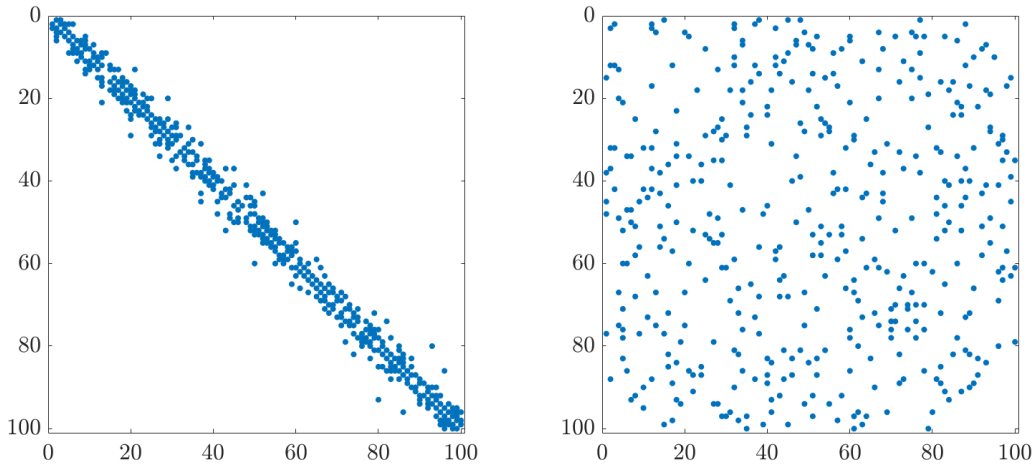
A summarising flowchart of the entire SA algorithm can be found in Figure 24 (see Appendix). The flowchart includes additional state variables $\{(\boldsymbol{s}_C, e_C), (\boldsymbol{s}_N, e_N), (\boldsymbol{s}_B, e_B)\}$ that denote current, neighbour and best values respectively.

3 EVALUATION PRELIMINARIES

3.1 Disordering

In order to assess the efficacy of the SA algorithm at reconstructing spatial structure, we randomly permute the $n \times n$ adjacency matrices of the networks generated according to section 1.3. This is achieved by generating a random permutation vector of length n and using this to map the true ordering $(1, ..., n)$ of the adjacency matrix to the corresponding indices of the permutation resulting in a randomly ordered adjacency matrix. This provides raw network topology data that is certainly correlated to space, but not in a way than can immediately be perceived by inspection of the adjacency matrix. We are left with networks that are amenable to spatial optimisation, and we can therefore apply reordering techniques to these networks to try to retrieve the true ordering.

Retaining the inverse of the random permutation used to shuffle the ordered adjacency matrices specifies the ideal reordering permutation of a perfect reordering. The evaluation of the SA algorithm's efficacy at determining the true spatial embedding of the model networks can therefore be evaluated by comparing the reordering permutation to this inverse permutation.



(a) True (b) Disordered

Figure 8: Adjacency matrices for a 100 vertex network generated by the RDRN model with $\alpha = 0.9$, $\lambda = 0.6$; originally and following a random permutation to a disordered spatial embedding.

3.2 Counting Inversions

The collection of all possible orderings (states) of the proposed SERNs compose their state space. In order to evaluate the algorithms effectiveness, it is necessary to compare the final state produced by the algorithm to the systems true state. To establish a metric for comparison we could consider the number of pairwise vertex swaps (transpositions) needed to permute the final state back to the original state. In the case of an exact reordering no transpositions are needed, and we might suppose that an increasing number of transpositions for a given number of vertices reflects a decreasing quality of reordering. However, identical reverse arrangements reflect an equivalent reordering quality but would in most cases require a different number of transpositions. We can counter this by considering both orientations of the final state. Lastly, the quality of reordering may be better captured by considering the number of transpositions of adjacent vertices required to permute the final state to the original. This metric will then depend on the distance of vertices from their true embedding down to an individual vertex level.

A permutation $\boldsymbol{\sigma}$ can be specified by the sequence $(\sigma_1, ..., \sigma_n)$ of its images (the outputs of the permutation $\sigma(i)$ for each index in the sequence of a natural order $(1, ..., n)$). A permutation has an inversion where two of its elements are out of their natural order. The set of inversions of a permutation correspond to all the pairs of indices (i, j) with $i < j$ that have corresponding images satisfying $\sigma_i > \sigma_j$. We are interested in calculating the minimum number of adjacent transpositions of vertices required to permute the final reordered state back to the original, or more succinctly, the total inversions required to transform the final permutation to the original permutation [18]. Note that the identity permutation here represents the networks original embedding.

Lehmer code

To calculate the inversions between states we utilise the Lehmer code [18], a method of encoding the possible permutations of a sequence of n numbers. For a permutation $\boldsymbol{\sigma}$, the Lehmer code is the sequence

$$L(\boldsymbol{\sigma}) = (L(\sigma_1), ..., L(\sigma_n)), \quad \text{where } L(\sigma_i) = \#\{j > i : \sigma_j < \sigma_i\}. \tag{14}$$

Essentially, the sequence is constructed by counting the number of subsequent images of the permutation sequence that are less than the image at each index, and is therefore identical to the right-inversion count [19]. The Lehmer code enables the efficient calculation of the total number of inversions I of a permutation, as this is equal to the sum of the elements of the Lehmer code for the permutation

$$I(\boldsymbol{\sigma}) = \sum_{i=1}^{n} L(\sigma_i). \tag{15}$$

The average number of inversions I_μ for a permutation of length n is given by (16) [20] which provides an upper benchmark for comparison of true and reordered states

$$I_\mu(n) = \frac{n(n-1)}{4}. \tag{16}$$

3.3 Connectedness and Orientation

Strictly, a network is connected if and only if it is comprised of a single component. Conversely, this means that any network that is not connected must be comprised of multiple components, each of which amenable to local spatial optimisation.

If a one-dimensional network is disconnected than the energy function will certainly have more than two global optima (at least one globally optimal state and it's equivalent reverse state are guaranteed). Since the components of a spatial network can be locally optimised individually this leads to issues when assessing the quality of the global reordering for disconnected networks.

As an example, Figure 10 shows a comparison between two possible states of a five component network. The two states of the networks here have equal energy evaluations, yet clearly inversions exist between them. The total inversions of a permutation is therefore not an adequate assessment of reordering quality for disconnected networks, and optimisation should instead be focused locally on individual components. Considering this, we empirically identify the areas of the (α, λ) parameter space that give rise to connected RDRNs.

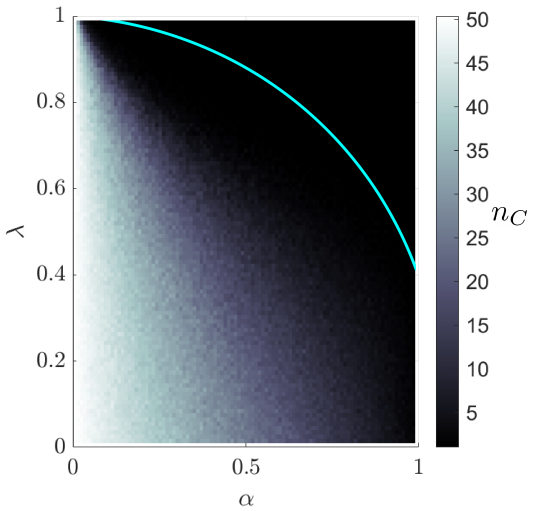


Figure 9: Mean number of network components n_C for a 50 vertex network with approximate 95% confidence boundary for connected networks based off 100 parameter sweeps.

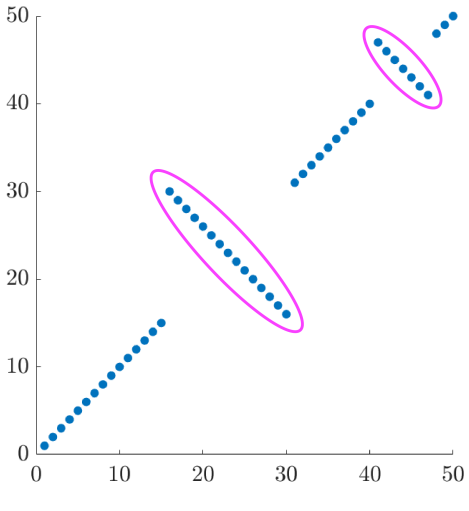


Figure 10: Spatial embedding comparison between two states of a disconnected network of five components. Here two components (circled) have reversed spatial orientations between the two networks.

4 RESULTS

Here, we quantitatively examine the influence of model parameters on the effectiveness of SA on reconstructing spatial structure on each class of SERN defined in Section 1.3, assuming the networks are generated by a RDRN model. We compare the central tendency, measured by the arithmetic mean, of the following quantities we refer to collectively as evaluation metrics:

- The system energy of the true state $e_T = E(\boldsymbol{s}_T)$ according to (9).
- The system energy of the minimum energy state estimated through simulated annealing $e_{SA} = E(\boldsymbol{s}_{SA})$ according to (9).
- The system energy of the disordered state $e_D = E(\boldsymbol{s}_D)$ according to (9).
- The total inversions of the permutation mapping the true state to the minimum energy state estimated through simulated annealing $I_{TSA} = I(\boldsymbol{\sigma}_{TSA})$, according to (15). That is, the total inversions for the permutation satisfying $\boldsymbol{\sigma}_{TSA}(\boldsymbol{s}_T) = \boldsymbol{s}_{SA}$.

For each network class we consider network realisations of 30 vertices for parameter sweeps. We omit results for points that do not give rise to connected networks, for the reasons described in Section 3.3.

4.1 Range-Dependent Random Networks

Here we evaluate the performance of SA on the reordering of one-dimensional RDRNs that obey the power-law in (2). We evaluate the mean of the evaluation metrics at points uniformly spread over the (α, λ) parameter space.

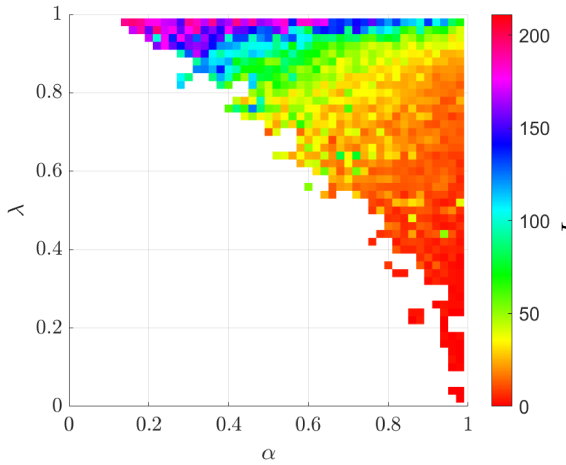


Figure 11: Mean of I_{TSA} for linearly spaced points selected over the entire (α, λ) parameter space for connected RDRNs with 30 vertices.

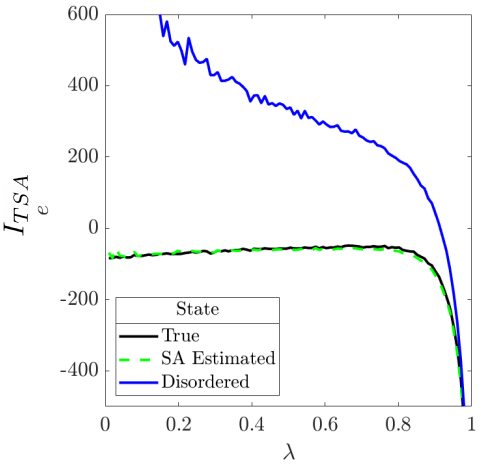
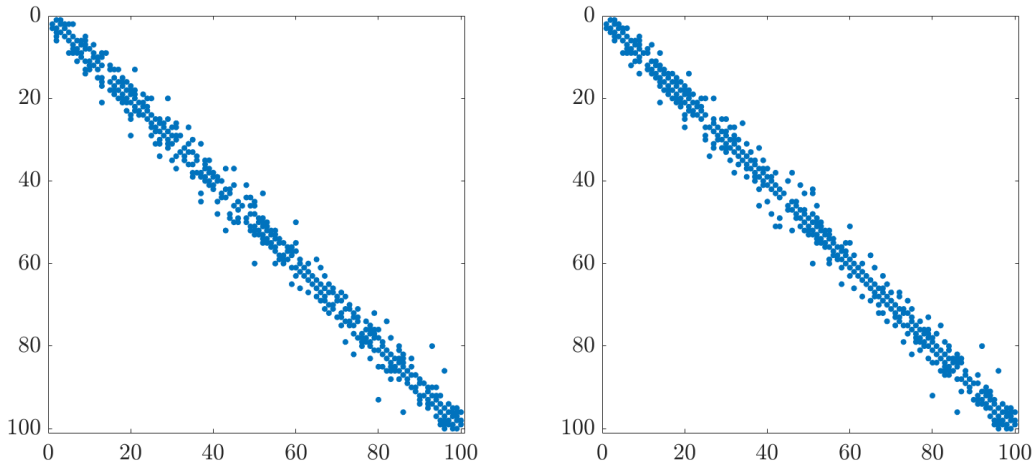


Figure 12: Mean of e_T, e_D and e_{SA} for fixed $\alpha = 0.95$ and varying $\lambda \in (0, 1)$ for connected RDRNs with 30 vertices.

Figure 11 shows the mean of I_{TSA} for connected networks across the parameter space. Generally I_{TSA} decreases with decreasing λ and increasing α for these RDRNs[2], with I_{TSA} reaching zero for certain parameter values that approach $(\alpha = 1, \lambda = 0)$ — indicating perfect reorderings ($s_T = s_{SA}$).

Figure 16 shows the mean of e_T, e_D and e_{SA} over a slice of the parameter space ($\alpha = 0.95$ and $\lambda \in (0, 1)$). Each of the system energies begin to decline sharply with increasing λ at $\lambda \approx 0.85$. e_D essentially forms a surface of values above e_T and e_{SA} over the entire parameter space, except in the parameter limit $\lambda \to 1$ (where each energy surface coincide). e_T and e_{SA} take similar mean values over the parameter space however the two energy surfaces intersect transversely at various points. This intersection is visible at $\lambda \approx 0.3$ in Figure 16.

To help contextualise the above results, we detail an example reordering of an 100 vertex RDRN. Figure 13 shows the true ordered adjacency matrix of the RDRN compared with the ordered adjacency matrix estimated through SA.



(a) $\boldsymbol{A}_{ij}^{(\boldsymbol{s}_T)}$ (b) $\boldsymbol{A}_{ij}^{(\boldsymbol{s}_{SA})}$

Figure 13: Visualisation of non-zero entries of the ordered adjacency matrices for a 100 vertex network generated by the RDRN model with $\alpha = 0.9$, $\lambda = 0.6$. For states $\boldsymbol{s}_T$ and $\boldsymbol{s}_{SA}$.

Figure 13 shows that SA estimates an ordered adjacency matrix that resembles the original for this 100 vertex network. The reordering in this instance results in a total of 48 inversions to the original state in contrast with 2475 inversions — the average number of inversions for a permutation of length 100 calculated from (16). The quality of reordering can also be visualised by plotting the original against the reordered state of the network, shown in Figure 14. Here, points that lie on the diagonal represent exact matches, enabling the verification of similarity in the ordered adjacency matrices.

[2]Similar results were obtained for samples of higher numbers of vertices (tested up to 150 vertices), although, due to the factorial growth of the state space with increasing n, the minimum energy configurations take a much greater number of iterations to reach for larger networks. This led us to consider smaller networks for parameter sweeps, due to computational constraints.

For the example network, this verification test shows that SA correctly estimates the exact spatial embedding of approximately half of the vertices of the network (49 out of 100), the remaining vertices are estimated to lie at most four indices away from their true embedding.

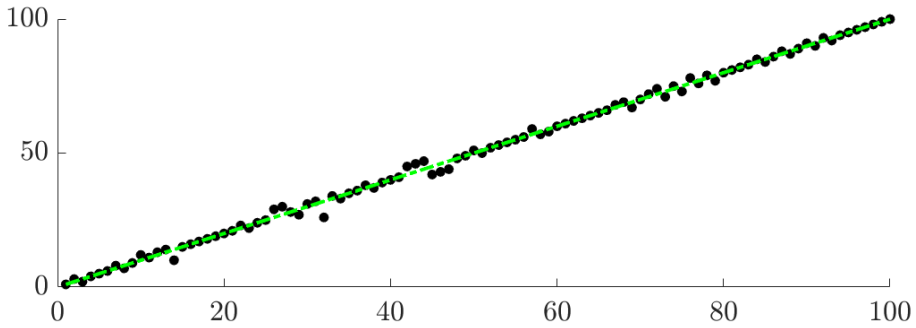


Figure 14: Original against achieved states for a 100 vertex network generated by the RDRN mode with $\alpha = 0.9$, $\lambda = 0.6$. Both originally and following a random permutation and reordering by the SA algorithm.

4.2 Watts–Strogatz Networks

Here we evaluate the performance of SA on the reordering of one-dimensional WS networks generated according to Section 1.3. Note the algorithm is derived inline with assumptions that the network it is applied to assumes a power-law (2) for the existence of edges across possible inter-vertex ranges. We evaluate the mean of the evaluation metrics at linearly spaced values of the parameter $p \in (0, 1)$ for four different values of the parameter $z/2 = 2, 3, 4$, corresponding to the mean degrees $z = 4, 6, 8$. Values for $z/2$ lower than those specified often result in disconnected networks, whereas, higher values can generate dense netowrks.

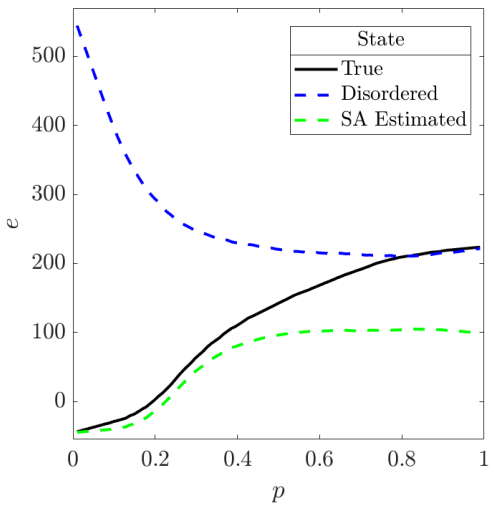


Figure 15: Mean of e_T, e_D and e_{SA} for fixed $z/2 = 4$ and varying $p \in (0, 1)$ for connected WS networks with 30 vertices.

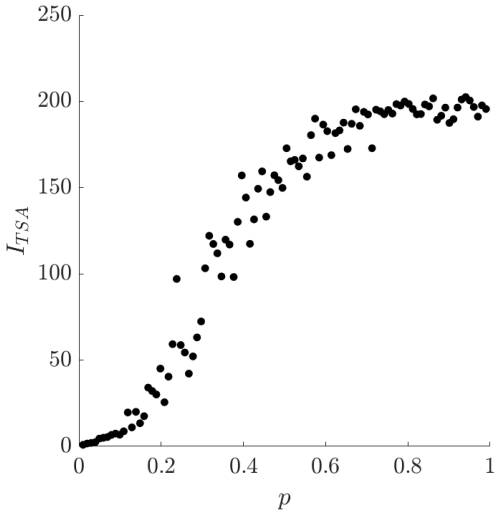


Figure 16: Scatter diagram of I_{TSA} for fixed $z/2$ and varying $p \in (0, 1)$ for connected WS networks with 30 vertices.

Figure 15 shows the mean of e_T, e_D and e_{SA} for $z/2 = 4$ and $p \in (0, 1)$. Here, e_D declines with increasing p and is greater than e_{SA} over all p. e_T increases for increasing p, converging to match values of e_D for larger p (both equalling approximately 200 for $p \in [0.75, 0.1)$).

e_{SA} also increases with increasing p and coincides with e_T for $p \in (0, \approx 0.3)$ before diverging, potentially indicating high quality reordering in this range. The rate of change of e_{SA} with respect to p approaches zero for increasing p outside this range, settling at $e_{SA} \approx 100$ for $p > 0.4$.

We find, that decreasing $z/2$ within the prescribed values simply decreases the system energies evaluated across possible p, and results in minimal other discernible changes to the topography of the system energies across p (see appendix).

4.3 Empirical Network — Zachary's Karate Club

Zachary's karate club (ZKC) is a network introduced and compiled by Wayne W. Zachary [21] based on Zachary's observations, over two years, of the interactions of 34 members of a university karate club. The network is a popular social network of study [1], and has been shown to exhibit community structure [22]. Communities within networks can be characterised as sets of densely interconnected vertices [1]. Barthelémy argues that community structures can often be determined by "geographical factors" [6] using the case of air transportation networks [23]. Since geographical factors are inherently spatial, the presence of community structure in ZKC alludes to the possibility of spatial structure in the network, and therefore makes the network an interesting candidate to subject to spatial optimisation.

Our interest here lies in whether the proposed SA algorithm can uncover a 'hidden' spatial structure within ZKC. An important trait of the network that potentially could have significant implications on any results is the network's degree distribution. ZKC has a distinctly non–homogeneous degree distribution shown in Figure 17, in contrast to the model networks we investigate that have relatively homogeneous degree distributions.

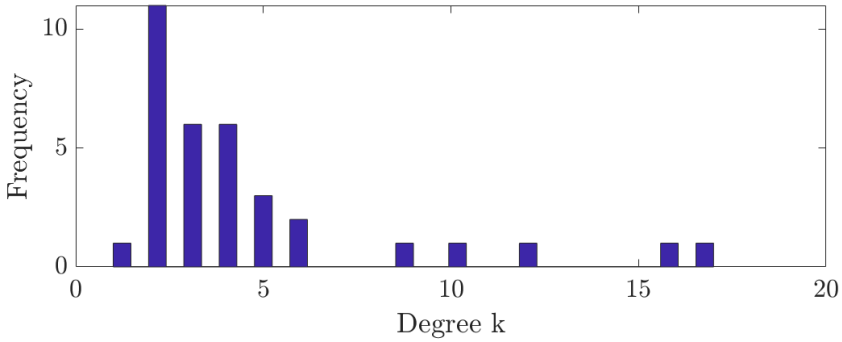


Figure 17: Histogram showing ZKC's degree distribution.

In order to discern whether SA can identify any spatial structure within ZKC it is necessary to specify a set of networks that can serve as a control group. The control group we specify is a set of random networks using the Chung–Lu model [24] that account for the non-homogeneity in ZKC's degree distribution. The resulting ensemble of networks have the same expected degree distribution as ZKC but, crucially, do not share the same topology. These networks serve as a control group by ensuring that any differences in the energy of the final ZKC compared to a random network of the same number of vertices are not due to differences in the degree distribution.

We therefore compare the central tendency of the system energy (assuming the networks were generated under a RDRN model) of the final state estimated through simulated annealing e_{SA} for ZKC with e_{SA} for a large number of random networks generated by the Chung–Lu model. The Chung–Lu model constructs networks according to a probability p_{ij} dictated by a networks expected degree distribution where

$$p_{ij} = \frac{k_i k_j}{\sum_{l=1}^{n} k_l}.$$

For each of the pairs of vertices i, j the Chung-Lu model constructs a networks adjacency matrix according to:

$$\forall_{i>j} \quad \boldsymbol{A}_{ij} = \boldsymbol{A}_{ji} = \begin{cases} 1 & \text{with probability } p_{ij}, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 18 shows a histogram comparing the results for e_{SA} for both ZKC and the control group, Figure 19 shows a visualisation of the ordered adjacency matrix $\boldsymbol{A}_{ij}^{(\boldsymbol{s}_{SA})}$ for ZKC.

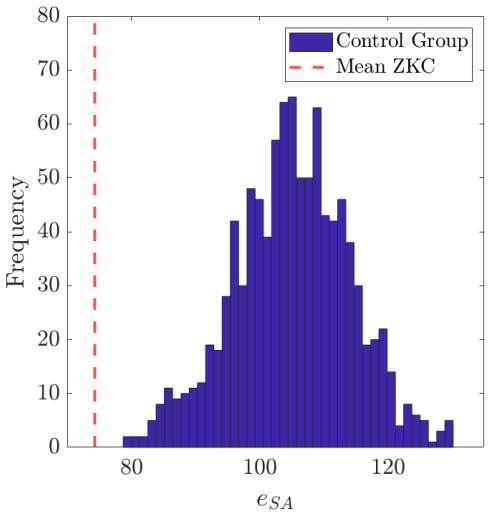


Figure 18: Histogram of the minimum energy estimate through simulated annealing e_{SA} for 1000 control networks labelled with the mean e_{SA} for ZKC.

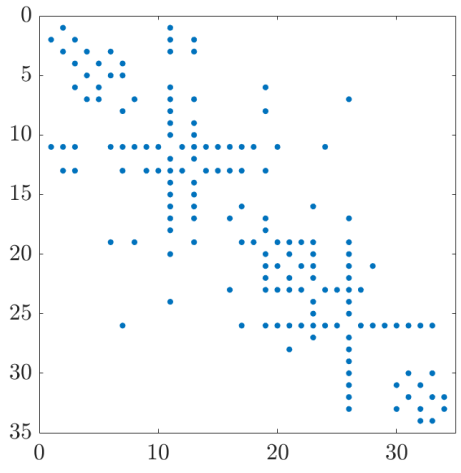


Figure 19: Ordered adjacency matrix visualisation of the minimum energy state estimate through simulated annealing s_{SA} of ZKC.

The histogram shows that the mean e_{SA} for ZKC is lower than that for any of the 1000 control networks, suggesting that ZKC is in fact spatial. The energies of the control group appear to be normally distributed, fitting a normal distribution to the control group returns a mean of 105.114 and standard deviation 9.08854.

To test the null hypothesis $\boldsymbol{H_0}$ that the mean e_{SA} of ZKC may have come from such a distribution ($\mathcal{N}(105.114, 9.08854^2)$) we carry out a z–test. Performing a z-test returns a p–value of 0.000693, indicating that we can reject the null hypothesis at the 0.1% significance level.

The ordered adjacency matrix of e_{SA} for ZKC appears to have elements that occur closer to the leading diagonal than for a random network. We would expect that if there were no spatial structure that the ordered adjacency would resemble that of a disordered arrangement of a SERN (see Figure 8b).

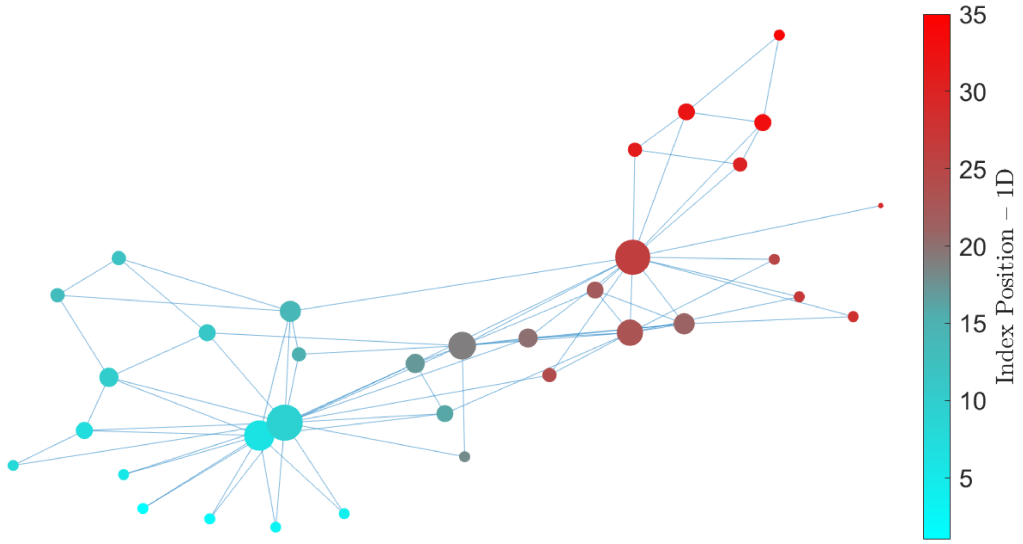


Figure 20: The network topology of ZKC with vertices colour–coded to the corresponding one–dimensional index position of $\boldsymbol{s}_{SA}$. The size of each vertex is representative of it's degree (the larger the size the higher the degree).

Figure 20 shows the network topology of ZKC with each of the vertices colour-coded to match their corresponding index position in the one-dimensional state estimated through SA — $\boldsymbol{s}_{SA}$. Here the two vertices of the largest degree (reflected by relative size) are of a distinctly different colour. The immediate neighbours of both of these vertices match the colour of the corresponding larger vertex they are connected to, with the exception of certain vertices that are adjacent to both. Figure 20 also provides a visualisation of centrality, vertices that are less discernibly coloured toward either side of the spectrum are estimated to be more central (spatially) through SA.

5 DISCUSSION

State Estimation of Model Networks

Generally the results show that, for each SERN we investigate, we are able to infer spatial structure over some portion of the respective parameter space.

For RDRNs, parameters sweeps evaluating I_{TSA} (Figure 11) indicate that states of RDRNs generated with model parameter values that approach $(\alpha = 1, \lambda = 0)$ are the most reliably well–estimated, as I_{TSA} tends to the lowest value. This confirms intuition since networks generated in the parameter limit $(\alpha \to 1, \lambda \to 0)$ produce 'string–like' networks. These are networks in which each vertex is connected to the vertices in the adjacent index positions (where they exist) only; $\lambda \to 0$ precludes edges over ranges exceeding one. String–like networks can therefore only have a single minimum energy configuration (ignoring global orientation) which concurs with the reliable estimation of network states at this limit.

In the limit $\lambda \to 1$, as described in Section 1.3, the RDRNs approach Erdős-Rényi networks — a degenerate case of RDRNs. Edges in an ER network ensembles exist with a probability independent to range and therefore the true state of such networks is impossible to estimate through SA. This is again concurrent with results, with the mean of I_{TSA} for a RDRN of 30 vertices with $\lambda \to 1$ equalling the average inversions of a permutation of length 30 $I_\mu(30) = 217.5$ calculated by (16).

What is of most interest here is the efficacy of SA at estimating minimum energy states for the non-degenerate class of RDRNs. That is, those generated with intermediate parameter values, as these correspond to network ensembles with spatial structure. Figure 28 (see appendix) shows the energies associated with a slice of the intermediate parameter space. Comparing this with Figure 29 shows that λ values from 0 to 0.9 result in an approximately linear increase in I_{TSA} (from 20 to 35) — matching the steady decrease of the system energy e_D and convergence of all system energies shown in Figure 28. For RDRNs the increase in I_{TSA} is certainly correlated to the convergence of e_T and e_D, although also depends on the similarity between e_T and e_{SA}. The comparison of Figure 28 to Figure 29 allows us to infer that parameter values giving rise to small differences between e_T and e_{SA} and large differences between e_D and both e_T and e_{SA} constitute the parameter values that SA is able to infer spatial structure best, where this does not hold we observe convergence of I_{TSA} to I_μ.

Parameter sweeps concerning WS networks of $p \in (0, \approx 0.3]$ for $z/2 = 4$, shown in Figure 15, reveal that e_{SA} and e_T are similar and are both significantly lower than e_D. There clearly exist distinctly lower energy states than random states ($\boldsymbol{s}_D$) for these parameter values, and similar energies can be achieved through SA. This indicates that it may be possible to infer spatial structure for WS networks generated under these parameter values as the differences in the system energies are analogous to the differences in system energies of successfully reordered RDRNs. As with RDRNs, we are able to verify success by investigating I_{TSA} over the range $p \in (0, 0.5]$, highlighted by the similarity of e_{SA} and e_D in this range.

The results for the WS networks in Figure 15 show that e_D and e_T converge at $p \approx 0.75$. Since these two energy evaluations are the same for $p > 0.75$, then neither the true or disordered states are more likely, according to SA, under the RDRN model. The system energy e_{SA} remains lower than both e_D and e_T over this range. The edges of WS networks assume a range dependence due to the initial formation of a lattice of vertices connected to a certain number of nearest neighbours on either side. The parameter p essentially controls the randomness of the formation of edges. Increasing p increases the randomness of the network by perturbing the topology of the network ensembles away from spatially structured lattices — the edges of which clearly exhibit range dependence (see appendices, Figure 25). This helps to explain why both e_D and e_T converge and why e_{SA} remains lower than both e_D and e_T. The former occurs since the generated networks approach random networks and therefore s_D is the result of a random permutation of a state s_T that is determined randomly to begin with. It stands to reason therefore that the energies associated with s_T and s_D — e_T and e_D are similar for high p. The energy e_{SA} remains lower than both e_D and e_T since there is a high probability that there exist states with lower energies than a given random state, regardless of p.

In summary, the results associated with the model SERNs verify that the proposed SA algorithm can be used to reconstruct the underlying spatial structure of certain spatial networks. The verification of our method opens the potential to explore other complex systems using SA. We can estimate the spatial-embedding of synthetic power-law RDRNs through SA although the capability to infer spatial-structure does not seem to be limited to power-law RDRNs. The method can also reconstruct the correct ordering of WS networks.

State Estimation of an Empirical Network

The results of the application of SA to the empirical network ZKC, suggest that ZKC is in fact spatial. We show, through a z–test, that the probability of a minimum energy configuration arising through SA as low as that of ZKC from a random network ensemble with the same expected degree distribution is less than 0.001. We must, however, consider whether there is anything other than the possibility that ZKC is spatial that may contribute to the fact that ZKC has a distinctly lower minimum energy state than it's simulated, random equivalents (the control group).

It is important to note that the control group we compare the results of ZKC with is a random network ensemble with the same *expected* degree distribution as ZKC as opposed to the same *exact* degree distribution. The configuration model [25, 26] is a method of generating ensembles of networks that have exactly a prescribed degree distribution. The configuration model generates every possible network with the specified degree distribution with equal probability [26], which naturally includes networks with duplicate and self–connections. To generate realisations of the configuration model that do not have duplicate or self–connections the defining algorithm of configuration model (that constructs edges one-by-one) must be restarted altogether as soon as a violating edge is generated, which has implications. Restarting the

algorithm ensures that the resulting realisations of networks are sampled uniformly across possible acceptable networks. Imposing the restriction to this subset of acceptable networks, however, results in a low probability of generating any acceptable networks when using the configuration model prescribed with ZKC's degree distribution. Synthesising networks under the configuration model in this case does not produce a single viable network in over 1 million trials, a problem addressed by utilising the Chung–Lu model (that guarantees networks without duplicate and self–links) with the expected degree distribution.

We are unable to successfully synthesise a null model of a network ensemble generated under the configuration model. Despite this, it remains that we can infer that ZKC seems to possess a spatial structure, based on the z-test. The key aspect to investigate is what the space we have attributed to the network might represent. We could speculate that the space reflects, perhaps, a preferred style of karate or preference of teaching instructor within the karate club. Without any additional information regarding the club besides its topology, determining the true nature of the underlying space is clearly conjectural. Importantly, however, Zachary observed a disagreement between the clubs instructor and administrator. The disagreement resulted in a split in the karate club network [1, 21], a spilt acknowledged and investigated in numerous studies concerning ZKC as well as community structure [22, 27]. The split of the 34-vertex network resulted in the formation of two sub-networks or factions, surrounding either the clubs administrator or instructor. Detail of the split was recorded by Zachary down to the faction affiliation of each individual member of the original network. Figure 21 shows the observed factions involved in the split of ZKC.

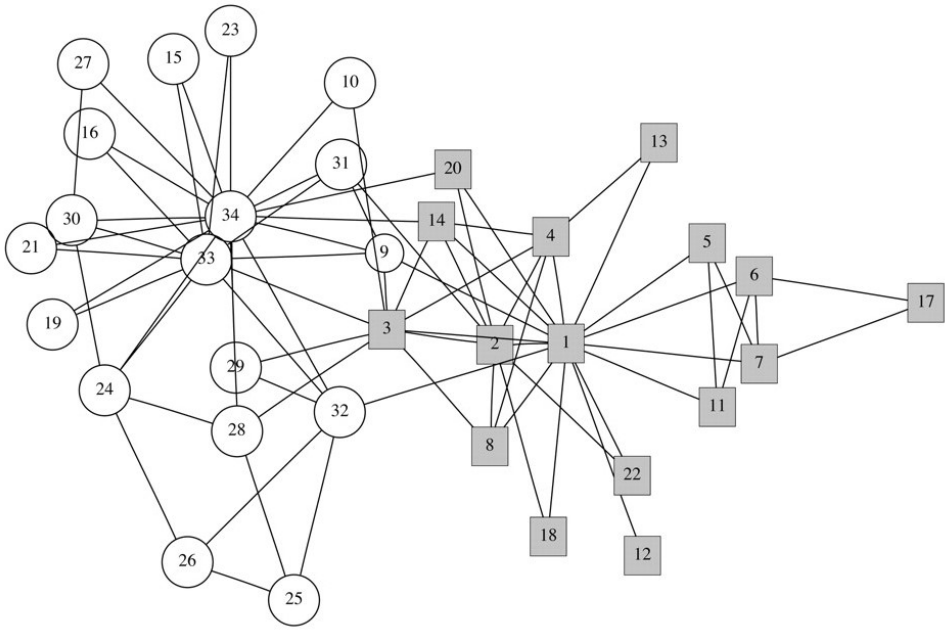


Figure 21: An unweighted version of the social network as described in Zachary's original paper [21]. Circles represent those associated with the club administrator's faction. Squares represent those associated with the instructor's faction. Taken from [22].

A notable community-detection technique able to detect these factions is the *Girvan–Newman* algorithm proposed by Michelle Girvan and Mark Newman [22]. The algorithm works by progressively removing edges from the original network based on what Girvan and Newman define as *edge betweenness* [22]. The edge betweenness is a generalisation of Freeman's *betweenness centrality* [28] and is defined for an edge as "the number of shortest paths between pairs of vertices that run along it".

We can compare the actual observations of the division of the network in Figure 21 to the minimum energy state $\boldsymbol{s}_{SA}$ estimated for ZKC through SA. The SA algorithm returns a one-dimensional minimum energy state $\boldsymbol{s}_{SA}$ and we would expect that vertices lying at opposing ends of the state to be the most distant in whatever hidden space we are attempting to uncover. Figure 22 shows $\boldsymbol{s}_{SA}$ with the vertices depicted according to the factions observed by Zachary and suggests that the SA algorithm we present produces an ordering that corresponds to the observed factions.

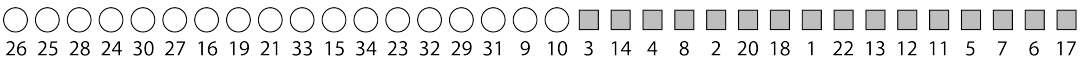


Figure 22: The minimum energy state $\boldsymbol{s}_{SA}$ of ZKC obtained through SA, with vertices depicted according to their corresponding faction affiliation. Circles represent those associated with the club administrator's faction. Squares represent those associated with the instructor's faction.

The minimum energy state $\boldsymbol{s}_{SA}$ of ZKC does not exhibit any overlap between the two factions, suggesting the space we have attributed to the network may in fact correlate to the sociological aspects of the club responsible for the formation of the factions themselves. We can compare $\boldsymbol{s}_{SA}$ of ZKC to the community-structure identified by the Girvan-Newman algorithm, shown by a hierarchical tree in Figure 23.

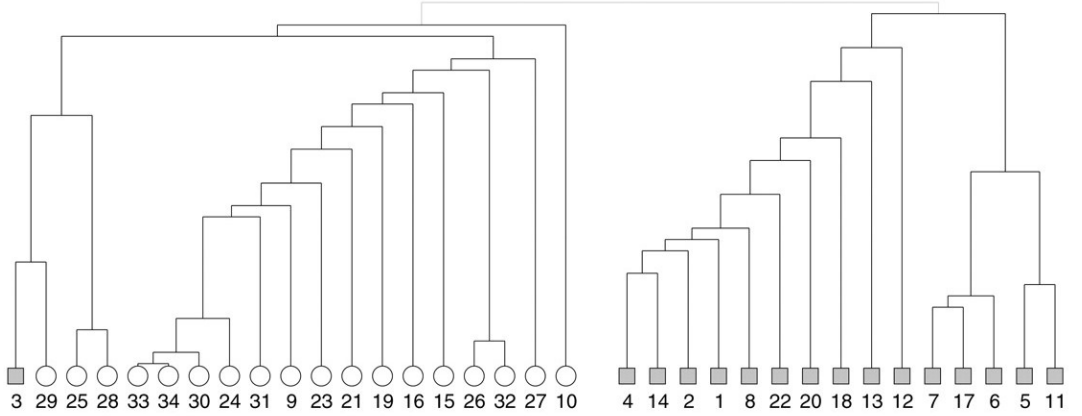


Figure 23: Community structure for ZKC calculated using the Girvan–Newman algorithm and presented as a hierarchical tree. Circles represent those associated with the club administrator's faction. Squares represent those associated with the instructor's faction. The top-level split of the network into two factions is in agreement with the actual factions observed by Zachary, with the exception of vertex 3. Taken from [22].

The results of the Girvan–Newman algorithm shows that only a single vertex, vertex 3, is misclassified. It is interesting to note that the misclassified vertex corresponds to the vertex within minimum energy state $\boldsymbol{s}_{SA}$ for ZKC that is positioned at the border between two factions. SA does not explicitly highlight communities within a network; however, the minimum energy state clearly corresponds to the factions described by Zachary. We might expect that vertices within $\boldsymbol{s}_{SA}$ lying closer to the border to be the most weakly affiliated with their faction as these are the vertices that are estimated to lie nearest, spatially, to the alternative faction. Assuming that bordering vertices are the least strongly affiliated with the faction is in agreement with the results obtained by the Girvan and Newman in Figure 23, as vertex 3, a bordering vertex, is the only vertex that is misclassified.

6 CONCLUSION & FURTHER WORK

We have investigated the problem of discerning spatial structure within a selection of SERNs and introduced a method to detect such structure using the information contained in each network's topology exclusively. The SA method we propose is designed to identify and reconstruct the spatial-embedding of RDRNs that obey a power-law edge probability. We verify our methods by demonstrating the reconstruction of the underlying spatial structure of synthetic power-law RDRNs, with a high degree of success over a significant portion of the parameter space. The method is effective at reconstructing the correct order of synthetic power-law RDRNs but does not seem to be limited to them. It can also reconstruct the correct order of synthetic WS networks, over a portion of the associated parameter space.

We detail the application of the method to a well documented empirical network, known to exhibit community structure. We find that the method produces a minimum energy configuration of the network lower than that of an associated control group with the same expected degree distribution, indicating a spatial structure. Comparing the results to ethnographic data associated with the empirical network provides insight into the space that simulated annealing attributes to the network. The comparison suggests that the space corresponds to factions within the network, factions detected by existing community-structure algorithms. The results coincide with that of the community-structure algorithm and potentially provides insight into why community-detection can fail. Spatial structure may therefore be meaningful in networks in the same way that community structures can be meaningful. The attributed spatial structures may contribute to, or have the advantage over, community detection in that they exhibit "shades of grey" rather than black and white exclusively. These grey areas essentially reflect a metric for vertex centrality.

The verification of our method opens the potential to explore other complex systems using simulated annealing and we expect that it may be possible to identify spaces associated with other mesoscale properties of networks using the method proposed here, such as core-periphery structures. In this interest, future work may involve testing the method on empirical networks known to exhibit other mesoscale properties.

The proposed method is shown to be effective for uncovering spatial structure in specific networks. More work is required to discern how the method we present generalises to reconstructing the spatial structure of other model networks — that is, networks generated by different rules and edge probabilities (that reflect other spatial and real systems). The method certainly needs adaptation to be optimised for other classes of synthetic networks to account for these differences. Although, it should be noted that the method may still correctly infer some level of spatial structure in it's current form for other synthetic networks, as it can for a subset of Watts–Strogats networks. Additionally, we anticipate the need to expand the one-dimensional implementation of the method to higher dimensions in order to capture more nuanced spatial structures than those identified in this article.

References

[1] M. E. J. Newman, "Networks: an introduction," 2010.

[2] J. C. Lang, H. De Sterck, J. L. Kaiser, and J. C. Miller, "Analytic models for SIR disease spread on random spatial networks," *Journal of Complex Networks*, vol. 6, no. 6, pp. 948–970, 2018.

[3] P. Erdős and A. Rényi, "On random graphs 1," *Publ. Math. Debrecen*, vol. 6, pp. 290–297, 1959.

[4] M. E. J. Newman, "Random graphs as models of networks," 2002.

[5] J. Leskovec, J. Kleinberg, and C. Faloutsos, "Graph evolution: Densification and shrinking diameters," 2006.

[6] M. Barthélemy, "Spatial networks," *Physics Reports*, vol. 499, no. 1, pp. 1 – 101, 2011.

[7] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world'networks," *nature*, vol. 393, no. 6684, p. 440, 1998.

[8] C. Moore and M. E. J. Newman, "Epidemics and percolation in small-world networks," *Phys. Rev. E*, vol. 61, pp. 5678–5682, 2000.

[9] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, pp. 3200–3203, Apr 2001.

[10] Y. Wang, G. Xiao, J. Hu, T. H. Cheng, and L. Wang, "Imperfect targeted immunization in scale-free networks," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 12, pp. 2535–2546, 2009.

[11] P. Grindrod, "Range-dependent random graphs and their application to modeling large small-world proteome datasets," *Phys. Rev. E*, vol. 66, p. 066702, 2002.

[12] J. Jia and A. R. Benson, "Random spatial network models for core-periphery structure," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining*, pp. 366–374, ACM, 2019.

[13] M. E. Newman, "Assortative mixing in networks," *Physical review letters*, vol. 89, no. 20, p. 208701, 2002.

[14] D. Bertsimas, J. Tsitsiklis, *et al.*, "Simulated annealing," *Statistical science*, vol. 8, no. 1, pp. 10–15, 1993.

[15] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *science*, vol. 220, no. 4598, pp. 671–680, 1983.

[16] E. Aarts and J. Korst, "Simulated annealing and boltzmann machines," 1988.

[17] H. Bayram and R. Şahin, "A new simulated annealing approach for the travelling salesman problem," *Mathematical and Computational Applications*, vol. 18, pp. 313–322, 12 2013.

[18] D. H. Lehmer, "Teaching combinatorial tricks to a computer," in *Proc. Sympos. Appl. Math. Combinatorial Analysis*, vol. 10, pp. 179–193, 1960.

[19] D. E. Knuth, "The art of computer programming, vol. 3: Searching and sorting," *Reading MA: Addison-Wisley*, 1973.

[20] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, "Introduction to algorithms," 2009.

[21] W. W. Zachary, "An information flow model for conflict and fission in small groups," *Journal of anthropological research*, vol. 33, no. 4, pp. 452–473, 1977.

[22] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[23] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proceedings of the National Academy of Sciences*, vol. 102, no. 22, pp. 7794–7799, 2005.

[24] F. Chung and L. Lu, "Connected components in random graphs with given expected degree sequences," *Annals of combinatorics*, vol. 6, no. 2, pp. 125–145, 2002.

[25] E. A. Bender and E. R. Canfield, "The asymptotic number of labeled graphs with given degree sequences," *Journal of Combinatorial Theory, Series A*, vol. 24, no. 3, pp. 296–307, 1978.

[26] M. Molloy and B. Reed, "A critical point for random graphs with a given degree sequence," *Random structures & algorithms*, vol. 6, no. 2-3, pp. 161–180, 1995.

[27] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," *Science*, vol. 328, no. 5980, pp. 876–878, 2010.

[28] L. C. Freeman, S. P. Borgatti, and D. R. White, "Centrality in valued graphs: A measure of betweenness based on network flow," *Social networks*, vol. 13, no. 2, pp. 141–154, 1991.

7 APPENDICES

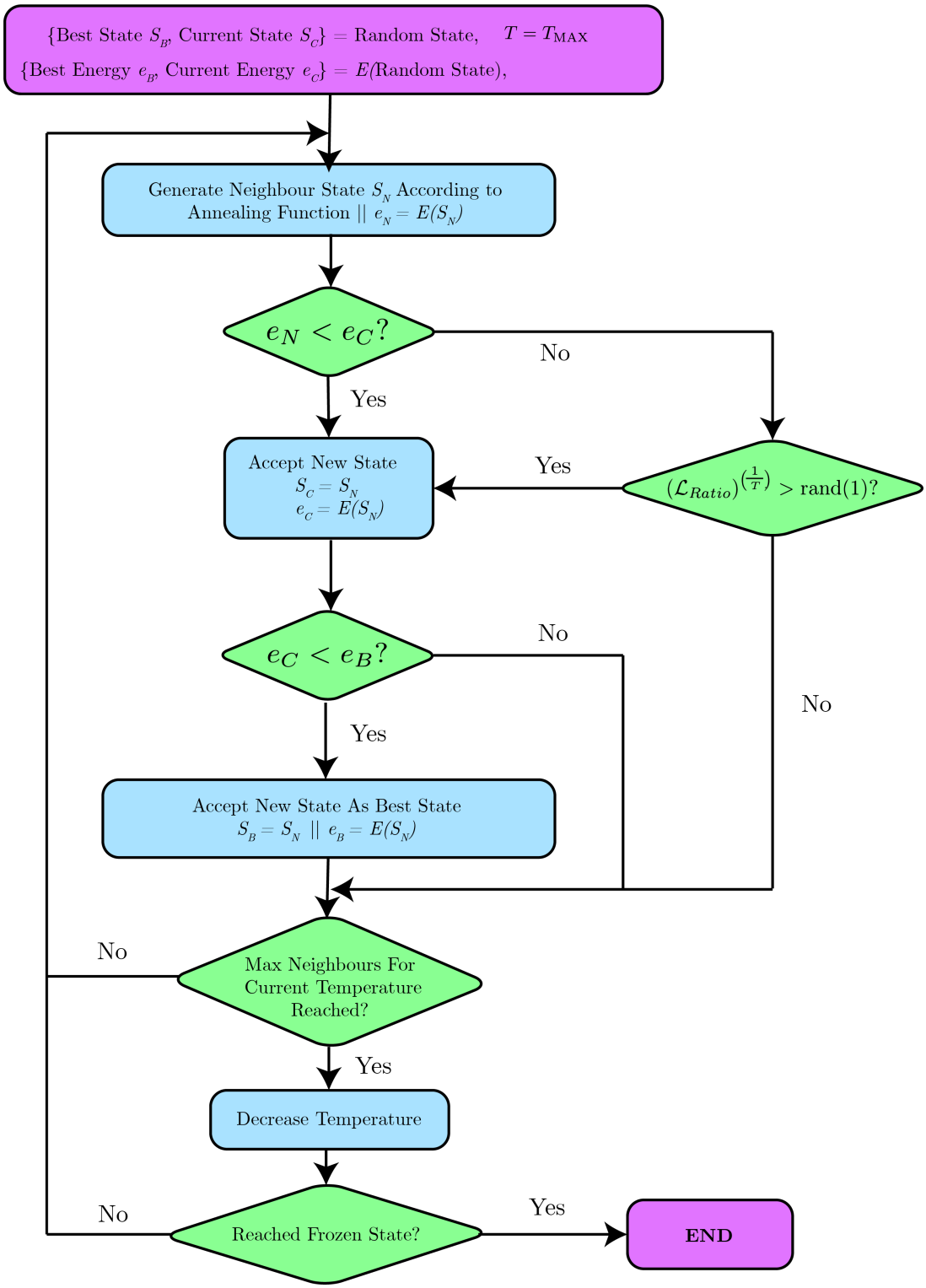


Figure 24: Schematic representation of the Simulated Annealing algorithm with short–term memory. Note "Frozen State" actually represents the minimum energy state and corresponds to the final state returned by the algorithm.

Supplementary Watts-Strogatz Network Results and Diagrams

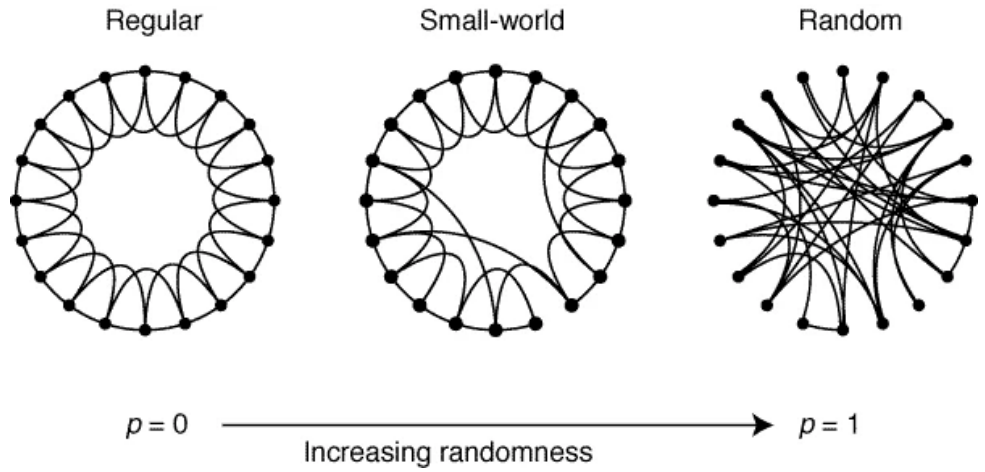
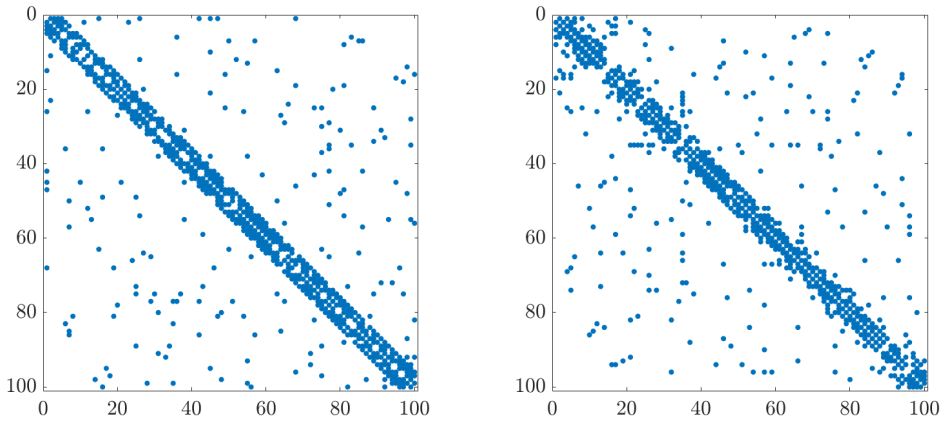


Figure 25: Example realisations of the Watts-Strogatz network to demonstrate the effect of rewiring probability p. Copied from [7].



(a) $\boldsymbol{A}_{ij}^{(\boldsymbol{s}_T)}$ (b) $\boldsymbol{A}_{ij}^{(\boldsymbol{s}_{SA})}$

Figure 26: Visualisation of non-zero entries of the ordered adjacency matrices for a 100 vertex network generated by the WS model with $z/2 = 4$, $p = 0.2$. For states $\boldsymbol{s}_T$ and $\boldsymbol{s}_{SA}$.

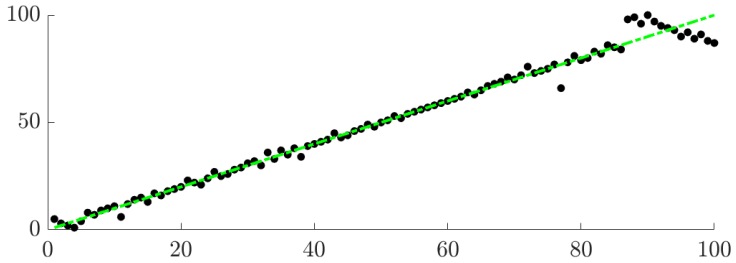


Figure 27: Original against achieved states for a 100 vertex network generated by the WS model with $z/2 = 4$, $p = 0.2$. Both originally and following a random permutation and reordering by the SA algorithm.

Supplementary Range-Dependent Random Network Results

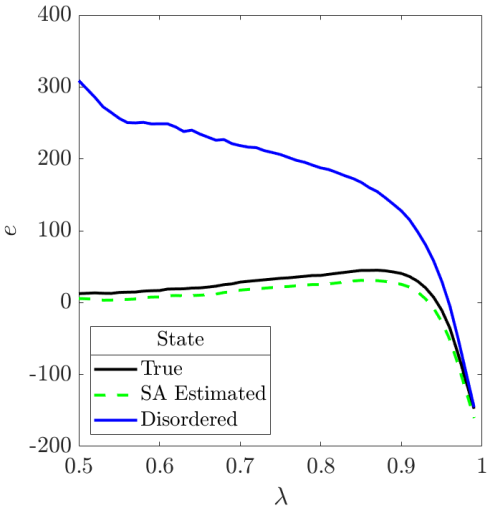


Figure 28: Mean energies e_T, e_D and e_{SA} for fixed $\alpha = 0.7$ and varying $\lambda \in [0.5, 1)$ for connected RDRNs with 30 vertices.

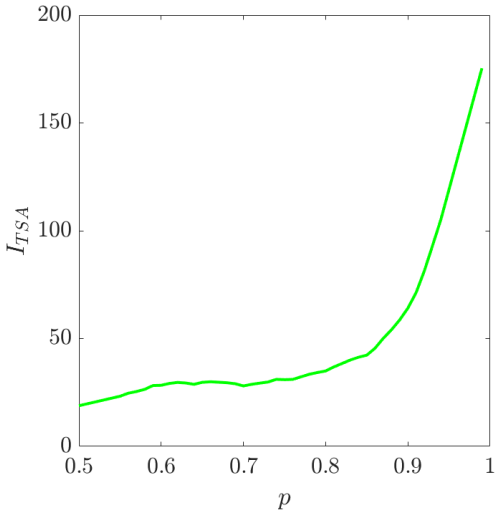


Figure 29: Mean inversions I_{TSA} of for linearly spaced points selected over $p \in [0.5, 1)$ for connected RDRNs of 30 vertices.