

# 1

## GRAPH THEORY

---

A review of graph theory lays the foundation for the mathematical considerations in this work.

In this chapter we review basic graph theory and explain how these terms are applicable in the context of biological neural networks. We begin with the definition of directed graphs:

### 1.1 DIRECTED GRAPHS

**Definition 1.1** (Directed graphs). A *directed pseudograph*  $G$  consists of two finite  $V$ , the *set of vertices* of  $G$ , and  $E$ , the *set of edges* of  $G$ , and two maps  $s, t : E \rightarrow V$ , the *source* and *target functions* of  $G$ . A *directed multigraph* is a directed pseudograph without loops, that is the map  $d = (s, t) : E \rightarrow V^2$  already maps to  $V^2 \setminus \Delta_V$ , where  $V^2 = V \times V$  denotes the cartesian product and  $\Delta_V = \{(x, x) \mid x \in V\} \subseteq V^2$  the diagonal. Similarly, a *directed loop graph* is a directed pseudograph where  $d$  is injective. Finally, a *simple directed graph* can be defined as a directed pseudograph where  $d$  is both injective and already maps to  $V^2 \setminus \Delta_V$ .

Thus, in simple directed graphs, neither parallel edges nor loops - edges between the same vertex - are allowed, whereas directed multigraphs and directed loop graphs admit one of them respectively.

Say something about what "directed graph" means here, do all definition until random graphs work for all types of directed graphs?

Given a directed graph  $G$ , we denote with  $V(G)$  the set of vertices of  $G$  and call it the **vertex set** of  $G$ . Analogously, the **edge set**  $E(G)$  of  $G$  denotes the set of edges of  $G$ . This means, for a directed graph specified as  $G = (V, E, s, t)$ , we have  $V(G) = V$  and  $E(G) = E$ .



**Figure 1.1:** Typical examples of the directed graph types **A**) directed pseudograph **B**) directed multigraph **C**) directed loop graph **D**) simple directed graph.

A **morphism**  $\phi : G \rightarrow H$ , between two directed graphs  $G = (V_G, E_G, s_G, t_G)$  and  $H = (V_H, E_H, s_H, t_H)$ , consists of a pair of maps  $\phi_V : V_G \rightarrow V_H$  and  $\phi_E : E_G \rightarrow E_H$ , such that

$$s_H \circ \phi_E = \phi_V \circ s_G \quad \text{and} \quad t_H \circ \phi_E = \phi_V \circ t_G,$$

that is such that the following diagram commutes:

$$\begin{array}{ccc} E_G & \xrightarrow{\phi_E} & E_H \\ s_G \downarrow t_G & & \downarrow s_H t_H \\ V_G & \xrightarrow{\phi_V} & V_H \end{array}$$

A morphism  $\varphi : G \rightarrow H$ , between two directed pseudographs  $G$  and  $H$  is an **isomorphism**, if the maps  $\varphi_V : V_G \rightarrow V_H$  and  $\varphi_E : E_G \rightarrow E_H$  are bijective. Two directed pseudographs are called *isomorphic* if there exists an isomorphism inbetween them.

*Remark.* The last definition implies that, if there exists an isomorphism  $\varphi : G \rightarrow H$ , an isomorphism  $\psi : H \rightarrow G$  can be found. This isomorphism is, of course, easily constructed via  $\psi_V : V_H \rightarrow V_G, v \mapsto \varphi_V^{-1}(v)$ ,  $\psi_E : E_H \rightarrow E_G, e \mapsto \varphi_E^{-1}(e)$ .

**Definition 1.2** (Weighted directed graphs). An *edge-weighted directed graph* is a directed graph  $G$  along with a mapping  $\omega : E(G) \rightarrow \mathbb{R}$ , called the *weight function*. Similarly, a *vertex-weighted directed graph* is a directed graph with a mapping  $\nu : V(G) \rightarrow \mathbb{R}$ .

*Remark.* (heavy draft)

Certainly, a weighted directed pseudograph is the most fitting mathematical modelling to the biological situation, since self connections and multiple synapse are not only plausible (source?) but the rule. However there is one abstraction we can make by adding together the synaptic weights - NEST is doing it. While linear synaptic integration was shown to be, it is a prevalent model in network (ref Nest).

Also think about **inhibitory, excitatory**. Suggestion: edge weights  $\omega : E(G) \rightarrow \mathbb{R}^+$  and vertex weights  $\nu : V(G) \rightarrow \{-1, 1\}$ . Synaptic weight  $\text{syn}(e)$  for edge  $e$  is then

$$\text{syn}(e) = \nu(s(e)) \omega(e).$$

Benefit: Synapse from one neuron are either excitatory or inhibitory but not mixed as in bio.

*Remark* (Equivalent definiton for directed loop graphs). A directed loop graph  $G$  can be equivalently defined as a pair of finite(, non-empty?) sets  $V$ , the *set of vertices* of  $G$ , and  $E \subseteq V^2$  the *set of edges* of  $G$ . For an edge  $(x, y) \in E$ , we call  $x$  the *source* and  $y$  the target of the edge  $(x, y)$ .

Source and target functions are then uniquely determined as the projections on the first and second component,  $s = \text{pr}_1, t = \text{pr}_2 : E(G) \rightarrow V$ . Conversely, the edge set  $E(G) \subseteq V^2$  can be determined from the source and target functions as  $E := \{(s(e), t(e)) \mid e \in E\}$ . The trivial identities  $(x, y) = (\text{pr}_1(x, y), \text{pr}_2(x, y))$  and  $\text{pr}_1(s(e), t(e)) = s(e)$  with  $\text{pr}_2(s(e), t(e)) = t(e)$  quickly verify the equivalence of the definitions.

Given a directed loop graph  $G$ , we often assume the graph to be given in this form and write edges as  $e = (x, y)$ . Note that this concept is more complicated to introduce for directed pseudographs, since parallel edges  $e$  and  $e'$  should to be differentiated in the egde set of  $G$ , establishing the need for  $E(G)$  to be a multi- or indexed set, notions we are trying to avoid in this document.

From now on any *directed graph* is assumed to be a directed loop graph. Although most, if not all, concepts work for directed pseudographs just as well, we want to start to heavily use the canonical edge representation, which when talking about pseudographs makes problems as mentioned before.

*Remark* (More Notation). - Check, do I really need this? For a pair of vertex sets  $X, Y \subseteq V(G)$  of a directed graph  $G$  we write

$$(X, Y)_G = \{(x, y) \in E(G) \mid x \in X, y \in Y\}$$

for the set of edges with source in  $X$  and target in  $Y$ . For vertex sets with a single element  $X = x$ , we also write  $(x, Y)_G$  and mean the edges with source  $x$  and target in  $Y$ .

**Definition 1.3** (In- and out-degree). For a directed graph  $G$  the **in-degree**  $d_G^-(x)$  of a vertex  $x$  is defined as the number of edges of  $G$  with target  $x$ , that is

$$d_G^-(x) = |(V(G), x)_G|.$$

Similarly, the **out-degree**  $d_G^+(x)$  of  $x$  is defined as

$$d_G^+(x) = |(x, V(G))_G|,$$

the number of edges in  $G$  with source  $x$ .

*Remark* (Side). In some literature about directed graphs (Bang-Jensen), loops are *not* counting towards the in- or out-degree of vertex. In the light of neural network however, we specifically want to count loops as well.

A basic property of the in- and out-degree in directed graphs is that number of in-degrees of every vertex, as well the sum of every out-degree, equal the total number of edges:

**Proposition 1.4.** *In every directed graph  $G$ , we have*

$$\sum_{x \in V(G)} d^-(x) = \sum_{x \in V(G)} d^+(x) = |E(G)|.$$

*Proof.* Since  $(V(G), x)_G \cap (V(G), y)_G = \emptyset$  for  $x \neq y$ , we can write

$$\sum_{x \in V(G)} d^-(x) = \left| \bigcup_{x \in V(G)} (V(G), x)_G \right| = |(V(G), V(G))_G| = |E(G)|.$$

Analogously for the out-degree.  $\square$

## 1.2 WALKS AND DISTANCES

Let  $G$  be a directed graph (what does it mean here?). A **walk**  $W$  in  $G$  is an alternating sequence  $(x_1, e_1, x_2, e_2, x_3, \dots, x_{n-1}, e_{n-1}, x_n)$  of vertices  $x_i$  and edges  $e_i$  from  $G$ , such that

$$s(e_i) = x_i \quad \text{and} \quad t(e_i) = x_{i+1}, \quad \text{for } i = 1, \dots, n-1,$$

that is, such that the vertices are connected by the edges inbetween them. We denote the set of vertices  $(x_1, \dots, x_n)$  of  $W$  as  $V(W)$  and the sequence of edges  $(e_1, \dots, e_{n-1})$  as  $E(W)$  (need it?).

The vertices  $x_1$  and  $x_n$  are called the *end vertices* of  $W$  and we also say that  $W$  is an  $(x, y)$ -walk. The **length** of  $W$  is defined as the length of the sequence of edges; a walk consisting of only one vertex has length zero. colon, really?

**Definition 1.5** (Distance). The **distance** of two vertices  $x, y$  in a directed graph  $G$  (means?), is defined as the minimum length of an  $(x, y)$ -walk, if any such walk exists, otherwise  $\text{dist}(x, y) = \infty$ . In short,

$$\text{dist}(x, y) = \inf\{|E(W)| \mid W \text{ is } (x, y)-\text{walk}\}.$$

$|E(W)|$  is not explained. Necessary?

**Proposition 1.6.** *The distance function  $\text{dist} : V(G) \times V(G) \rightarrow \mathbb{N}$  of a directed graph  $G$  satisfies the triangle equality,*

$$\text{dist}(x, z) \leq \text{dist}(x, y) + \text{dist}(y, z), \text{ for } x, y, z \in V(G).$$

*Proof.* Let  $x, y, z$  be vertices in  $G$ . If either no  $(x, y)$ -walk or  $(y, z)$ -walk exists, the inequality holds by definition. Other wise, let  $W$  be an  $(x, y)$ -walk of minimal length and let  $U$  be a  $(y, z)$ -walk of minimal length. Certainly, by concatenating  $W$  and  $U$  we obtain an  $(x, z)$ -walk of length  $|E(W)| + |E(U)| = \text{dist}(x, y) + \text{dist}(y, z)$ , proofing that

$$\text{dist}(x, z) \leq \text{dist}(x, y) + \text{dist}(y, z).$$

□

More to do:

- summarize category of directed (weighted) pseudographs
- weights!
- vertices will also be called nodes and neurons, edges will also be connections or synapses.
- subgraphs
- ~~vertex set, edge set  $E(G), V(G)$ .~~
- $\omega(e)$  is weight, connection strength or synaptic weight (as a side remark)
- extend to category of weighted directed pseudographs (isomorphisms)
- path
- adjacency matrix
- converses of graph related to opposite category?
- ~~in-and out-degree~~
- ~~triangle inequality for distance,  $\text{dist}(x, z) \leq \text{dist}(x, y) + \text{dist}(y, z)$~~

References for this chapter: <http://nlab.mathforge.org/nlab/show/graph>, <http://nlab.mathforge.org/nlab/show/quiver>, (Bang-Jensen and Gutin 2008)

### 1.3 RANDOM GRAPH THEORY

*focus on labeled,  
simple directed  
graphs*

From this chapter on, as it is common and practical when talking about random graph models, we move away from the abstract notion of graphs and their equivalence classes and consider *labeled graphs*, where the edge set of a graph with  $n$  vertices takes the form  $V = \{1, \dots, n\}$ . Simple directed graphs constitute the fundamental mathematical object underlying the concepts developed in this work and if not specified otherwise, all graphs are assumed to be labeled and simple directed.

The concept of a random graph was first formally introduced by Erdős and Rényi (1959). In their  $G(n, N)$  model, a graph with  $n$  vertices and  $N$  edges is randomly and with equal probability selected from the set of such possible graphs. In the same year Gilbert (1959) independently introduced his  $G(n, p)$  model, realizing edges between vertex pairs with a fixed probability  $p$ . The two models are closely related (Łuczak 1990) and overlap in literature, both at times being referred to as *Erdős-Rényi graphs*. Here we focus on the  $G(n, p)$  model, as it closer in concept to a computational implementation of a random graph. Defining it in detail in 1.7, we will refer to the random graph model as the *Gilbert random graph model*.

In general, a random graph model is a probability space over a set of graphs (Janson et al. 2000). Rather than specifying the sample space and probability measure explicitly, random graph models are often defined by a random process that generates such graphs, leaving probability measure and sample space implicit (Bollobás 2001). The term *random graph*, in the graph theoretical context, refers to the random graph model itself. Especially in the computational context however, a random graph often refers to a single graph generated by a random process. Here we try to avoid this ambiguity and strictly refer to the mathematical object as a random graph model.

Keeping in mind that the term *graph* now refers to labeled, simple directed graphs if not otherwise specified we define  $G^n$  to be the set of simple directed graphs with  $n$  vertices,

$$G^n := \{G \mid G \text{ graph}, |V(G)| = n\}.$$

We first introduce Gilbert's random graph model  $G(n, p)$  by explicitly defining a probability space over  $G^n$  and show later how the model may be realized as a random process.

**Definition 1.7** (Gilbert random graph model). Let  $n \in \mathbb{N}$  and  $0 \leq p \leq 1$ . The *Gilbert random graph model*  $G(n, p)$  is a discrete probability space over  $G^n$  with event space  $\mathcal{P}(G^n)$  and probability measure  $P$ ,

such that every graph  $G$  with  $|E(G)| = k$  edges appears with equal probability

$$P(G) = p^k(1-p)^{n(n-1)-k},$$

for  $0 \leq k \leq n(n-1)$ .

*Remark.* Clearly  $G(n, p)$  is well-defined, as there exist  $\binom{n(n-1)}{k}$  distinct labeled graphs with  $n$  vertices and  $k$  edges and thus

$$\sum_{G \in G^n} P(G) = \sum_{k=0}^{n(n-1)} \binom{n(n-1)}{k} p^k (1-p)^{n(n-1)-k} = 1,$$

after the binomial theorem.

Equivalently, the Gilbert random graph model can be defined as a stochastic process; to an empty graph with  $n$  vertices, for each of the  $n(n-1)$  vertex pairs an edge is added at random and independently with probability  $p$ . The probability to obtain a specific graph  $G$  with  $k$  edges is then obviously  $p^k(1-p)^{n(n-1)-k}$ , already proofing the equivalence, since assuming a process as above with edge probability  $p'$  such that the induced probability measure on  $G^n$  equals  $P$  from 1.7, already yields  $p = p'$  in the choice of  $n = 2$  and  $k = 1$ .

equivalent  
definition as  
random process

**Proposition 1.8.** *In- and out-degree distribution of vertices in the Gilbert random graph model are binomial.*

*Proof.* Let  $X$  be a random variable on the random graph model, mapping to the in-degree (out-degree)  $d_G(v)$  of a vertex  $v$  of a graph  $G \in G^n$ . There are  $n-1$  other vertices that, with probability  $p$ , project to  $v$  (receive input from  $v$ ), thus

$$P(X = k) = \binom{n-1}{k} p^k (1-p)^{n-1-k},$$

showing that  $P^X = \mathcal{B}_{n-1,p}$ . □

The Gilbert random graph model is therefore also often referred to as *binomial random graph*. As typical neuronal networks are large ( $n \geq 10^3$ ) with sparse connectivity ( $p \approx 0.1$ ), in- and out-degree distribution can be approximated by a Poisson distribution,  $P^X(k) \approx \text{Pois}_\lambda(k)$ , with  $\lambda = (n-1)p$ , after the Poisson limit theorem.

Most results in the study of random graph models consider  $n \rightarrow \infty$ . In this study we are mostly interested in patterns of connectivity that arise in local circuits, leaving behind limit considerations and employ the Gilbert random graph model as a reference for the development of more detailed and specific random graph models.

## 1.4 GEOMETRIC GRAPHS

The theory of geometric graphs address the embedding of graphs in  $\mathbb{R}^d$ . Planar graphs are graphs that can be drawn on a surface with their edges drawn as straight lines between the vertices, such that no two lines intersect (Diestel 2000). Here we are interested in graphs embedded on surface. In the models introduced, connectivity then depends on geometric properties, such as distance between nodes or relative orientation between vertices. The basic graph type to allow for such connectivity rules is a geometric graph.

**Definition 1.9** (Geometric directed graph). A *geometric directed graph*  $G_\Phi$  is a directed graph  $G$  paired with a map

$$\Phi : V(G) \rightarrow [0, 1]^2,$$

representing vertex positions on the unit square.

This definition diverges from the usual notion of a geometric graph, which determines the existence of edges only between nodes within a spatial distance  $x$  in a specified norm (Penrose 2003). Moreover, geometric graphs are usually only discussed in the context of random graph models, a concept first introduced by Gilbert (1961). Denote the set of geometric graphs with  $n$  edges by  $G_\Phi^n$ . The spatial embedding of geometric graphs allows us to define (random) connectivity depending on vertex positions and inter-vertex distances. Central to this study is the distance-dependent random graph model, in which edges are added with probability  $p(x)$  depending on the distance  $x$  between vertex pairs:

**Definition 1.10** (Distance-dependent random graph model). Let  $n \in \mathbb{N}$  and  $C : [0, \sqrt{2}] \rightarrow [0, 1]$  a continuous map. A *distance-dependent geometric random graph model*  $G_\Phi(n, C)$  is a random graph model over  $G_\Phi^n$ , generated by distributing uniformly at random the  $n$  vertices on the unit square and adding an edge from  $v$  to  $w$  for each vertex pair  $(v, w) \in V(G_\Phi)^2 \setminus \Delta_{V(G_\Phi)}$  with a probability  $C(x)$ , depending on the distance  $x = \|\Phi(v) - \Phi(w)\|$  between the vertices.

We call the function  $C$  the graph's *distance-dependent connection probability profile*. Clearly, connectivity statistics in the distance-dependent graph model intrinsically depend on the choice  $C$ . To develop a thorough understanding of connectivity in the model however, mapping the distribution of inter-vertex distances is equally important. Here we calculate the distribution of the distance between two random points in a square of side-length  $s$ . Being able to identify distributions of transformed random variables is integral to the calculation:

**Lemma 1.11.** *Let  $X, Y$  be independent continuous random variables with values in  $\mathbb{R}$ , denote with  $f_X$  and  $f_Y$  their probability distribution functions.*

- (1) *The distribution of the random variable  $X + Y$  is given by the probability density function*

$$f_{X+Y}(x) = \int_{\mathbb{R}} f_X(z) f_Y(z-x) dz.$$

- (2) *The distribution of the random variable  $X^2$  is given by the probability density function*

$$f_{X^2}(x) = \begin{cases} \frac{1}{2\sqrt{x}} (f_X(\sqrt{x}) + f_X(-\sqrt{x})) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

- (3) *Let  $X$  only take positive values. Then the distribution of the random variable  $\sqrt{X}$  is given by the probability density function*

$$f_{\sqrt{X}}(x) = \begin{cases} 2x f_X(x^2) & x \geq 0 \\ 0 & x < 0 \end{cases}$$

**Theorem 1.12.** *Let  $D$  be a random variable mapping to the distance of two random points in the square  $[0, s]^2$  of side-length  $s$ . Then the distribution of  $D$  is given by the probability density function*

$$f(x) = \begin{cases} \frac{2x^4 - 8sx^3 + 2\pi s^2 x}{s^4} & x \in [0, s] \\ \frac{H(x)}{s^4} & x \in [s, s\sqrt{2}), \\ 0 & x \notin [0, s\sqrt{2}) \end{cases} \quad (1.1)$$

where

$$H(x) = 8sx\sqrt{x^2 - s^2} - 2x^3 - 4s^2x \left( 1 + \arcsin \left( 1 - \frac{2s^2}{x^2} \right) \right).$$

*Proof.* We follow the approach described by Moltchanov (2012). Consider two independently and uniformly distributed random points  $p_1 = (x_1, y_1)$  and  $p_2 = (x_2, y_2)$  in  $[0, s]^2$ . The distance between  $p_1$  and  $p_2$  is given as

$$\|p_1 - p_2\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

As a first step we calculate the distribution for  $\Delta_x = x_1 - x_2$ . Denote with  $f_{\Delta_x}$  its probability density. Then, since  $f_{-x_2}(z) = f_{x_2}(-z)$ ,

$$f_{\Delta_x}(d) = f_{x_1 + (-x_2)}(d) = \int_{\mathbb{R}} f_{x_1}(z) f_{x_2}(z-d) dz \quad (1.2)$$

after Lemma 1.11. Density functions for  $x_1$  and  $x_2$  are given by

$$f_{x_1}(z) = f_{x_2}(z) = \begin{cases} \frac{1}{s} & \text{for } z \in [0, s] \\ 0 & \text{otherwise.} \end{cases}$$

Thus we may only obtain non-zero values in (1.2) for  $d \in (-s, s]$ , as otherwise either one of the factors in the integrand is zero. In full we obtain the triangular distribution (Simpson 1755),

$$f_{\Delta_x}(d) = \begin{cases} 0 & d \notin (-s, s] \\ \frac{s+d}{s^2} & d \in (-s, 0] \\ \frac{s-d}{s^2} & d \in (0, s]. \end{cases}$$

Next we calculate the distribution for  $\Delta_x^2 = (x_1 - x_2)^2$ . Using Lemma 1.11 once again we obtain for  $d > 0$

$$\begin{aligned} f_{\Delta_x^2}(d) &= \frac{1}{2\sqrt{d}} \left( f_{\Delta_x}(\sqrt{d}) + f_{\Delta_x}(-\sqrt{d}) \right) \\ &= \frac{1}{2\sqrt{d}} \left( \frac{s - \sqrt{d}}{s^2} + \frac{s + (-\sqrt{d})}{s^2} \right) = \frac{1}{s\sqrt{d}} - \frac{1}{s^2}, \end{aligned}$$

and  $f_{\Delta_x^2}(d) = 0$  for  $d \leq 0$ . Note that of course,  $f_{\Delta_x^2} = f_{\Delta_y^2}$  and we will refer to this density function as  $f_{\Delta}^2$ . Convolution yields again the probability density function for the sum of the random variables  $\Delta_x^2$  and  $\Delta_y^2$ ,

$$f_{\Delta_x^2 + \Delta_y^2}(d) = \int_{\mathbb{R}} f_{\Delta^2}(z) f_{\Delta^2}(d-z) dz.$$

Finally Lemma 1.11 lets us compute

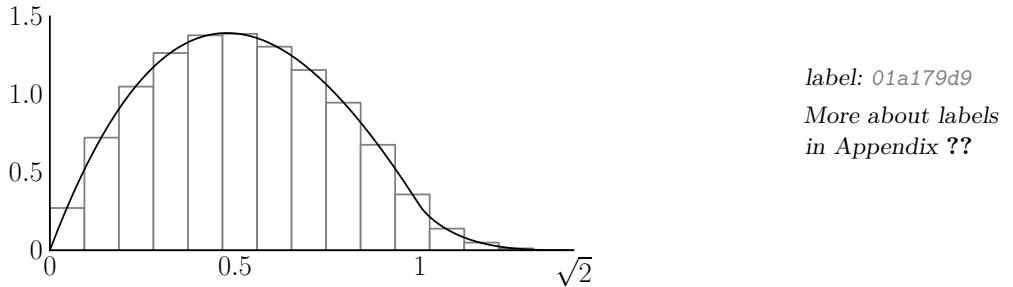
$$f_D(d) = f_{\sqrt{\Delta_x^2 + \Delta_y^2}}(d) = 2d f_{\Delta_x^2 + \Delta_y^2}(d^2),$$

for  $d \geq 0$ , yielding expression (1.1) for probability density function of  $D$  (Mathematica ??, ??).  $\square$

The distribution for the distance between two random points in the unit square  $[0, 1]^2$ , is then obtained from (1.1) by setting  $s = 1$ . The probability density function  $f$  becomes

$$f(x) = \begin{cases} 2x^3 - 8x^2 + 2\pi x & x \in [0, 1) \\ 8x\sqrt{x^2 - 1} - 2x^3 & x \in [1, \sqrt{2}) \\ -4x - 4x \arcsin\left(1 - \frac{2}{x^2}\right) & x \in [\sqrt{2}, 1] \end{cases} \quad (1.3)$$

Plotting  $f$  in combination with the results of a numerical simulation (10000 points randomly distributed on unit square, extracted distances binned to a of width  $\sqrt{2}/15$ ) verifies our calculation:



Density functions 1.1 and 1.3 are of high importance. Here we use 1.3 to compute the probability that a random pair of vertices in the distance-dependent random graph model is connected:

**Corollary 1.13.** *Let  $v \neq w$  be vertices in an arbitrary realization of  $G_\Phi(n, C)$ . The probability  $p$  for an edge from  $v$  to  $w$  is given by*

$$p = \int_0^{\sqrt{2}} C(x) f(x) dx,$$

where  $f(x)$  is the probability density function (1.3).

**Example** Let  $C(x) = 1 - \frac{x}{\sqrt{2}}$ . Then we compute the probability to find an edge between a random vertex pair in  $G_\Phi(n, C)$  according to Theorem 1.13 as

$$p = \int_0^{\sqrt{2}} f(x) dx - \frac{1}{\sqrt{2}} \int_0^{\sqrt{2}} x f_D(x) dx = 1 - \frac{1}{\sqrt{2}} \mathbf{E}[D].$$

The expected distance between two random points on the unit square is computed as  $\mathbf{E}[D] \approx 0.521405$  (??), a result confirmed by Philip (2007). Thus we obtain the probability for connection of  $p \approx 0.631311$



# 2

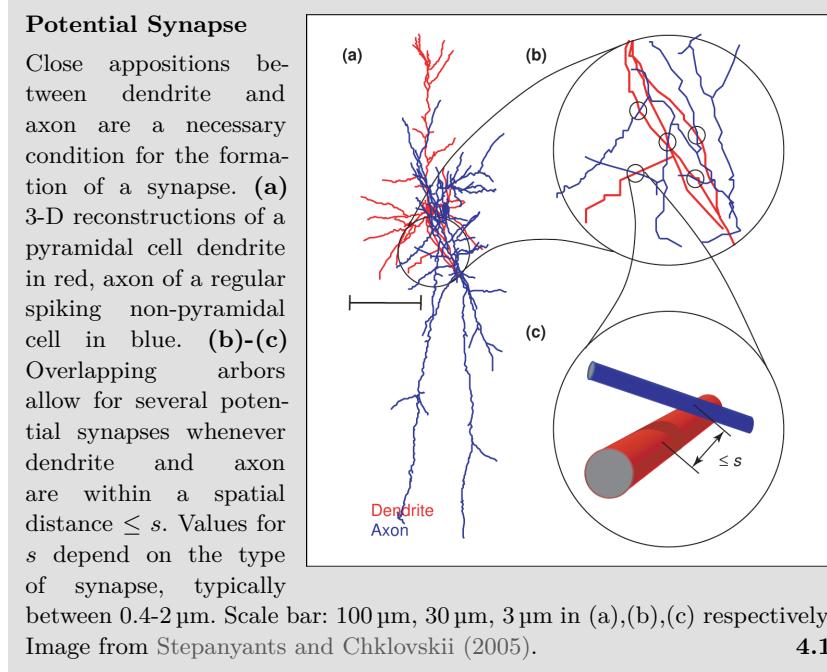
## NETWORK MODEL

---

Referring to anisotropic characteristics in local cortical circuits of the rat's brain, a network model implementing anisotropic tissue geometry is developed. The introduction of a rewiring algorithm and qualitative anisotropy measure lay the foundation for the analysis of structural aspects of this model in Chapter ??.

## 2.1 ANISOTROPY IN NEURAL CONNECTIVITY

Neurogeometry addresses the problem of inferring synaptic connectivity from the geometric shapes of axon and dendrites. A fundamental concept in this field is that of a *potential synapse* (Stepanyants et al. 2002). Defined as the potential axonal-dendritic connection of two neurons, present whenever the axon of one neuron is within a spatial distance  $s$  of the dendrite of the other, it is a necessary, although not sufficient, condition for the formation of a synaptic connection (Figure 2.1). The existence of such close appositions solely depends on dendritic and axonal anatomy; identification of defining morphological characteristics in both axon and dendrite would therefore allow for a model of local network connectivity, assuming for example that a certain ratio  $r$  of potential synapses turn into active contacts independently. It is the hope that such a model, motivated from the geometry of a neuron's functional compartments, not only displays inherent patterns of connectivity similar to what has been observed in biological networks, but also proofs itself as a testing ground for how this connectivity may affect network dynamics.

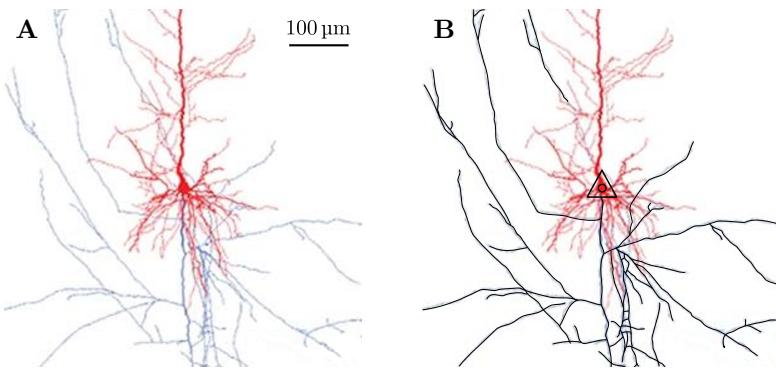


high variability in  
axonal  
morphology

Finding stereotypical anatomical characteristics however is difficult, as axonal morphology is, in general, highly diverse (Debanne 2004). Across different species, distinct regions in the central nervous system and different neuron types, axons display a wide variety of shapes characterized by morphometric parameters such as total length, branching complexity and axonal extent (Ropireddy et al. 2011). Typical exam-

ples of distinct morphology include the T-shaped axons of cerebellar granule cells branching only at a singular point (Ramon and Cajal 1911), and axons of hippocampal CA3 pyramidal cells, which, in stark contrast, may feature up to 40 branches resulting in a total length of axon collaterals of up to 12 mm (Ishizuka et al. 1990).

It is therefore imperative to confine this analysis to a specific brain region and neuron type. In this study, we set the focus on circuits of pyramidal cells in the mammalian cortex. More specifically, local circuits of thick tufted layer V pyramidal neurons in the rat's somatosensory cortex have been the target of advanced experimentation (Song et al. 2005; Perin et al. 2011; Romand et al. 2011; Ramaswamy et al. 2012), and will serve as a benchmark for results in neural morphology and network connectivity in this report.



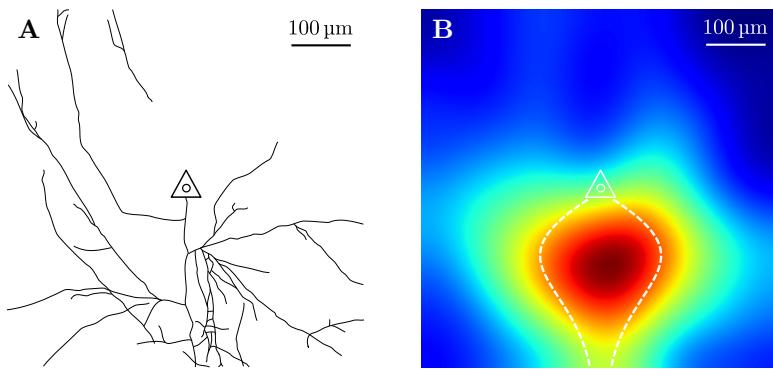
**Figure 2.2: Tracing axonal branching of a pyramidal cell** In a 3-D model reconstructed from biocytin-labeled thick-tufted layer V pyramidal cells in the somatosensory cortex of postnatal (day 14) Wistar rats, Romand et al. (2011) depict dendritic compartments in red, axonal compartments in blue. **A)** A 600  $\mu\text{m}$  window centered on the soma of the pyramidal cell shows the main stem of the cell's axon projecting downwards in a straight line, collaterals branching at various angles. **B)** Using image manipulation software, axon morphology was manually traced and is emphasized in black.

Axonal morphology of pyramidal cells in the cerebral cortex is well described. From the soma the single main stem of the axon originates and projects downwards, describing a trajectory closely resembling a straight line (Braitenberg and Schüz 1998). At arbitrary points along this path, collaterals branch off at various angles and constitute themselves linear paths until they further ramify or terminate. Displaying a high degree of ramification, axonal trees of cortical pyramidal cells build, in general, complex structures (Petersen et al. 2003; Ramaswamy et al. 2012). Cortical slice experiments analyzing neural anatomy are typically constrained by a slice thickness of 300  $\mu\text{m}$ . On this scale, 3-D reconstruction from labeled thick tufted layer V pyramidal cells reveals

cortical axons  
form straight lines,  
arborize profusely

characteristic morphology of the axonal tree (Figure 2.2). The downwards projecting, straight axon branches at several points, forming collateral branches that travel in linear path as well.

In a statistical view, this characteristic axonal morphology results in high axon branch densities along the main stem, whereas distant regions display a relatively low density (Figure 2.3). Specifically, axon collaterals do not cluster around the soma but align with the main stem's projection. As presence of an axonal branch constitutes a necessary condition for a potential synapse, a higher concentration of potential and, subsequently, realized synapses is expected in regions of high branch density. For a coherent picture of local connectivity profiles, however, dendritic morphology needs to be considered as well.

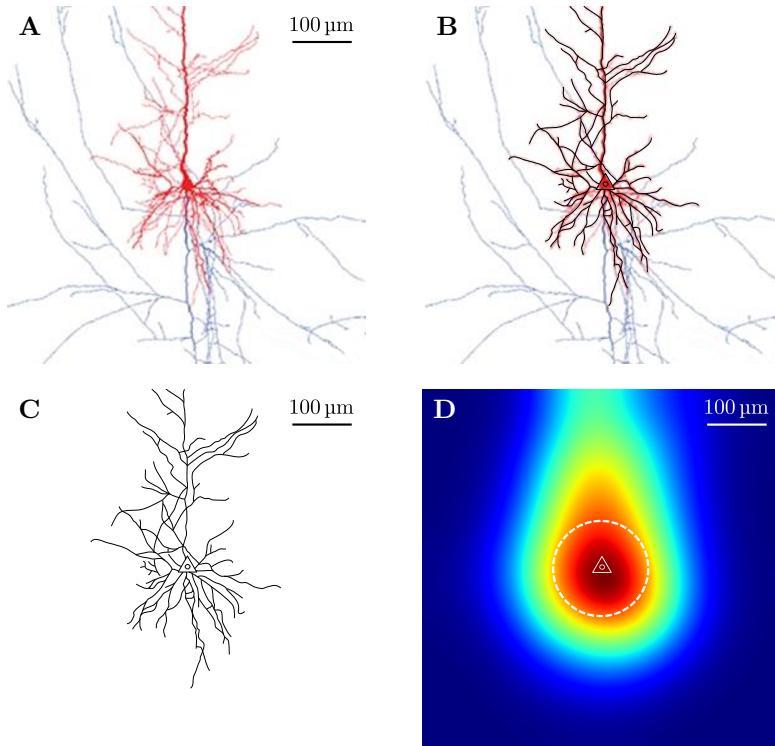


**Figure 2.3: Illustrating axonal branch density** In a sample of 5 reconstructions from thick-tufted layer V pyramidal cells (Romand et al. 2011), tracing axonal morphology illustrates characteristic branch density along the axon's main stem. **A)** Example of extracted axonal tree. Outline manually traced using image manipulation software. Soma indicated by triangle. Original data from Romand et al. (2011). **B)** Overlaying 5 axonal trees extracted as in A), applying a Gaussian filter and displaying high axon densities in warm colors, illustrates the characteristic higher branch densities along the axon's main stem.

*basal dendrites dominate local connectivity*

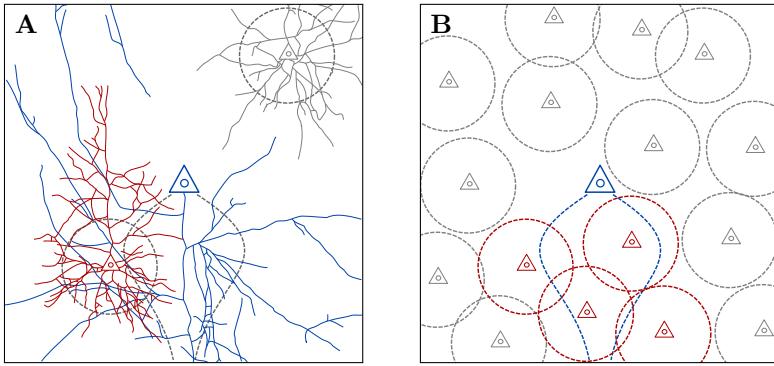
Dendritic anatomy of cortical pyramidal cells is inherently bipartite. From the soma several *basal dendrites* emerge and extend into arbitrary directions, branching profusely until they terminate. The single *apical dendrite* emerges from the apex of the pyramidal cell and ascends in a linear trajectory, forming occasional collateral branches until finally terminating into the apical tuft, where the dendrite branches several times to form a tree like structure (Feldman 1984). On the scale of typical cortical slice thickness, however, the apical dendrite is cut off and the basal dendrite dominates the dendritic morphology and potential of dendritic-axonal connections (Figure 2.4). The radial extension of dendritic branches results in a high concentration of dendritic branches

around the soma, much in the contrast to the findings of axonal branch densities before.



**Figure 2.4: Dendritic morphology and branch density** Using neuronal morphology of thick-tufted layer V pyramidal cells recorded by Romand et al. (2011), dendritic anatomy is traced and combined to illustrate high branch density around the soma. **A)** In a 600  $\mu\text{m}$  window centered on the soma, basal dendrites (red) are visible extending around the soma. The ascending thick apical dendrite (red) is cut off and apical tuft is not shown. **B)-C)** Manual tracing of dendritic outlines in five samples (one shown), allows for clearer identification of stereotypical morphology and later analysis. **D)** Combining 5 dendritic outlines as shown in C) and subsequent Gaussian filtering reveals the relatively high dendritic branch density around the soma.

Combining the above results of dendritic and axonal branch densities in the light of neurogeometry, a clear concept of anisotropy of neural connectivity emerges. As dendritic branches of potential post-synaptic targets extend radially from the soma and do not display a preferred direction, target neurons for outgoing synaptic contacts originating from a single pyramidal cell, cluster around the downwards projecting axon ([Figure 2.5](#)). In their in-depth study, Stepanyants and Chklovskii (2005) confirm the overrepresentation of potential synapses along the axon for pyramidal cells. Consistent with the notion that stereotypical morphology of pyramidal cells is intrinsic to the local network's connec-



**Figure 2.5: Connected neurons of a single pyramidal cell align with axonal projection** Reducing the full axonal (blue, cf. Figure 2.2) and dendritic trees (red, gray, cf. Figure 2.4) as shown for two neurons in A) to their stereotypical axonal (blue) and dendritic profiles (red, gray) in B), demonstrates how connected neurons (red) tend to cluster around the pre-synaptic axon's profile, as spatial closeness constitutes a necessary condition for the formation of contacts. Unconnected neurons (gray) are found distant from the axon's projection, but not necessarily distant from the soma.

tivity profile, they also find that anisotropy of this degree is *not* present in spiny stellate neurons located in lower-layer-4.

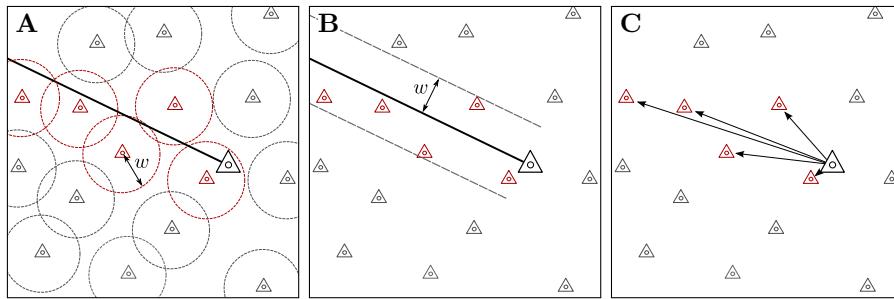
## 2.2 ANISOTROPIC GEOMETRIC NETWORK MODEL

In this section we formulate a model of network connectivity incorporating anisotropy as outlined in the last section. A

With this in mind, we propose the following model: On a square surface of side length  $s$ , a number of  $N$  point neurons are randomly, uniformly distributed. Connected neighbors are then calculated for each neuron separately and independently, by determining the randomly, uniformly distributed direction of the neuron's axon. In this direction the axon traverses over the surface describing a straight path, terminating only when an edge of the surface is reached. Directed contacts are made with every neuron that is within a width  $w(x)$  of the axon's trajectory, where in general  $w$  depends on the axon length  $x$  at this point (Figure 2.6).

random axonal orientation yields relevant connectivity

The implementation of arbitrary axonal orientation is crucial to the model. Although cortical axons are described as consistently projecting downwards (Braitenberg and Schüz 1998, cf. Section 2.1), combining exclusively vertically aligned axons with the simplified axonal and dendritic morphological profiles would result in a “vertically staggered connectivity” - neurons could then only project to targets located be-



**Figure 2.6:** Anisotropic geometric network model and interpretations of width parameter  $w$  Illustrating the process of finding connections for one neuron (large triangle, black), the axon describes a linear trajectory in an arbitrary direction and until terminating on the surface's edge. Target neurons (red) are encountered along the path within a distance  $w$ , which is in **A)** interpreted as a dendritic radius or, equivalently, in **B)** as a “bandwidth” of the axon. Connections to the encountered targets are then established as projections in **C)**, consistent with the directed nature of synapses in biological networks (cf. Chapter ??).

low them. It is in fact not a vertical alignment of axon orientation, but the anisotropy in neural connectivity - the observation of neuronal targets aligning with the axonal projection - that we try to capture and analyze in this model.

We will refer to the model as the *anisotropic geometric network model*. Mathematically it is interpreted as a random graph model, in which the resulting object of a realization of the random process is a geometric directed graph with a special mode of connectivity. We can formally define such realization as:

**Definition 2.1** (Anisotropic geometric graph). Let  $n \in \mathbb{N}$  and  $w \in (0, \infty)$ . An *anisotropic geometric graph*  $G_{n,w}$  then consists of a tuple  $(G, \Phi, a)$ , of a simple directed graph  $G$  with  $|V(G)| = n$  vertices and the maps

$$\Phi : V(G) \rightarrow [0, 1]^2, \quad a : V(G) \rightarrow [0, 2\pi),$$

such that for every vertex pair  $v, v' \in V(G)$  and edge  $e \in E(G)$  with  $s(e) = v$  and  $t(e) = v'$  exists if and only if  $\leq w$ . where denotes euclidean distance l2 norm

The anisotropic random graph model then is then giving the probability distribution over the set of anisotropic random graphs by describing a random process generating such graph.

**Definition 2.2** (Anisotropic random graph model). Let  $n \in \mathbb{N}$  and  $w > 0$ . The *anisotropic random graph model*  $G(n, w)$  is a probability space over the set of anisotropic geometric graphs with a probability

distribution induced by the following process: Let  $G$  be an empty graph with  $n$  vertices. Assign randomly and uniformly to every vertex  $v \in V(G)$  a position  $\Phi(v) \in [0, 1]^2$  and axonal orientation  $0 \leq a(v) < 2\pi$ . Then add edges such that  $(G, \Phi, a)$  is an anisotropic geometric graph  $G_{n,w}$ .

*anisotropic model  
is scale-free*

*Remark.* As with every geometric graph model introduced, we restrict the surface to be the unit square. This does not limit the model, as in fact only the relative width of the axon band in regards to the surface's side length is determining connectivity statistics; the expected number of connections is easily obtained by division of the area covered by the axon and the area of the square, making connectivity statistics int the anisotropic random graph model scale-free.

The subsequent is a study of anisotropic geometric graphs. To enable this analysis, some prior work which composes the rest of this chapter. Integral to is a numerical implementation. The anisotropic network model is

### 2.3 NUMERICAL IMPLEMENTATION

for numerical considerations was achieved in Python. Using convenient

$N$  Normal distribution.  $E$ . For a in  $N$

Computational implementation was achieved by

Implementing the model as an algorithm in Python, we obtain .

in connectivity stored in graph tool (Peixoto 2014)

Harnessing the computational implementation, we generated a sample of 25 networks with following parameters.

*determine  
parameter set to  
generate sample  
graphs*

To harness the numerical implemenation to generate networks, a set of parameters needs to be chosen. The network size  $N$  is solely governing computational efforts in subsequent calculations and has thus been set as  $N = 1000$ . As shown in ??, only the quotient of the side length parameter  $e$  and axon width  $w$  . As such, side length  $e$  is arbitrarily set as  $e = 100$  and leaves axon width for determination.

First, we determine  $w$  to be constant. Although simplistic profiles ([Figure 2.3](#)) and makes for characteristic distance distribution as we will see later ([Figure 2.8](#)), more in line with idea of abstractness and sim-

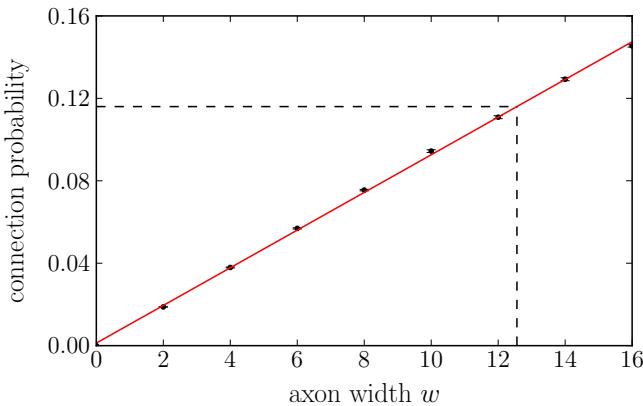
plasticity. For small  $w$  then, the overall connection probability  $p$  can be approximated as

$$p = \frac{Lw}{E^2},$$

where  $L$  is the average length of the axon until it terminates on a surface edge.

Having established the connection between ,

The final parameter is then the axon-profile width  $w$ . In their analysis of connectivity of thick-tufted layer V pyramidal cells in neonatal rats (day 14), Song et al. (2005) report an overall connection probability of  $p = 0.116$ , consistent with prior reports of a cortical connection probability of  $p \approx 0.1$ .



**Figure 2.7: Tracing axonal branching of a pyramidal cell** In a 3-D model reconstructed from biocytin-labeled thick-tufted layer V pyramidal cells in the somatosensory cortex of postnatal (day 14) Wistar rats, Romand et al. (2011) depict dendritic compartments in red, axonal compartments in blue. **A)** A 600  $\mu\text{m}$  window centered on the soma of the pyramidal cell shows the main stem of the cell's axon projecting downwards in a straight line, collaterals branching at various angles. **B)** Using image manipulation software, axon morphology was manually traced and is emphasized in black.

Having no further evidence at the time on how to select axon width, we set  $w(x)$  to be constant. This leaves

With this parameter set we generate a sample of 25 graphs

## 2.4 DISTANCE DEPENDENT CONNECTIVITY

random graph  
models in  
section 1.3

In Gilbert's random graph model  $G(n, p)$ , probability of connection  $p$  is independently chosen and a fixed value for all vertex pairs. The anisotropic geometric graph model introduced in section 2.2 is itself a random graph model - node positions as well as preferred directions of connection are uniformly at random distributed. In contrast to Gilbert's model however, neither is the probability of connection between a given vertex pair independent of the realization of other edges in the graph, nor is it a fixed value - probabilities strongly depend on internode distance in the anisotropic geometric graph model introduced.

Analyzing dependencies in the anisotropic model, specifically by identifying prevalent patterns of connectivity and relating these modes of non-randomness to biological findings, is the main focus of chapter ???. However, such structural correlations may not necessarily be an inherent feature of the network's anisotropy - distance dependent connectivity alone, as imposed by the model's specific geometry, may be the cause for emerging dependencies. It is therefore a crucial initial task to map the anisotropic model's distance dependent connection probability. Inferring connection probability as a function of internode distance and comparing it with computational results, in this section we explore distance connectivity of the anisotropic network model, securing a vital component in the analysis of structural features.

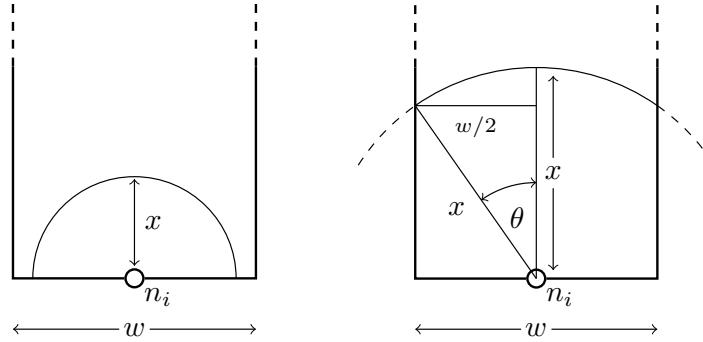
**Theorem 2.3.** *Let  $(G, \Phi, a)$  represent an arbitrary realization of the anisotropic random graph model  $G(n, w)$ . Define  $C : [0, \sqrt{2}] \rightarrow [0, 1]$  as the distance-dependent connection probability profile of  $(G, \Phi)$ , that is such that  $C(x)$  is the probability that for a vertex pair  $(v, v') \in V(G)^2 \setminus \Delta_{V(G)}$  in distance  $x = \|\Phi(v) - \Phi(v')\|$  the vertex  $v$  projects to vertex  $v'$ . Then*

$$C(x) = \begin{cases} 0.5 & \text{for } x \leq w/2 \\ \frac{1}{\pi} \arcsin\left(\frac{x}{2w}\right) & \text{for } x > w/2. \end{cases}$$

*Proof.* Let  $v, v'$  be a pair of vertices in  $V(G)^2 \setminus \Delta_{V(G)}$  in distance  $\|\Phi(v) - \Phi(v')\| = x$  of each other. The vertex  $v$  projects  $v'$  only for suitable axonal orientations  $a(v)$ .

To see this, consider a given source vertex  $v$  at  $(0, 0)$  and a possible target  $w$ , such that  $d(v, w) = x$ . We may then express the target coordinates as  $xe^{i\varphi}$ ,  $0 \leq \varphi < 2\pi$ .

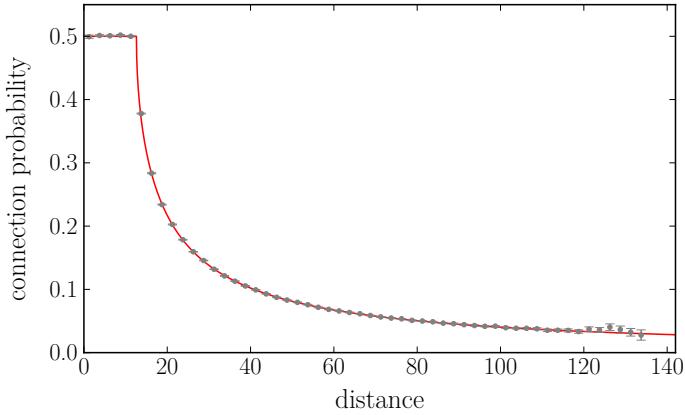
Figure ?? illustrates for which angles  $\varphi$  the node  $w$  becomes a valid target for an edge from  $v$ . This intervall



For a general  $v$  make coordinate transformation

□

We can verify this result by computationally extracting the distance dependencies in the sample graphs generated.



**Figure 2.8: Predicted distance-dependent connection probability profile is matched by numerical computation** In a 3-D model reconstructed from biocytin-labeled thick-tufted layer V pyramidal cells in the somatosensory cortex of postnatal (day 14) Wistar rats, Romand et al. (2011) depict dendritic compartments in red, axonal compartments in blue. **A)** A 600  $\mu\text{m}$  window centered on the soma of the pyramidal cell shows the main stem of the cell's axon projecting downwards in a straight line, collaterals branching at various angles. **B)** Using image manipulation software, axon morphology was manually traced and is emphasized in black.

## 2.5 REWIRING

It is in our highest interest to compare results to. To this end we introduce an algorithm that preserves distance-dependent connectivity as

*eliminate  
anisotropy  
through rewiring*

found in Proposition 2.3, but eliminates anisotropy in network connectivity by consecutively rewiring existing connections to new suitable targets.

**Algorithm 2.4.** Let  $N(n, e, w) = (G, P, a)$  be Then

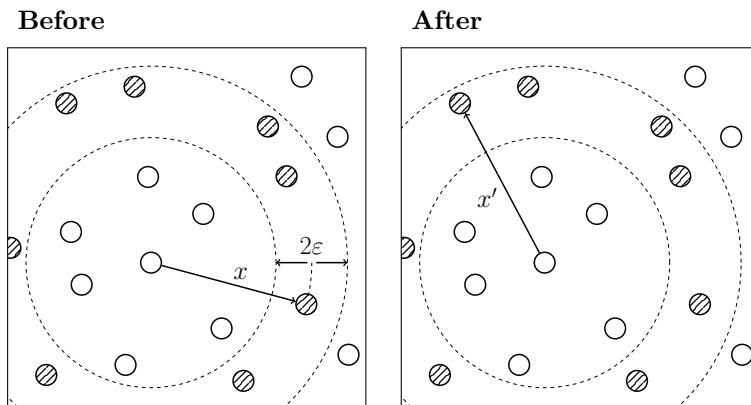
```

for  $v \in V(N_G)$  do
  for  $e \in E_{\text{out}}(v)$  do
     $x \leftarrow \|N_P(v) - t(e)\|$ 
     $T \leftarrow \{w \in V(N_G) \mid x - \varepsilon \leq \|N_P(v) - N_P(w)\| < x + \varepsilon\}$ 
     $t(e) \leftarrow \text{choice}T$ 
```

is defined.

**Proposition 2.5.** Preserves distance-connectivity.

for  $\in V(G)$  do



**Figure 2.9: Rewiring** finding new target in distance  $x'$  such that  $x - \varepsilon \leq x' < x + \varepsilon$ .

## 2.6 ANISOTROPY MEASURE

## 2.7 SUMMARY AND DISCUSSION

## BIBLIOGRAPHY

---

- Bang-Jensen, Jørgen and Gregory Z. Gutin (2008). *Digraphs: Theory, Algorithms and Applications*. 2nd ed. London: Springer. DOI: 10.1007/978-1-84800-998-1.
- Bollobás, Béla (2001). *Random graphs*. 2nd ed. Cambridge University Press.
- Braitenberg, Valentino and A Schüz (1998). *Cortex: Statistics and Geometry of Neuronal Connectivity*. 2nd ed. Berlin; New York: Springer.
- Debanne, Dominique (2004). Information processing in the axon. *Nature Reviews Neuroscience* 5.4, pp. 304–316.
- Diestel, Reinhard (2000). *Graph Theory*. 2nd ed. Vol. 173. Graduate Texts in Mathematics. Springer-Verlag, New York.
- Erdős, P. and A. Rényi (1959). On random graphs, I. *Publicationes Mathematicae (Debrecen)* 6, pp. 290–297.
- Feldman, M.L (1984). Morphology of the neocortical pyramidal neuron. *Cerebral cortex, Vol 1, Cellular components of the cerebral cortex*. Ed. by Alan Peters and Edward G Jones. New York: Plenum Press, pp. 123–200.
- Gilbert, E. N. (1959). Random Graphs. *The Annals of Mathematical Statistics* 30.4, pp. 1141–1144. DOI: 10.1214/aoms/1177706098.
- Gilbert, E.N. (1961). Random Plane Networks. *Journal of the Society for Industrial and Applied Mathematics* 9.4, pp. 533–543. DOI: 10.1137/0109045.
- Ishizuka, Norio, Janet Weber, and David G Amaral (1990). Organization of intrahippocampal projections originating from CA3 pyramidal cells in the rat. *Journal of Comparative Neurology* 295.4, pp. 580–623.
- Janson, S., T. Łuczak, and A. Rucinski (2000). *Random Graphs*. Wiley Series in Discrete Mathematics and Optimization. Wiley.
- Łuczak, Tomasz (1990). On the equivalence of two basic models of random graph. *Proceedings of Random graphs*. Vol. 87, pp. 151–157.
- Moltchanov, D. (2012). Distance distributions in random networks. *Ad Hoc Networks* 10.6, pp. 1146 –1166. DOI: 10.1016/j.adhoc.2012.02.005.
- Peixoto, Tiago (2014). *graph-tool - An efficient python module for manipulation and statistical analysis of graphs*.

- Penrose, Mathew (2003). *Random Geometric Graphs (Oxford Studies in Probability)*. Oxford University Press, USA.
- Perin, Rodrigo, Thomas K. Berger, and Henry Markram (2011). A synaptic organizing principle for cortical neuronal groups. *Proceedings of the National Academy of Sciences* 108.13, pp. 5419–5424. DOI: 10.1073/pnas.1016051108.
- Petersen, Carl C. H., Amiram Grinvald, and Bert Sakmann (2003). Spatiotemporal Dynamics of Sensory Responses in Layer 2/3 of Rat Barrel Cortex Measured In Vivo by Voltage-Sensitive Dye Imaging Combined with Whole-Cell Voltage Recordings and Neuron Reconstructions. *The Journal of Neuroscience* 23.4, pp. 1298–1309.
- Philip, J. (2007). The probability distribution of the distance between two random points in a box. TRITA MAT 07 MA 10.
- Ramaswamy, Srikanth, Sean L Hill, James G King, Felix Schürmann, Yun Wang, and Henry Markram (2012). Intrinsic morphological diversity of thick-tufted layer 5 pyramidal neurons ensures robust and invariant properties of in silico synaptic connections. *The Journal of physiology* 590.4, pp. 737–752.
- Ramon, Y and S Cajal (1911). Histologie du systeme nerveux de l'homme et des vertebres. *Maloine, Paris* 2, pp. 153–173.
- Romand, Sandrine, Yun Wang, Maria Toledo-Rodriguez, and Henry Markram (2011). Morphological development of thick-tufted layer V pyramidal cells in the rat somatosensory cortex. *Frontiers in Neuroanatomy* 5.5. DOI: 10.3389/fnana.2011.00005.
- Ropireddy, Deepak, Ruggero Scorcioni, Bonnie Lasher, Gyorgy Buzsáki, and GiorgioA. Ascoli (2011). Axonal morphometry of hippocampal pyramidal neurons semi-automatically reconstructed after in vivo labeling in different CA3 locations. *Brain Structure and Function* 216.1, pp. 1–15. DOI: 10.1007/s00429-010-0291-8.
- Simpson, Thomas (1755). A letter to the Right Honourable George, Earl of Macclesfield, President of the RoyalSociety, on the advantage of taking the mean ofa number of observations in practical astronomy. *Phil. Trans. R. Soc. Lond* 49, pp. 82–93.
- Song, Sen, Per Jesper Sjöström, Markus Reigl, Sacha Nelson, and Dmitri B Chklovskii (2005). Highly Nonrandom Features of Synaptic Connectivity in Local Cortical Circuits. *PLoS Biol* 3.3, pp. 507–519. DOI: 10.1371/journal.pbio.0030068.
- Stepanyants, Armen and Dmitri B. Chklovskii (2005). Neurogeometry and potential synaptic connectivity. *Trends in Neurosciences* 28.7, pp. 387 –394. DOI: 10.1016/j.tins.2005.05.006.
- Stepanyants, Armen, Patrick R. Hof, and Dmitri B. Chklovskii (2002). Geometry and Structural Plasticity of Synaptic Connectivity. *Neuron* 34.2, pp. 275 –288. DOI: 10.1016/S0896-6273(02)00652-9.

## INDEX

---

anisotropic geometric  
graph, 19  
random graph model, 19  
apical dendrite, 16  
  
basal dendrite, 16  
binomial random graph, 7  
  
directed graph, 1  
distance-dependent  
    connection probability, 8  
    random graph model, 8  
  
geometric graph, 8  
Gilbert random graph model, 6  
  
in-degree, 3  
  
loop graph, 1  
  
multigraph, 1  
  
neuro geometry, 14  
  
out-degree, 3  
  
potential synapse, 14  
pseudograph, 1  
  
random graph, 6  
  
simple graph, 1  
  
weighted graph, 2

<b>Symbol</b>	<b>Description</b>
$L$	Length
$Ma$	Mach number
$p$	Pressure