Stockholms universitet

Multi-Type Geometric Random Intersection Graphs

Thorfinn Laagus

Masteruppsats i matematisk statistik
Master Thesis in Mathematical Statistics

Masteruppsats 2024:8
Matematisk statistik
Juni 2024

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Stockholm University

Mathematical Statistics
Stockholm University
Master Thesis **2024:8**
http://www.math.su.se

Multi-Type Geometric Random Intersection Graphs

Thorfinn Laagus*

June 2024

Abstract

The geometric random intersection graph is a network model that connects vertices based on their geometric proximity to auxiliary vertices, which act as facilitators in the connection process. Building upon the recent work on the geometric random intersection graph, we define a generalised version of the model that includes multiple vertex types with differing connection functions. This added flexibility allows for the modelling of mixed populations with varying connection behaviours as well as differing types of facilitating vertices. We explore how this expanded definition affects the baseline properties of the model, including the general edge probability and expected degree, as well as the existence of non-trivial percolation transition parameters.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: thorfinn@tehis.net. Supervisor: Maria Deijfen.

Acknowledgements

I would like to thank my supervisor Maria Deijfen for valuable direction, guidance and feedback throughout this project.

Contents

1 Introduction

Networks are fundamental structures observable across a wide range of scientific disciplines. By constructing network models, we create crucial tools in understanding the interactions between entities in complex systems. While graph theory has existed for many centuries, the development of the theory behind random graphs only started in the mid twentieth century. Since then, the field has expanded with the formulation of a variety of modelling approaches. Still, given the complexity, diversity and scale of real-world networks, there remain many unexplored avenues within the field.

The geometric random intersection graph is a newly defined class of random graph model that integrates geometric proximity and group clustering structures into the connection process. The combination of these two features attempts to capture a portion of the intricate connection patterns inherent in empirical networks.

The specific geometric space that has been studied is the d-dimensional Euclidean space. The most obvious application of this geometry in network theory is geographic location. It is common for connection processes to be driven by location as nodes in a network are more likely to interact when they are closer together. While it seems quite natural to consider networks embedded on a geometric space, it adds an extra layer of complication as we must determine how to generate points on our chosen space as well as how to connect them according to proximity. Although we could consider fixing the location of points, a more flexible approach would be to generate points according to a random point process. In our model, we will only be concerned with the Poisson point process, but that does not rule out the possibility of using other point processes in other work.

In addition to accounting for geometric aspects of a network, we also use the intersection connection technique. This is used to improve the clustering structures within a model. Clustering in a network refers to the presence of subsets of nodes that have many connections to one other. There are many ways of connecting nodes that result in a highly clustered network. The intersection model recreates the scenario where nodes in a network connect through some auxiliary facilitator. To understand why we might want to do this, consider how social networks grow and evolve. Often, people tend to form connections through mutual membership to some kind of group or organisation, whether it be school, work, social clubs or some other structure. Therefore, to create representative models, we may want to incorporate an approximation of this procedure into our constructions.

The aim of this project is to build upon the theory of geometric random intersection graphs by exploring a generalisation of this model that includes the possibility of splitting the total population of the network, as well as these auxiliary facilitators, into subsets, each with different connection patterns. This introduces an element of inhomogeneity which allows for the modelling of mixed populations or mixed auxiliary structures, all while maintaining the core framework of the model.

In this paper, we define a model that has an arbitrary number of what we call primary vertex sets, which contain all of the nodes that will be present in the desired network, as well as an arbitrary number of auxiliary vertex sets, which contain vertices that will not be present in the desired network but will facilitate the connections between primary vertices. This is done by connecting primary and auxiliary vertices to one another using unique connection functions based on the vertex set memberships of both vertices before connecting two primary vertices if they both connect to the same auxiliary vertex on at least one occasion.

We study the process behind two primary vertices connecting to the same auxiliary vertex given that we know which vertex sets they belong to. This leads us to the conclusion that the number of auxiliary vertices shared by two primary vertices follows a Poisson distribution, the intensity of which is given by the sum of the intensities of the independent sub-models containing only one auxiliary vertex set, which were already known to be Poisson. The consequence of this is that two vertices in our model connect with a probability equal to the probability that this Poisson random variable is not zero. From this, we are able to characterise the expected degree, which we find to be the sum of the expected degrees of the sub-models including just one of the primary vertex sets.

The second part of our results relates to the percolation properties of the model, where we begin to investigate the component structure of the model. Since we are generating vertices throughout $\mathbb{R}^d$, we end up with a model that has an infinite number of vertices. This leads us to the question of how changing the parameter values of the model affects the probability that the largest connected component in the graph contains an infinite number of vertices, with the ultimate aim of identifying a transition boundary between parameter values that do not lead to an infinite component and those that do. This is done by fixing all parameters apart from one and finding the value at which the probability of an infinite component transitions to being strictly positive.

Results are initially given for models containing an arbitrary number of primary vertex sets along with just one auxiliary vertex set. These results show that the existence of a non-trivial phase transition, that is one that occurs in the open interval $(0, \infty)$, is guaranteed when there is at least one connection function that has unbounded support. Furthermore, for the case where all functions have bounded support, we find a sufficient condition on the fixed model parameters that gives an infinite critical value for the phase transition of the model. Finally, we show that when we swap the roles of primary and auxiliary vertex sets, the percolation properties of the model, now containing an arbitrary number of auxiliary vertex sets, are the same.

Before arriving at these results, we will explain all of the relevant background theory, starting with the fundamentals of random graph theory and progressing one-by-one through the various elements that will be needed to construct our model. We motivate the usefulness of each of these elements and illustrate how each layer of complexity affects the properties of the model. This leads into the definition and properties of the geometric random intersection graph with

generalised connection probabilities before we introduce our multi-type extension to the model. To reach a fully generalised model with arbitrary numbers of vertex types, the properties of the model are first developed for the addition of a single extra vertex set, first as primary vertices and then as auxiliary vertices. When looking at two types of each, the pattern becomes clear enough to expand the properties to the fully generalised model.

Having defined the multi-type geometric random intersection graph and investigated its core properties, we proceed to the topic of percolation. The work that has already been done on the geometric random intersection graph is introduced and explained before we present the results from our network model with multiple vertex sets. As was the case when exploring the fundamental properties of the model, we will build up to the final results by first understanding the effects of adding just one extra vertex set and then expanding to an arbitrary number.

2 Background

The models that we seek to construct and explore in this project are random networks. Random networks, or random graphs, build upon the mathematics of graph theory by considering such objects to be random. A graph is a fundamental construction in mathematics consisting of a set of vertices and a set of edges describing the connections between vertices. There are different ways of introducing randomness into a network model, the simplest of which is to consider a fixed number of vertices and assign some probability to the connection of each pair of vertices. Naturally, the probability that two vertices are connected becomes a central property of this construction.

2.1 Erdős-Rényi Graph

There are many ways that these probabilities can be defined, but perhaps the most natural starting point might be to connect each pair of vertices independently with some fixed probability, p, as introduced by Erdős and Rényi in 1960 [3]. From this construction, we can see that the event that two vertices connect is a Bernoulli random variable with probability p.

Another central property of networks is the degree of each vertex, that is the number of connections that each specific vertex has with other vertices in the graph. Given that each connection is random, this means that the degree of a vertex in a random graph is itself a random variable. If we have N vertices, the degree of an arbitrary vertex v can be expressed as

$$D = \sum_{i=1}^{N-1} \mathbb{1}_{\{v \leftrightarrow v_i\}},$$

where $\{v \leftrightarrow v_i\}$ denotes the event that the vertices v and v_i are connected. Notice here that because each pair of vertices connects independently with the same probability, p, the degree distribution of the Erdős-Rényi graph is just the sum of independent Bernoulli trials, which is a binomial distribution with $N-1$ trials and success probability p. Furthermore, this independence gives that the degree distribution is the same for every vertex in the graph.

However, for more complex networks, this distribution can be hard to characterise. In such cases, we can still investigate properties of the degree distribution in its stead. For example, we might consider the expected degree to help us understand the behaviour of an average vertex. If we observe that

$$\mathbb{E}[D] = \mathbb{E}\left[\sum_{i=1}^{N-1} \mathbb{1}_{\{v \leftrightarrow v_i\}}\right] = \sum_{i=1}^{N-1} \mathbb{E}\left[\mathbb{1}_{\{v \leftrightarrow v_i\}}\right] = \sum_{i=1}^{N-1} \mathbb{P}(v \leftrightarrow v_i),$$

then we can see that the expected degree is a quantity that relies exclusively on the connection probabilities between vertices, and thus is much more convenient to calculate. In the simple example we have looked at up to this point, we defined

the model such that these connection probabilities were a fixed constant p for all pairs of vertices. That is to say that

$$\mathbb{P}(v \leftrightarrow v_i) = p$$

for every pair of vertices. Therefore, the expected degree of the Erdős-Rényi graph is

$$\mathbb{E}[D] = (N-1)p,$$

the expectation of the aforementioned binomial distribution.

Due to the fact that the expected degree is dependent on the number of vertices in the graph, it is typical to define the connection probability as $p = \frac{c}{N}$ for some parameter c. This means that the expected degree is approximately c and is independent of the number of vertices. This is useful because it prevents the degree from diverging with the number of nodes in the network. In general, we want to construct graphs whose degree is finite, even when the number of vertices grows to infinity. This is because graphs where the expected degree diverges are not as interesting to study given how connected they are as well as being not very realistic when compared to most large scale empirical networks.

The Erdős-Rényi graph is one of the most basic models in random graph theory, making it simultaneously easy to analyse but too simple to effectively reflect the properties and behaviours of empirical networks.

2.2 Random Intersection Graph

When we construct new models, it is usually helpful to think about how the process or structures we are trying to model occur in the first place. For example, how realistic is it that individuals in a network have a uniform chance of connecting with any other individual in the network? When we observe networks empirically, one feature which is commonly found is the presence of 'clustering' or 'communities'. That is to say that there are often smaller subsets of vertices within a graph that have high connectivity. One reason for this feature might be that connections between individuals are often facilitated in an intermediate step by way of membership to some kind of group. This form of connection process is quite prevalent in social networks where people tend to connect through larger structures like schools, workplaces, social clubs, etc.

Introduced in 1995 [8], random intersection graphs are designed to mimic this connection process. For this model, we introduce a set of auxiliary vertices for the connection phase that will not ultimately be present in the final graph. The primary vertices, which are those that we are interested in studying, will be connected in two stages. The idea here is that we first connect vertices from the primary vertex set to vertices from the auxiliary vertex set randomly. The result of this is a bipartite graph between the set of primary and auxiliary vertices. A bipartite graph is a type of graph whose vertices can be divided into two disjoint sets such that no two vertices within the same set connect to one another. In the procedure that we described, we do not connect primary

vertices to each other, nor do we connect auxiliary vertices to each other. We only connect across the primary and auxiliary vertex set in the first stage and not within them, hence this graph is bipartite.

In the second step, we construct the final graph between vertices in the primary vertex set only. Each pair of primary vertices is connected in this final graph if there is at least one auxiliary vertex in the bipartite graph from the first step that connects to both primary vertices. In other words, two primary vertices are connected if there exists a path of length 2 between them in the bipartite graph, through some intermediate auxiliary vertex. The result of this is a graph over just the primary vertex set that includes local clustering structures in the form of fully connected subgraphs corresponding to each auxiliary vertex.

The randomness here comes exclusively from the first step, where primary and auxiliary vertices connect to one another with some probability. In its most basic form, we can connect each primary vertex to each auxiliary vertex independently with a fixed probability, p. The graph that arises from this first step alone can be thought of as a bipartite version of the simple Erdős-Rényi graph that was given as an example earlier. Therefore, the number of auxiliary vertices that a primary vertex connects to follows a binomial distribution with the number of auxiliary vertices as the number of trials and p as the success probability. Furthermore, the probability that two arbitrary primary vertices both connect to an arbitrary auxiliary vertex is p^2 as each primary vertex connects to an auxiliary vertex independently. Thus, we can also see that the number of auxiliary vertices that two arbitrary primary vertices both connect to is binomial with the number of auxiliary vertices as the number of trials and p^2 as the success probability.

Then, in the final graph constructed in the second step, the probability that any pair of primary vertices are connected is the same as the probability that there is at least one auxiliary vertex that both of them connect to in the first step. Since the number of auxiliary vertices that two arbitrary primary vertices both connect to is binomial, if we have M auxiliary vertices then the probability that two primary vertices connect in the second step is the probability that a binomial distribution with M trials and p^2 success probability is non-zero. This probability is thus

$$\mathbb{P}(v \leftrightarrow v_i) = 1 - (1 - p^2)^M.$$

As we have seen, the expected degree of a primary vertex in the final graph follows directly from these connection probabilities. Therefore, the expected degree for the basic random intersection graph is

$$\mathbb{E}[D] = (N - 1)\left(1 - (1 - p^2)^M\right).$$

As with the Erdős-Rényi graph, this expression for the expected degree depends on the number of vertices in the graph. To remove these dependencies, we could consider setting $M = aN$ and $p = \frac{b}{N}$, introducing two new constants, a and b. We can understand the asymptotic behaviour of the properties of the

model by approximating the connection probability using a Taylor expansion for a suitably small value of p, which gives

$$1 - (1 - p^2)^M = Mp^2 + O(M^2 p^4) = \frac{ab^2}{N} + O\left(\frac{1}{N^2}\right).$$

Therefore, the probability is of order $\frac{c}{N}$, where we have a constant $c = ab^2$. Inserting this approximation for the connection probability into the expression for the expected degree gives

$$\mathbb{E}[D] = (N-1)\left(\frac{ab^2}{N} + O\left(\frac{1}{N^2}\right)\right) = \frac{ab^2(N-1)}{N} + O\left(\frac{1}{N}\right),$$

which we can see will tend to the constant $c = ab^2$ as $n \to \infty$. This shows that it is also possible to control the expected degree of the random intersection graph so that the model scales well with the number of vertices in the graph.

Overall, the important thing to note here is that when we add an intersection step to a network model, the connection probabilities can be found by taking the complement of the probability that two primary vertices have no mutual connections among the auxiliary vertices in the bipartite graph.

2.3 Random Connection Model

Another aspect of empirical networks that we may be interested in is the effect of the geometric properties of vertices. It is often the case that nodes in a network have some kind of location. It is perhaps most intuitive to think about this in terms of geographic location in two dimensions. For example, people tend to be more likely to form connections with those who live nearby.

The random connection model is used for constructing a graph on a set of points lying in $\mathbb{R}^d$. In this model, vertices are taken to be the set of points from a stationary point process. These vertices are connected by some connection function, g, which is a mapping from some distance defined on $\mathbb{R}^d$ to a probability [5].

The simplest and most commonly used point process is the Poisson point process. This process generates points uniformly and independently over a space, with the number of points in any given subset of space adhering to a Poisson distribution.

More formally, let N be a point process on $\mathbb{R}^d$ and let $\mathcal{B}(\mathbb{R}^d)$ denote the Borel σ-algebra on $\mathbb{R}^d$. Then, N is a Poisson process with intensity λ if

1. For any Borel set $B \in \mathcal{B}(\mathbb{R}^d)$, the number of points in B, $N(B)$, follows a Poisson distribution with parameter $\lambda \cdot \text{vol}(B)$, where $\text{vol}(B)$ is the d-dimensional volume of B.

2. If $B_1, \ldots, B_n$ are disjoint Borel sets in $\mathcal{B}(\mathbb{R}^d)$, then $N(B_1), \ldots, N(B_n)$ are independent random variables. [7]

Then, to define a random connection model on a point process in $\mathbb{R}^d$, we require a connection function $g : \mathbb{R}^d \to [0, 1]$, which is non-increasing. Therefore, in this model, the connection probabilities are defined as

$$\mathbb{P}(v \leftrightarrow v_i) = g(v - v_i).$$

By defining the connection function in terms of the difference between the positions of the vertices, it is useful to assume that g is a radial function. A radial function is a function whose value depends only on the distance of the argument from the origin. Consequently $g(v - u) = g(u - v)$ under this assumption.

Notice that the number of points generated by a Poisson process in $\mathbb{R}^d$ is infinite and connection probabilities are no longer uniform across all pairs of vertices. This prevents us from calculating the expected degree as we did before. Instead, we will need to further explore the subject of Poisson point processes. Let $\mathcal{V}$ denote the set of points generated by a Poisson process with intensity λ. Furthermore, let us define $\mathcal{V}_v \subset \mathcal{V}$ to be the subset of points that connect to some vertex v. By definition, the number of elements in $\mathcal{V}_v$ is the degree of v.

First, notice that we can arrive at $\mathcal{V}_v$ from $\mathcal{V}$ by including each $v_i \in \mathcal{V}$ with probability $\mathbb{P}(v \leftrightarrow v_i)$. This is known as thinning a Poisson process and it is known that the result of this procedure is also a Poisson process where the new intensity is the product of the intensity of the original process and the probability of including a point in the thinned process [7]. Therefore, we can say that $\mathcal{V}_v$ is a Poisson process with intensity $\lambda\mathbb{P}(v \leftrightarrow x)$, where we can substitute in the connection probability that we have already found to get $\lambda g(v - x)$.

We have here an added complication in that the intensity of this thinned Poisson process is a function of x rather than a constant as we defined earlier. When a Poisson process has an intensity that is not constant, it is called inhomogeneous, or non-homogeneous. When we define an inhomogeneous Poisson process with intensity function $\lambda(x)$, we say that for any Borel set, $B \in \mathcal{B}(\mathbb{R}^d)$, the number of points in B, denoted by $N(B)$, follows a Poisson distribution with parameter $\int_B \lambda(x)dx$ [6]. From this, we can see that the expected number of points in an inhomogeneous Poisson process with intensity $\lambda(x)$ is the expectation of a Poisson random variable with intensity $\int_{\mathbb{R}^d} \lambda(x)dx$.

Now we can return to $\mathcal{V}_v$, the set of points that connect to v. We stated previously that the number of points in this set is the degree of our vertex v. But since this is the point set of an inhomogeneous Poisson process with intensity $\lambda g(v - x)$, the number of points in this set follows a Poisson distribution with intensity $\lambda \int_{\mathbb{R}^d} g(v - x)dx$. Thus, the expected degree of a random connection model on a homogeneous Poisson process for some vertex v is

$$\mathbb{E}[D_v] = \lambda \int_{\mathbb{R}^d} g(v - x)dx,$$

since the expectation of a Poisson random variable is its intensity parameter.

As we have seen with the previous models, we want to control the expected degree to ensure that it remains finite, even when the number of vertices is not. An easy way to enforce this in a random connection model is to require that

$\int_{\mathbb{R}^d} g(x)dx < \infty$. It is clear to see from the expression for the expected degree that this condition will achieve this.

A simple example of a random connection model is the Poisson Boolean model. In the Poisson Boolean, we generate points in $\mathbb{R}^d$ using a Poisson process and then connect points if and only if they are within some distance r of each other [4]. This means that a point connects to every other point within the ball of radius r centered at its location. The connection function is therefore $g(x) = \mathbb{1}_{\{x<r\}}$. In addition to being an illustrative example of how we can think about connection functions in the random connection model, the Poisson Boolean model will be useful as a reference model when we come to the topic of percolation later on.

Relating to the content of the current section, we have introduced the idea of generating vertices through a Poisson process as well the challenges this presents when finding an expression for the expected degree. This challenge can be overcome through the technique of thinning the Poisson process according to the connection probability.

2.4 Geometric Random Intersection Graph

By combining the random connection model with the random intersection graph, we can construct a hybrid model that incorporates both geometric and clustering structures. The random connection model is itself capable of producing some level of clustering based on the distances between points in the point process, but this can be improved upon by integrating an intersection step to create more natural clusters.

Here we consider a vertex set $\mathcal{V}$ to be the set of points from a Poisson point process on $\mathbb{R}^d$ with intensity λ. To make this an intersection graph, we also require an auxiliary vertex set $\mathcal{U}$, which we generate using another Poisson point process independent of $\mathcal{V}$, this time with intensity μ.

Just as with the random connection model, we connect vertices using some non-increasing, radial connection function $g : \mathbb{R}^d \to [0, 1]$. For every combination of $v \in \mathcal{V}$ and $u \in \mathcal{U}$ we connect them with probability $g(v - u)$. One final constraint we will place on the connection function is that $\int_{\mathbb{R}^d} g(x)dx < \infty$, which will help control the expected degree. In the same way as the non-geometric random intersection graph, the result of the connection procedure is a bipartite graph between primary and auxiliary vertex sets.

The second step is to connect primary vertices with each other. We do this with the same rule as the non-geometric random intersection graph, which is that two vertices $v_1, v_2 \in \mathcal{V}$ are connected if there exists some $u \in \mathcal{U}$ that connects directly to both of them. With this, we arrive at our final graph with vertex set $\mathcal{V}$.

In [2], the connection probability and expected degree of the model are given. When characterising these properties, the authors place a primary vertex at the origin and base their results around this vertex. Adding this point to the Poisson process generating the primary vertices creates what is known as a Palm version of the process. Formally, the Palm version is conditioned to have a point at the

origin and, for a Poisson process, can be shown to be equal in distribution to the original process with this added point at the origin [1]. Including a point at the origin allows us to view the properties of the graph from the perspective of a vertex, which is helpful because the characteristics of the model do not rely on the position of the vertex, only the distance between vertices. Therefore, this choice simplifies the notation slightly without changing the results. We will follow the same convention when generalising these properties later on.

Let 0 denote the vertex at the origin. To determine the probability that this vertex connects to some other vertex $v \in \mathcal{V}$, we need to find the probability they both connect to the same auxiliary vertex at least once. For some auxiliary vertex $u \in \mathcal{U}$, the probability that 0 connects to u is $g(u)$ while the probability that v and u are connected is $g(v-u)$. Given that they connect independently, the probability that both 0 and v are connected to u is $g(u)g(v-u)$.

If we can determine the number of auxiliary vertices that both 0 and v are connected to, we can simply compute the probability that this number is not zero to get the connection probability. Define $\mathcal{U}_{0,v} \subseteq \mathcal{U}$ to be the subset of auxiliary vertices that both 0 and v connect to and recall that we can construct $\mathcal{U}_{0,v}$ from $\mathcal{U}$ by thinning the Poisson process corresponding to $\mathcal{U}$ according to the probability that 0 and v both connect to some point at x. This means that $\mathcal{U}_{0,v}$ is the point set of a Poisson process with intensity $\mu g(x)g(v-x)$. It follows then that the number of auxiliary vertices that both 0 and v connect to adheres to a Poisson distribution with intensity $\mu \int_{\mathbb{R}^d} g(x)g(v-x)dx$.

The final step is to find the probability that the number of auxiliary vertices shared by 0 and v is non-zero. Since this number has a Poisson distribution, the general connection probability is thus,

$$\mathbb{P}(0 \leftrightarrow v) = 1 - e^{-\mu \int_{\mathbb{R}^d} g(x)g(v-x)dx}.$$

The process of finding the expected degree of the zero-vertex, 0, is similar to what we saw in the random connection model, which is where we originally introduced the concept of thinning a Poisson process. Recall that the set of primary vertices that 0 connects to is the point set of an inhomogeneous Poisson process with intensity $\lambda\mathbb{P}(0 \leftrightarrow x)$ and therefore the number of vertices that 0 connects to follows a Poisson distribution with intensity $\lambda \int_{\mathbb{R}^d} \mathbb{P}(0 \leftrightarrow x)dx$. Using that the expectation of a Poisson variable is its intensity and substituting in the expression for the connection probability, we get that the expected degree of the geometric random intersection graph is

$$\mathbb{E}[D] = \lambda \int_{\mathbb{R}^d} \left(1 - e^{-\mu \int_{\mathbb{R}^d} g(x)g(v-x)dx}\right) dv.$$

3 The Multi-Type Model

One limitation of the geometric random intersection graph might be in modelling networks of mixed populations where different members of the network exhibit different connection characteristics. In the following section we will generalise the geometric random intersection graph to consider the case where the vertex sets are generated from independent combined Poisson processes, starting with the addition of one extra vertex set and leading up to an arbitrary number of vertex sets.

3.1 Two Primary Vertex Sets

We will begin by replacing the primary vertex set of the geometric random intersection graph with two primary vertex sets, $\mathcal{V}_1$ and $\mathcal{V}_2$, generated from independent Poisson processes with intensities λ_1 and λ_2 respectively. Here we maintain the single auxiliary vertex set $\mathcal{U}$. This means that our primary vertices are now split across two sets generated by different Poisson processes.

Notice that, at this point, we have not made any consequential changes to the model. This is because if we consider the combined primary vertex set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$, this is still the point set of a Poisson process, now with intensity $\lambda_1+\lambda_2$. This property of combined Poisson processes will be useful later on when we begin deriving our results, but for now it still remains to construct a model where different types of vertices exhibit different connection patterns. This is done by defining two connection functions, g_1 and g_2, each being non-increasing radial functions from $\mathbb{R}^d$ to a probability as before.

As one might expect, we will connect primary vertices in $\mathcal{V}_1$ to auxiliary vertices in $\mathcal{U}$ using the function g_1 and primary vertices in $\mathcal{V}_2$ to auxiliary vertices in $\mathcal{U}$ using the function g_2. This will mean that primary vertices will connect to auxiliary vertices differently depending on which primary vertex set they belong to.

The second step in constructing the geometric random intersection graph remains unchanged. We no longer need to distinguish which vertex set each primary vertex belongs to and we can simply connect two vertices from the combined primary vertex set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ if they both connect to the same auxiliary vertex at least once.

With this extension to the geometric random intersection graph now defined, we can turn to extending the properties of the model relating to an arbitrary vertex. As in [2], we will consider a point located at the origin, denoted by 0. In this new model, having two sets of primary vertices adds an extra layer of complication to determining these properties as this will affect the functions used to connected them to auxiliary vertices. To make things simpler, we will assume that we know which point set each primary vertex belongs to.

First, note that $\mathcal{V}_1$ and $\mathcal{V}_2$ are independent and that vertices in $\mathcal{V}_1$ are connected to auxiliary vertices in $\mathcal{U}$ independently of vertices in $\mathcal{V}_2$. Therefore, it follows directly from the basic geometric random intersection graph that the

auxiliary vertices shared by 0 and v form an inhomogeneous Poisson process on $\mathbb{R}^d$ with intensity

$$\mu g_k(x) g_l(v-x)$$

if $0 \in \mathcal{V}_k$ and $v \in \mathcal{V}_l$, with $k, l \in \{1, 2\}$. This expression is similar to the single primary vertex set case except we now have up to two different functions contributing to this intensity depending on which primary vertex sets 0 and v belong to.

Once the auxiliary vertices shared by 0 and v have been determined, the probability that they are connected in the final graph can be found in the same way as before because the second step of the graph's construction is unchanged in this new model. Therefore, we have that

$$\mathbb{P}(0 \leftrightarrow v | 0 \in \mathcal{V}_k, v \in \mathcal{V}_l) = 1 - e^{-\mu \int_{\mathbb{R}^d} g_k(x) g_l(v-x) dx}$$

with, once again, $k, l \in \{1, 2\}$.

Finally, we can find the expected degree. We will still assume that we know the primary vertex set which the zero vertex belongs to. Notice that a connection between two primary vertices is still independent of all other primary vertices in this model. Therefore, if we focus on the subgraph that contains only the primary vertices from $\mathcal{V}_1$, this is exactly the geometric random intersection graph with a single primary vertex set and connection function g_1. The vertices from $\mathcal{V}_2$ have no effect on how vertices from $\mathcal{V}_1$ connect to each other, nor does the function g_2. This property is of course mirrored for vertices from $\mathcal{V}_2$.

As it relates to the expected degree, let D_1 be the number of vertices in $\mathcal{V}_1$ that connect to the zero vertex and let D_2 be the equivalent quantity for vertices in $\mathcal{V}_2$. The zero vertex and all primary vertices in $\mathcal{V}_1$ form a geometric random intersection graph. Thus, by using the expression for the expected degree of a geometric random intersection graph with a single primary vertex set, the expectation of D_1 is

$$\mathbb{E}[D_1 | 0 \in \mathcal{V}_k] = \lambda_1 \int_{\mathbb{R}^d} \left(1 - e^{-\mu \int_{\mathbb{R}^d} g_k(x) g_1(v-x) dx}\right) dv$$

Likewise, expectation of D_2 is

$$\mathbb{E}[D_2 | 0 \in \mathcal{V}_k] = \lambda_2 \int_{\mathbb{R}^d} \left(1 - e^{-\mu \int_{\mathbb{R}^d} g_k(x) g_2(v-x) dx}\right) dv$$

Finally, notice that the degree of the zero vertex over all primary vertices is $D = D_1 + D_2$. The linearity of expectation gives us that

$$\begin{aligned}\mathbb{E}[D | 0 \in \mathcal{V}_k] =& \lambda_1 \int_{\mathbb{R}^d} \left(1 - e^{-\mu \int_{\mathbb{R}^d} g_k(x) g_1(v-x) dx}\right) dv \\ &+ \lambda_2 \int_{\mathbb{R}^d} \left(1 - e^{-\mu \int_{\mathbb{R}^d} g_k(x) g_2(v-x) dx}\right) dv.\end{aligned}$$

Thus, we have shown that splitting our primary vertex set into two disjoint subsets generated by independent Poisson processes and connected to auxiliary vertices using different connection functions does not add much complexity to the basic properties of the model.

3.2 Two Auxiliary Vertex Sets

We can now look at the 'opposite' case where we return to having a single primary vertex set, $\mathcal{V}$, and instead introduce two auxiliary vertex sets, $\mathcal{U}_1$ and $\mathcal{U}_2$, generated from independent Poisson processes with intensities μ_1 and μ_2 respectively. Since these are independent Poisson processes, we can consider the point set of combined process, $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$, with intensity $\mu_1 + \mu_2$.

We still have two connection functions, g_1 and g_2, with primary vertices connecting to auxiliary vertices from $\mathcal{U}_1$ using the function g_1 and auxiliary vertices from $\mathcal{U}_2$ using the function g_2. Then, in step 2, primary vertices in $\mathcal{V}$ connect to each other if they both connect to the same auxiliary vertex from $\mathcal{U} = \mathcal{U}_1 \cup \mathcal{U}_2$ at least once, regardless of which of the two vertex sets the auxiliary vertex comes from.

Let $N^{0,v}$ be the number of auxiliary vertices shared by the zero vertex and some other vertex v. Furthermore, let $N_1^{0,v}$ be the number of auxiliary vertices from $\mathcal{U}_1$ shared by 0 and v and let $N_2^{0,v}$ be the number of auxiliary vertices from $\mathcal{U}_2$ shared by 0 and v, so that $N^{0,v} = N_1^{0,v} + N_2^{0,v}$. Since $\mathcal{U}_1$ and $\mathcal{U}_2$ are generated independently, $N_1^{0,v}$ and $N_2^{0,v}$ are independent random variables.

We have seen that when we have one primary vertex set and one auxiliary vertex set, the auxiliary vertices that 0 and v share constitute an inhomogeneous Poisson process with intensity $\mu g(x) g(v-x)$ and this means that the number of auxiliary vertices shared by 0 and v follows a Poisson distribution with intensity

$$\mu \int_{\mathbb{R}^d} g(x) g(v-x) dx.$$

Therefore, $N_1^{0,v}$, the number of auxiliary vertices from $\mathcal{U}_1$ shared by 0 and v, follows a Poisson distribution with intensity

$$\mu_1 \int_{\mathbb{R}^d} g_1(x) g_1(v-x) dx$$

and likewise $N_2^{0,v}$ also has a Poisson distribution with intensity

$$\mu_2 \int_{\mathbb{R}^d} g_2(x) g_2(v-x) dx.$$

And given that $N^{0,v} = N_1^{0,v} + N_2^{0,v}$, the total number of auxiliary vertices shared by 0 and v also follows a Poisson distribution with intensity

$$\mu_1 \int_{\mathbb{R}^d} g_1(x) g_1(v-x) dx + \mu_2 \int_{\mathbb{R}^d} g_2(x) g_2(v-x) dx,$$

since the sum of Poisson random variables is also Poisson.

Now that we know that the number of auxiliary vertices shared by 0 and v still follows a Poisson distribution, we can deduce that the probability that 0 and v connect is

$$\mathbb{P}(0 \leftrightarrow v) = 1 - e^{-\mu_1 \int_{\mathbb{R}^d} g_1(x)g_1(v-x)dx - \mu_2 \int_{\mathbb{R}^d} g_2(x)g_2(v-x)dx}.$$

With a single primary vertex set, it has already been shown that the expected degree follows from the individual connection probabilities, and so we get

$$\mathbb{E}[D] = \lambda \int_{\mathbb{R}^d} \left(1 - e^{-\mu_1 \int_{\mathbb{R}^d} g_1(x)g_1(v-x)dx - \mu_2 \int_{\mathbb{R}^d} g_2(x)g_2(v-x)dx}\right) dv.$$

Our final generalisation of this model will be to have an arbitrary number of primary and auxiliary vertex sets. In order to develop this model, we should first try to understand how the model behaves when we have just two primary vertex sets and two auxiliary vertex sets. The properties of this model can then inform our conclusions about the most general case.

3.3 Combining Two Primary and Auxiliary Vertex Sets

Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be our primary vertex sets generated with intensities λ_1 and λ_2 respectively and let $\mathcal{U}_1$ and $\mathcal{U}_2$ be our auxiliary vertex sets with intensities μ_1 and μ_2. To be able to connect vertices from both primary vertex sets with vertices from both auxiliary vertex sets, we will need to introduce further connection functions, giving us four in total. Define g_{11} for $\mathcal{V}_1$ and $\mathcal{U}_1$, g_{12} for $\mathcal{V}_1$ and $\mathcal{U}_2$, g_{21} for $\mathcal{V}_2$ and $\mathcal{U}_1$ and g_{22} for $\mathcal{V}_2$ and $\mathcal{U}_2$.

With these definitions we can proceed to constructing the model as before, where primary and auxiliary vertices connect according to the connection function corresponding to their respective vertex set memberships and then primary vertices are connected if they share at least one auxiliary vertex.

To find the connection probabilities between primary vertices, we will once again investigate the distribution of the number of auxiliary vertices shared by those two primary vertices. Given that we have returned to a situation with multiple primary vertex sets, we will continue with conditioning on the vertex sets that each primary vertex belongs to. Let $N^{0,v}|0 \in \mathcal{V}_k, v \in \mathcal{V}_l$ denote the number of auxiliary vertices shared by the zero vertex and some other vertex given that we know which vertex sets they come from, where k and l are either 1 or 2.

Just as in the previous case, we can split this quantity into two parts, $N_1^{0,v}|0 \in \mathcal{V}_k, v \in \mathcal{V}_l$ and $N_2^{0,v}|0 \in \mathcal{V}_k, v \in \mathcal{V}_l$, corresponding to the two auxiliary vertex sets we have in this model. Through conditioning on the primary vertex types and splitting the auxiliary vertex types, we can see that these two variables follow Poisson distributions with intensities

$$\mu_1 \int_{\mathbb{R}^d} g_{k1}(x)g_{l1}(v-x)dx$$

and

$$\mu_2 \int_{\mathbb{R}^d} g_{k2}(x) g_{l2}(v-x) dx$$

respectively. Thus $N^{0,v}|0 \in \mathcal{V}_k, v \in \mathcal{V}_l$ also has a Poisson distribution with the sum of these intensities as its parameter.

It follows that the conditional connection probability is

$$\mathbb{P}(0 \leftrightarrow v|0 \in \mathcal{V}_k, v \in \mathcal{V}_l) = 1 - e^{-\mu_1 \int_{\mathbb{R}^d} g_{k1}(x) g_{l1}(v-x) dx - \mu_2 \int_{\mathbb{R}^d} g_{k2}(x) g_{l2}(v-x) dx},$$

and subsequently, the conditional expected degree is

$$\begin{aligned}\mathbb{E}[D|0 \in \mathcal{V}_k] =& \lambda_1 \int_{\mathbb{R}^d} \left(1 - e^{-\mu_1 \int_{\mathbb{R}^d} g_{k1}(x) g_{11}(v-x) dx - \mu_2 \int_{\mathbb{R}^d} g_{k2}(x) g_{12}(v-x) dx}\right) dv \\ &+ \lambda_2 \int_{\mathbb{R}^d} \left(1 - e^{-\mu_1 \int_{\mathbb{R}^d} g_{k1}(x) g_{21}(v-x) dx - \mu_2 \int_{\mathbb{R}^d} g_{k2}(x) g_{22}(v-x) dx}\right) dv.\end{aligned}$$

The expressions for our properties are becoming lengthier as we add more vertex sets to the model, but at this point we can start to see patterns emerging as to how exactly they change. By conditioning on primary vertex set membership and splitting the number of auxiliary vertices shared by two primary vertices, we can always relate these properties back to combinations of the simplest version of the model with one type both primary and auxiliary vertices.

3.4 Generalised $n \times m$ Model

This leads us to the final step in characterising multi-type geometric random intersection graphs with independent combined Poisson processes, which is to expand to an arbitrary number of primary and auxiliary vertex sets. For this setup, let $\mathcal{V}_1, \dots \mathcal{V}_n$ be n independent Poisson processes on $\mathbb{R}^d$ with intensities $\lambda_1, \dots \lambda_n$, and let $\mathcal{U}_1, \dots \mathcal{U}_m$ be m additional independent Poisson processes on $\mathbb{R}^d$ with intensities $\mu_1, \dots \mu_m$. We will then define $n \times m$ non-increasing radial connection functions, $g_{ij} : \mathbb{R}^d \to [0,1]$, with $i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$, one for each combination of primary and auxiliary vertex sets, $\mathcal{V}_i$ and $\mathcal{U}_j$.

Once again, the graph is constructed in two steps, connecting primary and auxiliary vertices, and then using the auxiliary vertices to connect primary vertices with each other. First, for every $v \in \mathcal{V}_i$ and $u \in \mathcal{U}_j$, connect v and u with probability $g_{ij}(v-u)$. Repeat this for all $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, m\}$. The resulting graph is a bipartite graph between the combined primary vertex set, $\mathcal{V} = \bigcup_{i=1}^n \mathcal{V}_i$, and the combined auxiliary vertex set, $\mathcal{U} = \bigcup_{j=1}^m \mathcal{U}_j$.

Then, for every combination $v_1, v_2 \in \mathcal{V}, v_1 \neq v_2$, connect v_1 and v_2 if there exists some $u \in \mathcal{U}$ such that u connects to both v_1 and v_2. This can be done irrespective of which Poisson processes v_1, v_2 and u are generated from.

When considering the properties of this model, we still condition on which of the n vertex sets our vertices belong to. We have seen that the number of

auxiliary vertices from a single auxiliary vertex set, $\mathcal{U}_j$, that are shared by 0 and v has a Poisson distribution with intensity

$$\mu_j \int_{\mathbb{R}^d} g_{kj}(x) g_{lj}(v-x) dx.$$

Using the properties of Poisson distributions, we can take the sum over all auxiliary vertex sets to find distribution of the total number of auxiliary vertex sets.

Result: Number of Mutual Auxiliary Vertices

The total number of auxiliary vertices connected to both 0 and v has a Poisson distribution with intensity parameter

$$\sum_{j=1}^{m} \mu_j \int_{\mathbb{R}^d} g_{kj}(x) g_{lj}(v-x) dx.$$

It follows that the conditional connection probability is the probability that the aforementioned Poisson random variable is not zero.

Result: Connection Probability

$$\mathbb{P}(0 \leftrightarrow v | 0 \in \mathcal{V}_k, v \in \mathcal{V}_l) = 1 - e^{-\sum_{j=1}^{m} \mu_j \int_{\mathbb{R}^d} g_{kj}(x) g_{lj}(v-x) dx}$$

To find the conditional expected degree, we use the idea that the degree of a vertex over the total primary vertex set $\mathcal{V}$ can be split into the sum of the degrees of that vertex with each of the n primary vertex sets. This means that we can express the conditional expected degree as

$$\mathbb{E}[D | 0 \in \mathcal{V}_k] = \sum_{i=1}^{n} \mathbb{E}[D_i | 0 \in \mathcal{V}_k].$$

where D_i represents the degree of the zero vertex when considering only vertices from $\mathcal{V}_i$. Thus, we get the following conditional expected degree.

Result: Expected Degree

$$\mathbb{E}[D | 0 \in \mathcal{V}_k] = \sum_{i=1}^{n} \lambda_i \int_{\mathbb{R}^d} (1 - e^{-\sum_{j=1}^{m} \mu_j \int_{\mathbb{R}^d} g_{kj}(x) g_{ij}(v-x) dx}) dv$$

3.5 Removing the Vertex Set Membership Condition

All of the vertex properties we have studied in this section have relied upon knowing which vertex set these vertices belong to. This is a fairly reasonable to assume that we might have this information when studying a network. If we do have this information, then we are most likely to be interested in the properties pertaining to vertices from a specific vertex set. Nonetheless, we could still consider the unconditional properties of the model, where the vertex set memberships are not known.

We can start by looking at the unconditional probability $\mathbb{P}(0 \leftrightarrow v)$. Since the set of intersecting events $\{0 \in \mathcal{V}_k\} \cap \{v \in \mathcal{V}_l\}$ for all $k, l = 1, \ldots, n$ are mutually exclusive, we can say that

$$\mathbb{P}(0 \leftrightarrow v) = \sum_{k=1}^{n} \sum_{l=1}^{n} \mathbb{P}(0 \leftrightarrow v | 0 \in \mathcal{V}_k, v \in \mathcal{V}_l) \mathbb{P}(0 \in \mathcal{V}_k, v \in \mathcal{V}_l).$$

We have already found $\mathbb{P}(0 \leftrightarrow v | 0 \in \mathcal{V}_k, v \in \mathcal{V}_l)$, so we just require an expression for $\mathbb{P}(0 \in \mathcal{V}_k, v \in \mathcal{V}_l)$, the joint membership probability.

If we first look at a single membership probability $\mathbb{P}(v \in \mathcal{V}_l)$, recall that by the definition of a Poisson process with intensity λ_l, we expect an average of λ_l points to occur in a unit volume of $\mathbb{R}^d$. This is true for all n Poisson processes, and thus we expect a total of $\sum_{i=1}^{n} \lambda_i$ points over this unit area. Therefore, the probability of selecting a random point from a unit area is just the ratio of these quantities. Finally, since the numbers of points in disjoint regions are independent in a Poisson process, this probability holds for any and all areas. This allows us to conclude that the membership probability for a single primary vertex is

$$\mathbb{P}(v \in \mathcal{V}_l) = \frac{\lambda_l}{\sum_{h=1}^{n} \lambda_h}.$$

These membership probabilities are independent, which allows us to take their product to get the joint membership probability for two vertices

$$\mathbb{P}(0 \in \mathcal{V}_k, v \in \mathcal{V}_l) = \frac{\lambda_k \lambda_l}{\left(\sum_{h=1}^{n} \lambda_h\right)^2}.$$

This expression for the joint membership probability can be used directly in the previous expression for the unconditional probability, giving that

$$\mathbb{P}(0 \leftrightarrow v) = \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\lambda_k \lambda_l}{\left(\sum_{h=1}^{n} \lambda_h\right)^2} \left(1 - e^{-\sum_{j=1}^{m} \mu_j \int_{\mathbb{R}^d} g_{kj}(x) g_{lj}(v-x) dx}\right).$$

From the results for the conditional expected degree, notice that if we have a connection probability in the multi-type model, we can express the expected degree as

$$\mathbb{E}[D] = \sum_{i=1}^{n} \lambda_i \int_{\mathbb{R}^d} \mathbb{P}(0 \leftrightarrow v) dv.$$

Using this, we arrive at the follow expression for the unconditional expected degree of the multi-type geometric random intersection graph

$$\begin{aligned}\mathbb{E}[D] =& \sum_{i=1}^{n} \lambda_i \int_{\mathbb{R}^d} \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\lambda_k \lambda_l}{\left(\sum_{h=1}^{n} \lambda_h\right)^2} \left(1 - e^{-\sum_{j=1}^{m} \mu_j \int_{\mathbb{R}^d} g_{kj}(x) g_{lj}(v-x) dx}\right) dv \\ =& \sum_{i=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} \frac{\lambda_i \lambda_k \lambda_l}{\left(\sum_{h=1}^{n} \lambda_h\right)^2} \int_{\mathbb{R}^d} \left(1 - e^{-\sum_{j=1}^{m} \mu_j \int_{\mathbb{R}^d} g_{kj}(x) g_{lj}(v-x) dx}\right) dv.\end{aligned}$$

4 Percolation

Throughout the natural sciences, the study of percolation concerns behaviour relating to movement through some kind of structure. In graph theory, this manifests in the form of traversing the nodes of a graph using the edges that connect them. Traversing a graph in this way is one approach to understanding its component structure and connectivity. For example, if there exists a path (sequence of edges) that visits every vertex in the graph, then we can conclude that the graph has a single connected component.

Often, when we want to investigate the component structure of a random graph model, we consider the asymptotic component structure as the number of vertices grows to infinity. In this setting, a graph is said to percolate if the number of vertices in the largest component grows linearly with the number of vertices in the whole graph. In the model that we have studied, the number of vertices is infinite. Therefore, this model can be said to percolate if the largest component contains an infinite number of vertices.

When studying the general percolation properties of a model, we often want to characterise how changing the parameters within the model affects the percolation. Which model parameters cause the graph to percolate and which do not? This often gives rise to a phase transition boundary between a region in the parameter space where percolation can occur and the remaining region where it cannot. This transition boundary is often sufficient to describe the percolation characteristics of the model.

4.1 Geometric Random Intersection Graph

In this section we return to presenting the results from [2], before using them to investigate the percolation of the model defined in this paper. The approach that is used to study the percolation of the model is to choose a parameter to vary while all others are fixed. The geometric random intersection graph has three input parameters, the intensities of the Poisson processes generating the primary and auxiliary vertices, λ and μ, as well as the connection function, g. Of these three parameters, it is by far the most challenging to study how varying the connection function affects the percolation of the model. Therefore, the percolation properties are considered for some fixed function and only the Poisson intensities are varied.

Before we begin, we will introduce the following notation for the integral of g,

$$\|g\| := \int_{\mathbb{R}^d} g(x)dx,$$

since this integral is just the L_1-norm of g. Recall that we introduced the condition that $\|g\| < \infty$ when defining this model earlier in order to prevent the expected degree from being infinite. This is because the condition ensures that all primary vertices connect to a finite number of auxiliary vertices, and vice versa. If a function follows this condition, it is said to be integrable. As it

relates to this section, the condition also ensures that the percolation analysis is sensible by ruling out the possibility that one vertex could be solely responsible for creating an infinite component.

We already mentioned that the most reasonable approach is to fix the connection function g, which leaves a choice between the two Poisson intensity parameters, λ and μ. We will start by fixing λ and trying to identify the phase transition of μ. Let $C(\lambda, \mu, g)$ denote the number of vertices in the component which the zero vertex belongs to for a model with the given parameters. We are interested in the probability that the size of this component is infinite, $\mathbb{P}(C(\lambda, \mu, g) = \infty)$. We can then define

$$\mu_c(\lambda, g) = \sup\{\mu : \mathbb{P}(C(\lambda, \mu, g) = \infty) = 0\},$$

the largest μ such that there is no chance of percolation. The model is known to be ergodic, which implies that it percolates with probability 0 or 1 [5]. Therefore, $\mu_c(\lambda, g)$ defines a critical value describing the point of transition in μ where the model goes from not percolating to percolating almost surely. The transition is well defined because $\mathbb{P}(C(\lambda, \mu, g) = \infty)$ is increasing in μ.

An exact expression for the value of this transition point in terms of λ and g would be hard to find, so we will instead focus on the conditions required for the transition point to be non-trivial. By this, we mean that the transition does not occur at zero nor infinity, but rather somewhere in-between. If it were the case that $\mu_c(\lambda, g) = 0$, the model with parameters λ and g would always percolate regardless of the choice of μ whilst if $\mu_c(\lambda, g) = \infty$, the model would never percolate.

It turns out that this depends on whether the function g has bounded or unbounded support. When g has unbounded support, that means that points arbitrarily far away can still connect with some probability. The consequence of this is that $\mu_c(\lambda, g) \in (0, \infty)$ for all $\lambda > 0$.

However, when g has bounded support, this is not always the case. Here, primary and auxiliary vertices can only connect if the distance between them is within the support of g. As a result, sufficiently small values of λ will prevent percolation despite the size of μ.

We can observe this through the Poisson Boolean model. Recall that the Poisson Boolean model connects points from a Poisson process directly if and only if they are within a fixed distance of one another. It has been shown that, for some fixed radius r, the Poisson Boolean model has a non-trivial phase transition $\tilde{\lambda}_c(r) \in (0, \infty)$ [5]. That is that the model percolates when the Poisson process has intensity $\lambda > \tilde{\lambda}_c(r)$ and does not if $\lambda < \tilde{\lambda}_c(r)$.

Now, because g has bounded support, we can define $s_{\max} = \sup\{|x| : g(x) > 0\}$ as the furthest distance between primary and auxiliary vertices such that there is still positive probability that they connect. Beyond this value, we can be certain that points have no probability of connecting. Therefore, $2s_{\max}$ is the furthest distance that two primary vertices can be from each other and still be connected in the geometric random intersection graph. To see this, consider two primary vertices v_1 and v_2. If they are connected, then there exists at least one

auxiliary vertex u that both connect to in the initial bipartite graph. Because g is bounded, the furthest v_1 could be from u is $s_{\max}$, by definition, and likewise for v_2. If both v_1 and v_2 are within $s_{\max}$ of u, then v_1 and v_2 are within $2s_{\max}$ of each other.

Note that the random intersection graph with bounded g is a subgraph of the Poisson Boolean model with radius $2s_{\max}$ as two vertices in the intersection graph have to be within distance $2s_{\max}$ of each other to have even a chance of connecting while the Poisson Boolean model with this radius connects all such points. Thus, if there is not an infinite component in this Poisson Boolean model, then it is impossible to have an infinite component in the random intersection graph with bounded g. Since $\tilde{\lambda}_c(r) \in (0, \infty)$, there are some values of λ where percolation cannot occur in the Poisson Boolean model, specifically when $\lambda < \tilde{\lambda}_c(r)$. Therefore, when $\lambda < \tilde{\lambda}_c(2s_{\max})$, the model never percolates irrespective of μ, so $\mu_c(\lambda, g) = \infty$.

A further interesting property of this model with regards to percolation is the symmetry that exists between the parameters λ and μ, meaning that if we know whether $\lambda = a$ and $\mu = b$ percolates, then the model with $\lambda = b$ and $\mu = a$ exhibits the same percolation characteristics.

To see this, consider the bipartite graph constructed between primary vertices in $\mathcal{V}$ and auxiliary vertices in $\mathcal{U}$. Usually, we construct the random intersection graph by connecting vertices from $\mathcal{V}$ if they both connect to at least one mutual vertex from $\mathcal{U}$ in the bipartite graph. If, instead, we were to reverse the roles of $\mathcal{V}$ and $\mathcal{U}$ and connect vertices from $\mathcal{U}$ to each other if they have a mutual connection with at least one vertex from $\mathcal{V}$ and show that this model also percolates, then we would arrive at this symmetry property.

First, if we have an infinite component in the geometric random intersection graph of vertices from $\mathcal{V}$, then we also have an infinite component in the bipartite graph. This is because all of the vertices from $\mathcal{V}$ that are connected to each other in this infinite component are connected by paths of length two in the bipartite graph through vertices from $\mathcal{U}$. Furthermore, when $\|g\| < \infty$, vertices from $\mathcal{V}$ and $\mathcal{U}$ have finite degree in the bipartite graph, which means that an infinite component in the bipartite graph must contain infinite numbers of vertices from both $\mathcal{V}$ and $\mathcal{U}$. This means that if we were to construct the intersection graph with vertices from $\mathcal{U}$, there would be an infinite component in this graph as well.

To define the corresponding percolation transition parameter for the intersection graph of vertices from $\mathcal{U}$, we would fix μ and have

$$\lambda_c(\mu, g) = \sup\{\lambda : \mathbb{P}\left(C(\mu, \lambda, g) = \infty\right) = 0\},$$

The symmetry of the percolation parameters implies that, for any a, we have that $\lambda_c(a, g) = \mu_c(a, g)$. This gives us an understanding of the transition boundary in terms of both parameters.

4.2 The Multi-Type Model

In this section, we will attempt to characterise the percolation of the geometric random intersection graph with multiple vertex sets, which was defined earlier.

4.2.1 Two Primary Vertex Sets

We will begin, as we did when exploring the core properties of the model, with the simplest case where there are two primary vertex sets, $\mathcal{V}_1$ and $\mathcal{V}_2$ along with only one auxiliary vertex set, $\mathcal{U}$ and two integrable connection functions, g_1 and g_2. In this model, we have three intensity parameters to investigate, λ_1, λ_2 and μ. The obvious choice is to fix λ_1 and λ_2 and observe the percolation properties when varying μ.

As such, we can define

$$\mu_c(\lambda_1, \lambda_2, g_1, g_2) = \sup\{\mu : \mathbb{P}\left(C(\lambda_1, \lambda_2, \mu, g_1, g_2) = \infty\right) = 0\}$$

in an analogous way to baseline model. We saw that in that case, whether or not the connection function was bounded had an effect on when $\mu_c(\lambda_1, \lambda_2, g_1, g_2)$ was non-trivial. Here, this is potentially complicated by the presence of two connection functions. Therefore, we have four cases to think about:

1. g_1 and g_2 have unbounded support,
2. g_1 has unbounded support while g_2 has bounded support,
3. g_1 has bounded support while g_2 has unbounded support,
4. g_1 and g_2 have bounded support.

First, let us consider the case where at least one of the connection functions has unbounded support. Let us assume that it is g_1 that has unbounded support. Consider the subgraph containing only vertices from $\mathcal{V}_1$ and the connections between them. Since this is a random intersection graph with only one primary vertex set, we know that $\mu_c(\lambda_1, g_1) \in (0, \infty)$. In other words, the phase transition is non-trivial.

Next, note that

$$\mu_c(\lambda_1, \lambda_2, g_1, g_2) \leq \mu_c(\lambda_1, g_1).$$

This is because adding vertices through the second vertex set $\mathcal{V}_2$ would only increase the probability of an infinite component and therefore reduce the number of auxiliary vertices needed to achieve percolation. Since $\mu_c(\lambda_1, g_1) \in (0, \infty)$ and bounds $\mu_c(\lambda_1, \lambda_2, g_1, g_2)$ from above, we can conclude that $\mu_c(\lambda_1, \lambda_2, g_1, g_2) < \infty$ for all values of λ_1 and λ_2.

As an aside, an interesting follow-up question which we will not attempt to answer here is whether the inequality between the transition parameters of the subgraph and the whole model is a strict inequality or not. In other words,

are there values of λ_2 small enough that the vertices from $\mathcal{V}_2$ do not affect the probability of percolation in the wider model?

Returning to the investigation of the effects of defining one of the connection functions to have unbounded support, it remains to determine whether $\mu_c(\lambda_1, \lambda_2, g_1, g_2)$ is non-zero. Define a geometric random intersection graph with a single primary vertex set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ with intensity $\lambda_1 + \lambda_2$ and connection function $\tilde{g} = \max\{g_1, g_2\}$. Defining $\tilde{g}$ in this way gives a valid connection function since we still have that $\|\tilde{g}\| < \infty$. To see this, note that given $\|g_1\| < \infty$ and $\|g_2\| < \infty$, which we require by definition, then $\|g_1\| + \|g_2\| < \infty$. We can quickly see that

$$\begin{aligned}
\|g_1\| + \|g_2\| &= \int_{\mathbb{R}^d} g_1(x)dx + \int_{\mathbb{R}^d} g_2(x)dx \\
&= \int_{\mathbb{R}^d} \left(g_1(x) + g_2(x)\right) dx \\
&\geq \int_{\mathbb{R}^d} \max\{g_1(x), g_2(x)\}dx,
\end{aligned}$$

and so $\|\tilde{g}\| < \infty$.

This graph has the same vertex set, but each pair of vertices has a higher or equal probability of being connected since $\tilde{g} \geq g_1$ and $\tilde{g} \geq g_2$ everywhere. Because this graph is more connected than our model with two vertex types, it is more likely to have an infinite component. Therefore,

$$\mu_c(\lambda_1, \lambda_2, g_1, g_2) \geq \mu_c(\lambda_1 + \lambda_2, \tilde{g}).$$

Since the model with connection function $\tilde{g}$ is a geometric random intersection graph with only one primary vertex set, we know that $\mu_c(\lambda_1+\lambda_2, \tilde{g}) \in (0, \infty)$ and so it follows that $\mu_c(\lambda_1, \lambda_2, g_1, g_2) > 0$. Thus we have shown that the graph with two vertex sets also has a non-trivial phase transition. Note that this reasoning required no restrictions on g_2 so it could have had bounded or unbounded support. Therefore it is sufficient for one connection function to be unbounded to have a non-trivial transition value. This addresses the first 3 out of 4 cases that we set out to explore.

The final case is when both g_1 and g_2 have bounded support. With just a single primary vertex set, we saw that a connection function with bounded support combined with a suitably small λ ruled out percolation for all values of μ by comparing with the Poisson Boolean model. Now with two functions with bounded support, we will make a similar argument. Let $s_{1,\max} = \sup\{|x| : g_1(x) > 0\}$ and $s_{2,\max} = \sup\{|x| : g_2(x) > 0\}$. Then, define $s_{\max} = \max\{s_{1,\max}, s_{2,\max}\}$, being the furthest distance from either function between primary and auxiliary vertices such that there is still positive probability that they connect. Then, this model is contained within the Poisson Boolean model with radius $2s_{\max}$. Given that the vertex set $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ has intensity $\lambda_1 + \lambda_2$, we can say that $\mu_c(\lambda_1, \lambda_2, g_1, g_2) = \infty$ for $\lambda_1 + \lambda_2 < \tilde{\lambda}_c(2s_{\max})$.

Furthermore, there will exist sufficiently large values of λ_1 and λ_2 such that $\mu_c(\lambda_1, \lambda_2, g_1, g_2) \in (0, \infty)$. In fact, it should only be necessary for one of λ_1 or λ_2 to be large. For example, if at least one of λ_1 or λ_2 in combination with its connection function gives a non-trivial value of $\mu_c(\lambda, g)$ in the single vertex set case, then $\mu_c(\lambda_1, \lambda_2, g_1, g_2) \in (0, \infty)$ will also be true.

We have shown that the condition $\lambda_1 + \lambda_2 < \tilde{\lambda}_c(2s_{\max})$ is sufficient to make the conclusion that the model will not percolate for any value of μ. However, we did not show whether $\lambda_1 + \lambda_2 = \tilde{\lambda}_c(2s_{\max})$ is the point at which the critical value goes from being infinite to being finite. We will not attempt to answer this but it is still an interesting open problem relating to this particular model.

4.2.2 Multiple Primary Vertex Sets

Before we move on to the next simplest version of the model where there are two auxiliary vertex sets, let us first note that the properties we have just found for the two primary vertex set model actually extend for any number of primary vertex sets, as long as there is only one auxiliary vertex set. We will now demonstrate this by extending the same arguments that were just presented.

Let n be the number of primary vertex sets. Therefore, there are n primary intensity parameters, one auxiliary intensity parameter and n connection functions. Define $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_n]$, and $\boldsymbol{g} = [g_1, \ldots, g_n]$ to tidy up some of the notation. Furthermore, define $\mu_c(\boldsymbol{\lambda}, \boldsymbol{g})$ to be the largest value of μ such that the probability of an infinite component is zero.

Now, assume that one of the connection functions, for example g_1, has unbounded support. Using the same logic as before, the subgraph containing only vertices from $\mathcal{V}_1$ has a phase transition that is known to be non-trivial and greater that the phase transition of the model with n vertex sets. This allows us to conclude that $\mu_c(\boldsymbol{\lambda}, \boldsymbol{g}) < \infty$.

We can also say that the phase transition is non-zero since we can define a model with primary vertex set $\mathcal{V} = \bigcup_{i=1}^{n} \mathcal{V}_i$ with intensity $\sum_{i=1}^{n} \lambda_i$ and connection function $\max\{g_1, \ldots, g_n\}$. Since this model has a non-trivial phase transition that is less than the one for the multiple primary vertex set model, we can say that $\mu_c(\boldsymbol{\lambda}, \boldsymbol{g}) > 0$.

When all connection functions have bounded support, the model can be bounded by a Poisson Boolean model with radius equal to the largest value in the supports of any of the connection functions $g_1, \ldots, g_n$. If we define this value to be $s_{\max}$ and define $\tilde{\lambda}_c(s_{\max})$ to be the non-trival phase transition of the Poisson Boolean model, as before, then we can conclude that $\mu_c(\boldsymbol{\lambda}, \boldsymbol{g}) = \infty$ when $\sum_{i=1}^{n} \lambda_i < \tilde{\lambda}_c(s_{\max})$. These results are summarised below.

Result: Percolation for n Primary Vertex Sets

- If at least one connection function has unbounded support, then $\mu_c(\boldsymbol{\lambda}, \boldsymbol{g}) \in (0, \infty)$ for all $\boldsymbol{g}, \boldsymbol{\lambda}$.
- If all connection functions have bounded support, then $\mu_c(\boldsymbol{\lambda}, \boldsymbol{g}) = \infty$ when $\sum_{i=1}^{n} \lambda_i < \tilde{\lambda}_c(s_{\max})$.

4.2.3 Symmetry for Two Auxiliary Vertex Sets

Now we turn to the model with one primary vertex set, $\mathcal{V}$, and two auxiliary vertex sets, $\mathcal{U}_1$ and $\mathcal{U}_2$. The natural parameter to investigate here is λ, so we will define

$$\lambda_c(\mu_1, \mu_2, g_1, g_2) = \sup\{\lambda : \mathbb{P}\left(C(\lambda, \mu_1, \mu_2, g_1, g_2) = \infty\right) = 0\}$$

to be the phase transition for some fixed μ_1, μ_2, g_1 and g_2. We again want to determine the conditions on the fixed parameters that induce a non-trivial transition value. We could break this problem down into the four scenarios of bounded and unbounded supports that we did with the previous case, but it is worth looking into whether the percolation in the model with two primary vertex sets says anything about this model with two auxiliary vertex sets. The symmetry between parameters in the basic geometric random intersection graph hints that we may be able to obtain the results we are looking for in a more efficient way whilst concluding something more generally about the relationship between the two primary vertex set model and the two auxiliary vertex set model.

Consider a model consisting of primary vertex sets, $\mathcal{V}_1$ and $\mathcal{V}_2$, with intensities $\lambda_1 = a_1$ and $\lambda_2 = a_2$, and an auxiliary vertex set, $\mathcal{U}$, with intensity $\mu = b$. Our hypothesis is that if we switch the roles of the vertex sets, taking $\mathcal{U}$ to be the primary vertex set and $\mathcal{V}_1$ and $\mathcal{V}_2$ to be the auxiliary vertex sets, then percolation in this two auxiliary vertex set model with the vertex set role reversal is equivalent to percolation in the initial two primary vertex set model, provided that the connection functions, g_1 and g_2, are kept the same.

Our initial assumption is that we know whether the two primary vertex set model percolates. In the standard random intersection model, we reasoned that percolation in the resulting intersection graph containing only primary vertices is equivalent to percolation in the bipartite graph between primary and auxiliary vertices. This was because vertices that connect in the intersection graph are connected by a path of length two in the bipartite graph. When we have two primary vertex sets, this argument remains unchanged.

Next, recall that we construct the bipartite graph by connecting all $v_1 \in \mathcal{V}_1$ and $u \in \mathcal{U}$ with probability $g_1(v_1 - u)$ and connecting all $v_2 \in \mathcal{V}_2$ and $u \in \mathcal{U}$ with probability $g_2(v_2 - u)$. Now, if we reverse the roles of vertex sets and consider $\mathcal{U}$ to be the primary vertex set with auxiliary vertex sets $\mathcal{V}_1$ and $\mathcal{V}_2$, we would construct the appropriate bipartite graph by connecting all $u \in \mathcal{U}$ and $v_1 \in \mathcal{V}_1$ with probability $g_1(u - v_1)$ and connecting all $u \in \mathcal{U}$ and $v_2 \in \mathcal{V}_2$ with probability $g_2(u - v_2)$. Since g_1 and g_2 are radial, $g_1(v_1 - u) = g_1(u - v_1)$ and $g_2(v_2 - u) = g_2(u - v_2)$. The consequence of this is that the bipartite graph is the same for both the two primary vertex model and the two auxiliary vertex model, provided that the parameters are mirrored as we defined them and the connection functions are the same.

The final step is to argue that percolation in this bipartite graph gives percolation in the mirrored random intersection graph with two auxiliary vertex sets. This argument follows the same form as the single vertex set model. If we have that $\|g_1\| < \infty$ and $\|g_2\| < \infty$, then each vertex has a finite number of connections in the bipartite graph, meaning that an infinite component in the bipartite graph must contain an infinite amount of both primary and auxiliary vertices. Hence, when we construct the intersection graph using the vertices from $\mathcal{U}$ as primary vertices, those vertices that belong to an infinite component in the bipartite graph will belong to the same component in the intersection graph, and since there are infinitely many, the intersection graph percolates. Therefore, percolation in the two primary vertex set model with parameters $\lambda_1 = a_1$, $\lambda_2 = a_2$, $\mu = b$ is equivalent to percolation in the two auxiliary vertex set model with parameters $\lambda = b$, $\mu_1 = a_1$, $\lambda_2 = a_2$.

This also means that

$$\lambda_c(a_1, a_2, g_1, g_2) = \mu_c(a_1, a_2, g_1, g_2)$$

for constants a_1 and a_2 and thus we can extend the properties that we proved for $\mu_c(\lambda_1, \lambda_2, g_1, g_2)$. These being that when one of g_1 or g_2 has unbounded support, $\lambda_c(\mu_1, \mu_2, g_1, g_2) \in (0, \infty)$ for all μ_1 and μ_2 while when both g_1 and g_2 have bounded support, $\lambda_c(\mu_1, \mu_2, g_1, g_2) = \infty$ for $\mu_1 + \mu_2 < \tilde{\mu}_c(2s_{\max})$, where $s_{\max}$ is the furthest distance from either function between primary and auxiliary vertices such that there is still positive probability that they connect.

4.2.4 Symmetry for Multiple Auxiliary Vertex Sets

This property also extends to the case where we have an arbitrary number of primary vertex sets, through the same arguments as before. To summarise, if know whether the model with n primary vertex sets has an infinite component, then the vertices in this component are all part of the same component in the bipartite graph, and so the bipartite graph percolates. When we reverse the roles of the vertex sets to have $\mathcal{U}$ as the primary vertex set with n auxiliary vertex sets, the bipartite graph is the same because the all connection functions are radial. Finally, if all connection functions are integrable, then each vertex from $\mathcal{U}$ has a finite number of connections in the bipartite graph. Therefore, if there is an infinite component in the bipartite graph, all of these vertices must be in the same component in the intersection graph and thus the model with n auxiliary vertex sets percolates.

Result: Percolation for n Auxiliary Vertex Sets

Percolation in the graph with n primary vertex sets and one auxiliary vertex set implies percolation in the graph with one primary vertex set and n auxiliary vertex sets when the roles of the parameters are swapped.

4.2.5 Combining Multiple Primary and Auxiliary Vertex Sets

In these models where either the primary or auxiliary vertex set has no subdivision into types, we still have a clear candidate for a parameter to focus on when studying the percolation boundary. This choice of parameter becomes less clear once we have more than one type for both primary and auxiliary vertices. To illustrate this, we will look at the model with two primary vertex sets and two auxiliary vertex sets. Recall that this model included four intensity parameters, λ_1, λ_2, μ_1 and μ_2 as well as four connection functions g_{11}, g_{12}, g_{21} and g_{22}.

The first approach we might consider is to fix all parameters apart from one, as we have done up to this point. For example, let us pick μ_1 as our chosen parameter and fix the other three. As we vary μ_1, we will capture nothing about the interactions between auxiliary vertices from $\mathcal{U}_2$ and the primary vertex sets. Therefore, this approach does not capture the full dynamics of the model. This was not a problem we faced with previous models as there was always at least one parameter that could be varied while affecting the behaviour of the entire model and ensuring that all connection functions played a role in the phase transition. In the model with two primary vertex sets and one auxiliary vertex set, this was the parameter μ.

We will still attempt to characterise the percolation using this approach in order to demonstrate why it does not produce useful results. Define $\boldsymbol{\lambda} = [\lambda_1, \lambda_2]$, $\boldsymbol{\mu} = [\mu_1, \mu_2]$ and $\boldsymbol{G} = [g_{11}, g_{12}, g_{21}, g_{22}]$, once again for notational reasons. Here, we would define our transition parameter as

$$\mu_c(\boldsymbol{\lambda}, \mu_2, \boldsymbol{G}) = \sup\{\mu_1 : \mathbb{P}\left(C(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{G}) = \infty\right) = 0\}.$$

With the previous models, for every combination of the fixed parameters, we wanted to say whether the transition was non-trivial. However, since μ_2 is a fixed parameter along with λ_1 and λ_2, this method will no longer provide the properties that we are interested in. For example, fix λ_1, λ_2 and μ_2 such that the sub-model containing the corresponding vertex sets $\mathcal{V}_1$, $\mathcal{V}_2$ and $\mathcal{U}_2$ percolates. This means that the wider model that includes $\mathcal{U}_1$ must also percolate with an equal or higher probability, regardless of the value of μ_1, since adding more auxiliary vertices can only make the model more connected. Therefore, $\mu_c(\boldsymbol{\lambda}, \mu_2, \boldsymbol{G}) = 0$, before we even had a chance to consider the effects of μ_1. The problem here is that any choice of individual parameter leaves a subgraph which it does not affect, making the transition parameter highly dependent on the connectivity of this subgraph.

In an attempt to address this, another approach we might try is to investigate varying the sum of parameters, either from the primary or auxiliary vertex sets. For example, we could fix both λ_1 and λ_2 and investigate how changing the sum $\mu_1 + \mu_2$ affects the percolation of the model. In this case, we would define the percolation transition as

$$\mu_c(\boldsymbol{\lambda}, \boldsymbol{G}) = \sup\{\mu_1 + \mu_2 : \mathbb{P}\left(C(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{G}) = \infty\right) = 0\}.$$

The expansion of potential number of cases to investigate involving different

combinations of functions with bounded and unbounded support adds further complication.

One case that is not affected too much by these new complications is the one where we take all four connection functions to have bounded support. Here, we can bound this model using a Poisson Boolean model with radius equal to two times the maximum value in the combined support of the connection functions. Therefore, for a sufficiently small sum $\lambda_1 + \lambda_2$, the model cannot percolate no matter how large we make μ_1 and μ_2. This remains in-line with the results for the previous models where all connection functions were assumed to have bounded support.

The main task is to consider the effects of having functions with unbounded support. Define a single type geometric random intersection graph with intensities $\lambda_1+\lambda_2$ and $\mu_1+\mu_2$ and connection function $\tilde{g} = \max\{g_{11}, g_{12}, g_{21}, g_{22}\}$. This function $\tilde{g}$ has unbounded support as long as at least one of the four connection functions has unbounded support. Therefore, this single type model has a non-trivial phase transition. Moreover, since this model is more connected than the two-by-two model we are interested in, we get that $\mu_c(\boldsymbol{\lambda}, \boldsymbol{G}) > 0$ in the two-by-two model when at least one function has unbounded support.

Next, we want to check whether we can get an infinite critical value while at least one function has unbounded support. If it is the case that $\mu_c(\boldsymbol{\lambda}, \boldsymbol{G}) = \infty$, then we must be able to take either μ_1 or μ_2 to be arbitrarily large and have that the model still does not percolate. Assume that the model does not percolate for any value of μ_1 and consider the sub-model constructed using only the vertices from $\mathcal{V}_1$ and $\mathcal{U}_1$. If g_{11} has unbounded support, then we know from previous results that the critical value of this sub-model is finite. Furthermore, since the wider model contains this sub-model, it is more likely to contain an infinite component. Therefore, when μ_1 is greater than the phase transition of the sub-model containing vertices from $\mathcal{V}_1$ and $\mathcal{U}_1$, the model percolates. Therefore, μ_1 cannot cause the transition parameter, $\mu_c(\boldsymbol{\lambda}, \boldsymbol{G})$, to be infinite when g_{11} has unbounded support. This argument can be replicated when g_{21} has unbounded support.

However, it could also be the case that the model does not percolate for any value of μ_2. Applying the same arguments from the previous paragraph, we can rule this out when one of g_{12} or g_{22} has unbounded support. Therefore, if one of g_{11} or g_{21} has unbounded support and one of g_{12} or g_{22} has unbounded support, then the model has a non-trivial phase transition, $\mu_c(\boldsymbol{\lambda}, \boldsymbol{G})$, when defined in terms of the sum of auxiliary intensities.

Where this definition of $\mu_c(\boldsymbol{\lambda}, \boldsymbol{G})$ starts to break down is when we have a combination of functions with bounded and unbounded support and one of the two auxiliary vertex sets only interacts with connection functions with bounded support. This scenario might look something like the following. The functions g_{11} and g_{21}, which connect vertices from $\mathcal{U}_1$ to those from $\mathcal{V}_1$ and $\mathcal{V}_2$ respectively, have unbounded support, but the functions g_{12} and g_{22}, which likewise only interact with auxiliary vertices from $\mathcal{U}_2$, have bounded support.

Here, the sub-model constructed from the vertex sets $\mathcal{U}_2$, $\mathcal{V}_1$ and $\mathcal{V}_2$ is a two primary vertex set model where both connection functions have bounded

support. As such, we know that for a sufficiently small sum of primary vertex intensities, $\lambda_1 + \lambda_2$, the model does not percolate for any μ_2. We showed this earlier by comparing the model to a Poisson Boolean model. This means that to get $\mu_c(\boldsymbol{\lambda}, \boldsymbol{G}) = \infty$, we should be able to take μ_2 to be arbitrarily large with the sufficiently small values of λ_1 and λ_2 along with a small value of μ_1 that does not give percolation in the subgraph of vertex sets $\mathcal{U}_1$, $\mathcal{V}_1$ and $\mathcal{V}_2$. We know that such a value of μ_1 exists since the phase transition is non-trivial for models with a single auxiliary vertex set and at least one connection function with unbounded support. This means that it seems to be the case that there are certain conditions on λ_1 and λ_2 in this setup that lead to a trivial phase transition, though this has not been proven.

We are not going to look into the specifics of how these sub-models interacting affects the overall percolation and the exact conditions on the sub-model containing the functions with unbounded support needed to give an overall model with no percolation for some infinite large sum of auxiliary vertex intensities. This exploration of the two primary and two auxiliary vertex set model is meant to demonstrate why characterising the phase transition of the model for increasingly many vertex sets becomes less meaningful.

5 Conclusion

In this paper, we have expanded upon the known theory within the field of random graphs by extending the geometric random intersection graph and characterising the core properties of the extended model. The addition of multiple primary and auxiliary vertex sets to the geometric random intersection graph, each with unique general connection functions, provides a further layer of flexibility in modelling real-world networks while still maintaining workable properties that largely resemble those of the original model.

The connection probability and expected degree give an overview of how the edges and vertices behave within the model. In the original geometric random intersection graph, the connection probability decays exponentially with an increase in distance between the vertices in question or a reduction in μ, the intensity of the Poisson process generating auxiliary vertices. In our model, this exponential feature of the connection probability remains. Furthermore, we can view the probability that two vertices do not connect in the model with m auxiliary vertex sets as the product of the probabilities that they do not connect in each of the sub-models containing just one of the auxiliary vertex sets. Similarly, we found that the expected degree of our model is just the sum of the expected degrees of the sub-models including just one of the primary vertex sets. These properties extend nicely in this way due to the independence imposed when generating the vertex sets and connecting vertices.

Furthermore we were able to characterise certain properties relating to the percolation of the model with an arbitrary number of primary vertex sets and just one auxiliary vertex set. We distinguished the difference between having connection functions with bounded and unbounded support and showed that having at least one function with unbounded support is sufficient to guarantee a non-trivial phase transition. For the remaining case, where all functions have bounded support, we determined a condition on the sum of primary vertex intensity parameters that ensures the model will not percolate for any auxiliary vertex intensity parameter. This turned out be when this sum of primary intensity parameters was less than the phase transition of a Poisson Boolean model with radius equal to the largest value in the combined support of the two connection functions.

In our final result, we showed that the percolation properties from the model with n primary vertex sets can be mirrored to the model with n auxiliary vertex sets. This means that the percolation results we found for the model with n primary vertex sets can be directly applied to the model with n auxiliary vertex sets without any further investigation. This also means that any future results relating to the percolation of one these models may be applied to the other.

By building upon the theory of geometric random intersection graphs, we move ever closer to being able to model the diverse and complex structures of real world networks.

6 Future Work

There are still many opportunities relating to the work that has been done here to expand upon our knowledge of random graph structures and construct new models that follow the complex structures of real-world networks. Among these are a number of avenues of inquiry relating to the exact model that we have defined. We briefly mentioned two open problems relating to the percolation of the two primary vertex set model. In the first of these open problems, we were considering the case that one connection function has unbounded support and we speculated whether there could be any non-zero values of the opposite primary vertex set intensity parameter that do not affect the location of the phase transition. The other problem was set within the scenario where both connection functions had bounded support. We could ask whether we can improve upon the specificity of the bound set on the sum of primary vertex set intensities to guarantee a model that does not percolate for any auxiliary Poisson process.

When attempting to find the percolation properties of the generalised model, we ran into some practical problems caused by having a large number of parameters to deal with. Future work on this model might include attempting to overcome some of these challenges with a different approach to studying percolation than the one used in this paper.

It may also be useful or interesting to define new models that are similar or build upon the one defined here. For example, while the Poisson process is useful for its simplicity and homogeneity, it may not appropriately represent the way in which nodes are spread across a space. One could therefore consider constructing the same model using different point processes.

There has been some interest in the field surrounding inhomogeneous random graphs where each vertex is assigned a weight and the connection probabilities are a function of these weights. Applying this approach to the geometric random intersection graph is an unexplored alternative to defining multiple vertex sets as we have done.

One could also adapt the intersection step so that not all primary vertices who share the same auxiliary vertex are connected to each other automatically. Instead, it may make sense in certain situations to have some probability that each pair of primary vertices that share an auxiliary vertex are connected. In particular, this makes more sense when the average number of primary vertices connected to a single auxiliary vertex is quite high.

A further construction that might improve the ability of the model to handle complex network dynamics is to add an extra layer of intersection in a hierarchical way. This means that there would be two or more levels of auxiliary vertices, with primary vertices first connecting to lower level auxiliary vertices and then these lower level auxiliary vertices themselves connecting to higher level auxiliary vertices. This would mean that primary vertices would be connected if they shared a higher level auxiliary vertex through their connections to lower level auxiliary vertices. Moreover, this construction could be combined with the addition of probabilities for connecting in the intersection step, particularly for

primary vertices connecting to a higher level auxiliary vertex through different lower level vertices.

One final suggestion for future work is to explore different ways of constructing geometric random intersection graphs and even combining these methods with the model defined here. Consider studying a social network where we consider how children form connections through both school and social groups outside of school. We could say that modelling the extra-curricular activities using a geometric random intersection graph might be quite natural but modelling connections with schools may not fit so well. This is because we may want to enforce the rule that children attend no more than one school and that schools have a limit on their capacity. Therefore, we might want to connect children to schools using a configuration model, where the degree of each vertex is defined first from some distribution and then vertices are paired until each vertex has reached its specified degree. It would be interesting to study a version of the geometric random intersection graph where primary and auxiliary vertices are connected in this way and whether a mixed model could be defined such that there are multiple auxiliary vertex sets where some sets connect using the procedure investigated in this paper and some use a configuration approach.

References

[1] Jean-François Coeurjolly, Jesper Møller, and Rasmus Waagepetersen. "A tutorial on Palm distributions for spatial point processes". In: *International Statistical Review* 85.3 (2017), pp. 404–420.

[2] Maria Deijfen and Riccardo Michielan. "Geometric random intersection graphs with general connection probabilities". In: *Journal of Applied Probability* (2023).

[3] P. Erdős and A.Rényi. "On the Evolution of Random Graphs". In: *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* 5 (1960), pp. 17–61.

[4] E. N. Gilbert. "Random plane networks". In: *Journal of the Society for Industrial and Applied Mathematics* (1961).

[5] Ronald Meester and Rahul Roy. *Continuum Percolation*. Cambridge Tracts in Mathematics. Cambridge University Press, 1996.

[6] J. Moller and R.P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. CRC Press, 2003, p. 14. ISBN: 9780203496930.

[7] S. Resnick. *Adventures in Stochastic Processes*. Birkhäuser Boston, MA, 1992. Chap. 4.

[8] Karen B. Singer. "Random Intersection Graphs". PhD thesis. Johns Hopkins University, 1995.