An Alternative Method for Characterization and Comparison of Plant Root Shapes

A thesis submitted to the

College of Graduate and Postdoctoral Studies

in partial pulfillment of the requirements

for the degree of Master of Science

in the Department of School of Environment and Sustainability

University of Saskatchewan

Saskatoon

By

Yujie Pei

Contents

Existing Morphological Descriptors for Root Systems

1.1 Background

1.1.1 Importance of Roots

- Mechanical and functional abilities of plant roots

- Plant root plasticity in the resource-limited environment

1.1.2 Importance of Research

1.1.2.1 Demand in crop breeding programs

1.1.2.2 Environment and Sustainability

- reduce the negative impacts of fertilization

- high crop productivity to feed the increasing global population

1.1.3 Promising application in other branching structures

- Trees

- River networks in geography

- Blood vessels in medicine

- Leaf vein networks

1.2 Summary of Existed Descriptors

1.2.1 Metric

1.2.1.1 Basic Geometric Descriptors

- maximum depth

- maximum width

- etc.

1.2.1.2 Compound Descriptors (Computed From the Basic Descriptors)

- Density

- aspect ratio

- etc.

1.2.1.3 Weaknesses

- Rely on the resolution of the images [1]

- Only provide a general view of root morphology [1]

- Difficult to assess the spatial configuration of roots

- Fail to describe the full complexity of root systems

1.2.2 Non-metric

1.2.2.1 Topological Analysis

- Persistent homology [12]

    - Persistence barcode shows the number of branches.

    - Persistence barcode indicates how branched roots connect along the scale of the function (e.g. geodesic distance).

    - Compare the similarity of branching structures by a pair-wise distance matrix using the bottleneck distance method.

- Horton-Strahler index [16]

    - Categorize the topological complexity of the whole branching structure.

- Provide a numerical measure of connectedness and complexity of the branching at each vertex by a dimensionless ratio: bifurcation ratio.

- The range of index and the length ratio indicate the size of the branching structure.

- Fractal Analysis [15]

  - Measure the complexity of branching structures.

  - Measure self-similarity of branching structures by fractal dimension of the root systems.

1.2.2.2 Strengths

- Highly complementary to geometric descriptors to characterize how individual roots are connected through branching [3].

- Describe branching structures independent of transformation and deformation.

1.2.2.3 Problems and Weaknesses

- Only analyze the connectedness of branching structures, which is a portion of the complexity of plant root systems.

- Fail to characterize spatial distribution.

- Some biologically topological indices analyze the root growth qualitatively based on line-linked systems [5], but not characterize the mathematical topological properties.

- The fractal analysis aims to describe self-similar structures, which grow by continually repeating simple growth rules [6].

1.3   Problem Statements

...

An Alternative Mathematical Method for Shape Description

2.1 Kac's Idea: Can One Hear the Shape of a Drum? [10]

2.1.1 Interpretations of Kac's Problem

- When the drum vibrates, one can hear the sound, which is composed of tones of various frequencies. How much can shape features be inferred from hearing a discrete spectrum of pure tones produced by a drum?

- If a complete sequence of eigenvalues of the Dirichlet problem for the Laplacian can be obtained precisely, will people determine the shape of a planar?

2.1.2 Problem Statement

- Consider a simply connected membrane Ω in the Euclidean space bounded by a smooth convex curve $\partial\Omega$ (e.g. a drum without any holes)

- Find function ϕ on the closure of Ω, which vanishes at the boundary $\partial\Omega$, and a number λ satisfying $-\Delta\phi = \lambda\phi$.

  - Δ is the Laplace operator. e.g. $\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2}$ in Cartesian coordinate system.

  - If there exists a solution $\phi \neq 0$, the corresponding λ is defined as a Dirichlet eigenvalue.

  - For each domain Ω, there has a sequence of eigenvalues $\lambda_1, \lambda_2, \lambda_3, ...$ corresponding to a set of eigenfunction $\phi_1, \phi_2, \phi_3, ...$

  - ϕ_k form an orthonormal basis of $L^2(\Omega)$ of real valued eigenfunctions; the corresponding discrete Dirichlet eigenvalues are positive ($\lambda_k \in \mathbb{R}^+$).

- An important function [9]:

$$h(t) = \sum_{k=1}^{\infty} e^{-\lambda_k t} \tag{2.1}$$

  - It is a Dirichlet series.

  - It is called the spectral function or the heat trace.

  - It is smooth and converges for every $t > 0$.

2.1.3 Summarize the Results of Kac's Idea

$$h(t) = \sum_{k=1}^{\infty} e^{-\lambda_k t} \sim \frac{|\Omega|}{2\pi t} - \frac{L}{4}\frac{1}{\sqrt{2\pi t}} + \frac{1}{6} \tag{2.2}$$

- As $t \to 0^+$, the leading terms of the asymptotic expansion of $h(t)$ imply the geometrical attributes of Ω

- the total area

- the perimeter

- the curvature

- If the domain Ω has the polygonal boundary, the third term shows in the information about the interior angles of the polygon [9].

2.1.4 Conclusion

2.1.4.1 Advantages

- Kac proposed a novel analytical mathematical method for the shape description without using measuring tools, e.g. rulers.

- Other mathematicians extended Kac's idea in exploring the geometrical information of more complex domains with various boundary conditions [11][8][7] [18][14].

2.1.4.2 Limitations

- It is only available for the convex domain, which has a smooth or piecewise smooth boundary.

- Except in very few cases (i.e. rectangular, disk, certain triangles), the complete sequence of eigenvalues λ_k can not be calculated [9].

- Only the first few terms in the asymptotic expansion of $h(t)$ are explicitly available.

LRWs in Artificial Images

	test_statistic	p
Logrank	137.23	0.0
Tarone-Ware	134.31	0.0
Gehan-Breslow	123.83	0.0
Fleming-Harrington	123.83	0.0

Table 3.1: Survival functions for circle and rectangle are statistically different since p values equal zeros.

The fixed-time step Monte Carlo simulation, LRWs, has been validated in the annulus in Appendix B by comparing the analytical and numerical survival function. The further validation of LRWs is to distinguish the geometries and explore their structural features from both short and long time survival behaviours.

3.1   Circle and Rectangle ^{Yuge}

Given two simple convex shapes with the same area, circle and rectangle, we are interested in how and whether their corresponding survival curves differ from each other. For the equal-area geometries, rectangle and circle, Eq. 2.2 indicates that the survival function of the former decays faster than the latter as the time approaches zero.

The preliminary step of testing the research hypothesis is to generate two black-and-white images with the same dimensions as shown in Fig. C.1. In the binary images, circle and rectangle have an equal number of white pixels. For simplicity, the centroid of shapes located at the center of the image. Then, simulating LRWs in the images and estimating survival functions by Kaplan-Meier estimator.

3.1.1   Output Analysis

The differences between survival functions for the circle and rectangle are not visible. Moreover, the approximate 95% confidence intervals of the survival functions overlap. In this case, non-parametric statistical tests can be used to compare entire survival distributions and assess their dissimilarities. The logrank test has maximum power if the proportional hazards assumption is satisfied.

3.1.2   Conclusion

Although the proportional hazard assumption test is failed as shown in Fig. **??**, the weighted logrank tests indicate that the null hypothesis should be rejected. In conclusion, LRWs is an alternative tool to quantify and distinguish the geometries in the $2-$ dimensional image without measuring the predefined shape descriptors.
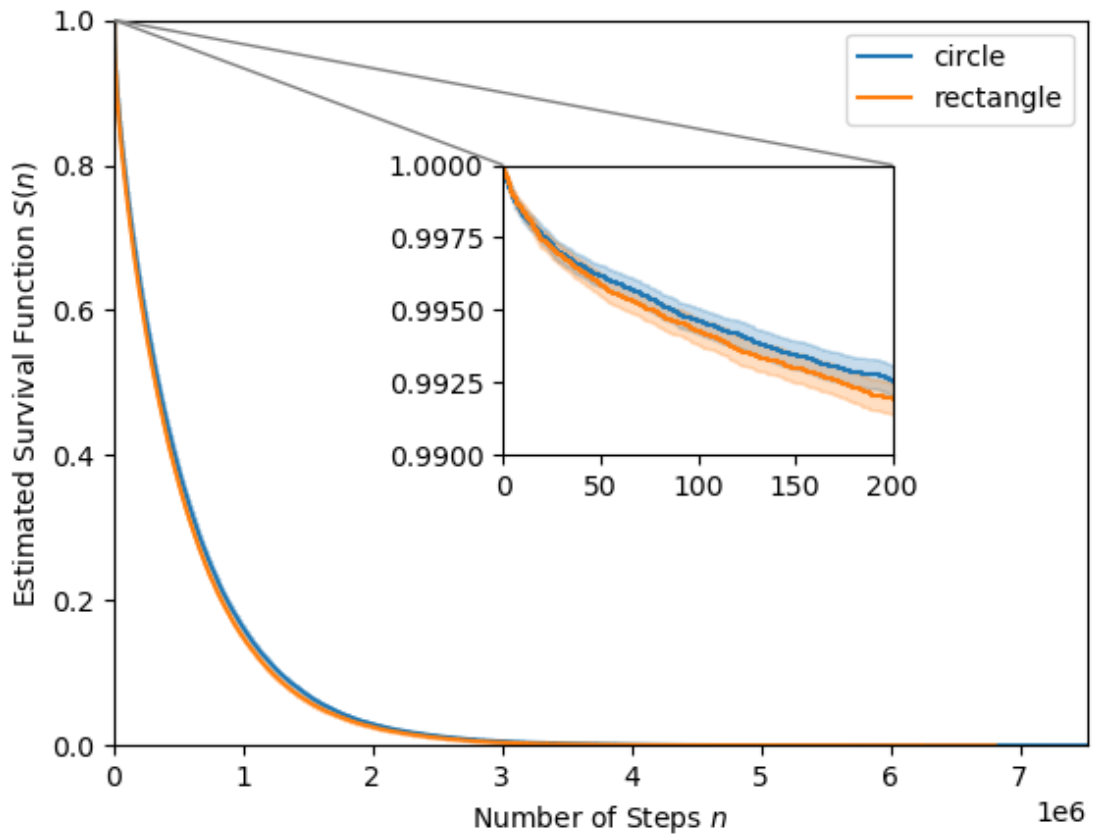
Figure 3.1: In the inset, the decay rate of the survival function for the rectangle is slightly larger than for the circle, which coincides with the theoretical result.
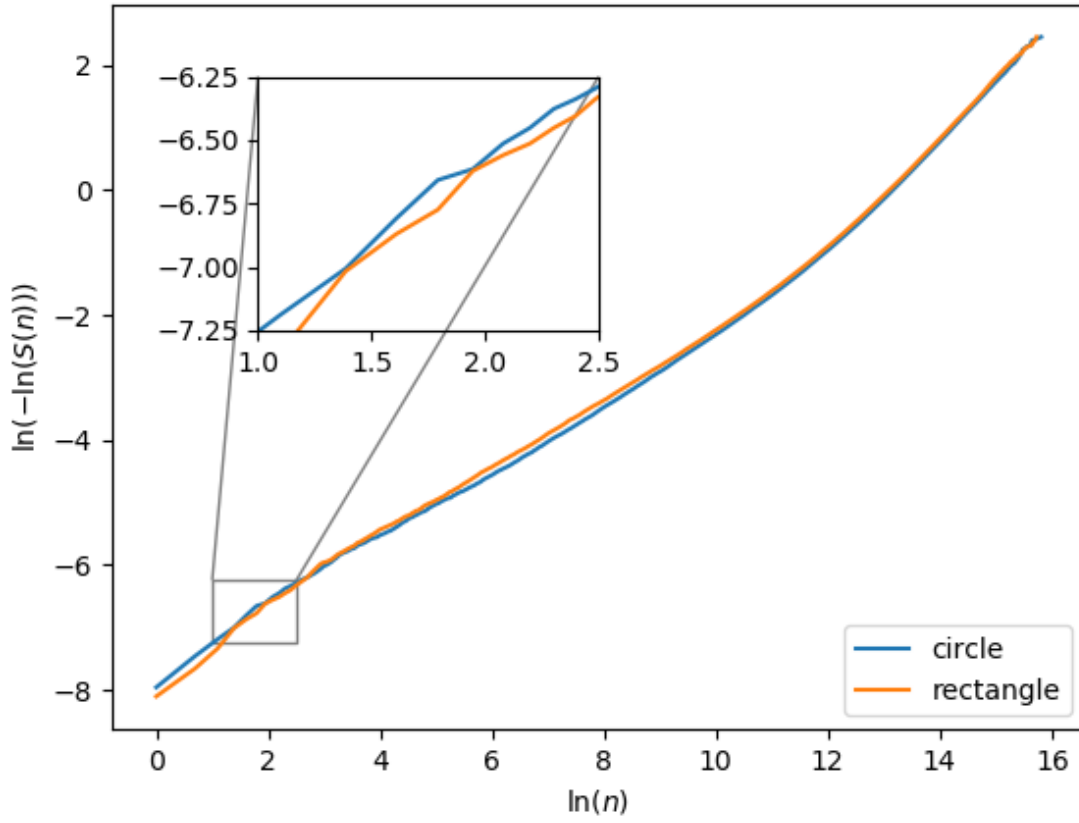
Figure 3.2: It is a graphical method for checking proportionality by looking for parallelism. As shown in the inset plot, two curves cross at some points and their shapes vary over time. Moreover, $p < 0.05$ in the non-proportional test. Thus, the survival functions for circle and rectangle do not satisfy the proportional hazard assumption.

3.2 Complicated Branching Structures

In the preceding section, the research hypothesis has been tested by the comparison of simple shapes. More complex branching structures, as shown in Fig. C.2 and Fig. C.3, are produced in this section for further validation of the research assumption and understanding of the survival curve. The branching structures in the images are equal-area and vertically symmetric, but the template in G_1 is shorter and narrower than in G_2. In G_i, $i = 1, 2$, the larger number of iterations j, the more complicated structures L_j, $j = 3, 4, 5, 6$, with more nodes.
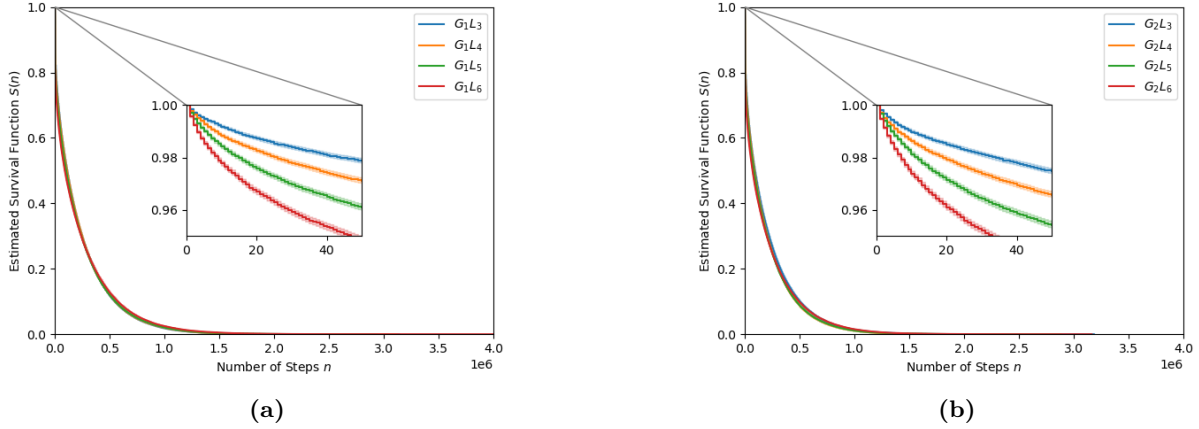
Figure 3.3: (a) and (b) are survival functions for branching structures in G_1 and G_2, respectively. n is the number of steps taken by the particle from the initial to the stop pixel in LRWs.

3.2.1 Output Analysis of $S(n)$ <mark>Yuge</mark>

The insets in Fig. 3.3 show that, within a group, the decay rates of $S(n)$ for L_j, $j = 3, ..., 6$, are significantly distinct as n approaches 0. The graphical representation of short-time behaviours of survival functions is consistent with the analytical result since bigger j results in the larger perimeter of the branching structure.

As shown in Fig. 3.4, the survival curve is divided into several coloured segments. In Fig. 3.5, Fig. 3.6, Fig. 3.7, Fig. 3.7, and Fig. 3.9, particles' initial and stop positions are painted with colour identical to the survival curve segment to understand the underlying stochastic process and its properties. Furthermore, the black region in the initial position plot is the target branching structure.

In reality, LRWs is latent and cannot be observed directly. However, those scatter plots are helpful to reveal the first-passage properties of particles in LRWs. For instance, particles will be absorbed within a smaller number of steps if their starting positions are closer to the target because they have less randomness. Moreover, more particles will be generated in the broader and longer in-between area of branches. Therefore, each segment of the survival curve carries massive geometric and spatial information about the unoccupied area of the binary image (i.e. the region with black pixels) and the boundary of the simply connected domain (i.e. the artificial branching structure with white pixels).

Some weighted log-rank tests can be utilized to detect the early or late differences between the pairwise overlapping or crossing survival curves. In Table 3.2 and Table 3.3, TW is the abbreviation for Tarone-Ware test, GB is for Gehan-Breslow test, and FH is for Fleming-Harrington test. However, p values in the tables are not too informative, and the log-rank test under conditions of non-proportional hazards leads to misleading results. Hence, distance measures are alternative methodologies for quantifying the discrepancy between survival functions.

It is assumed that $\widehat{S}_1(t)$ and $\widehat{S}_2(t)$ are the Kaplan-Meier estimators of the survival functions for random variable $T_1 > 0$ and $T_2 > 0$, respectively. Let $(\tau_j)_{j=1,2,...,N}$ are distinct increasing observed times when the
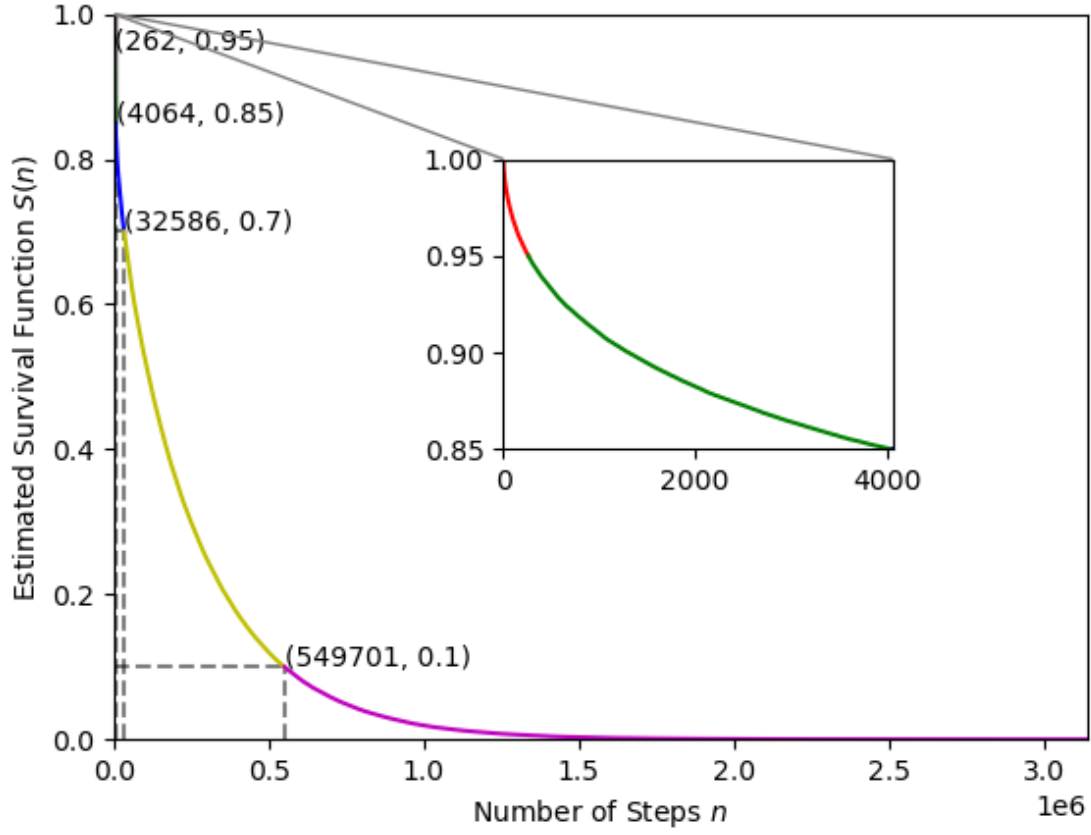
Figure 3.4: It is an estimated survival function for LRWs in $G_1 L_3$.

		p			
		Log-rank	TW	GB	FH
$G_1\ L_3$	$G_1\ L_4$	0.4393	0.0285	0.0005	0.0005
	$G_1\ L_5$	0.0	0.0	0.0	0.0
	$G_1\ L_6$	0.0	0.0	0.0	0.0
$G_1\ L_4$	$G_1\ L_5$	0.0007	0.0	0.0	0.0
	$G_1\ L_6$	0.0002	0.0	0.0	0.0
$G_1\ L_5$	$G_1\ L_6$	0.7223	0.0	0.0	0.0

Table 3.2: The differences between the pairwise survival functions for branching objects in G_1 are statistically significant based on TW, GB, and FH tests.
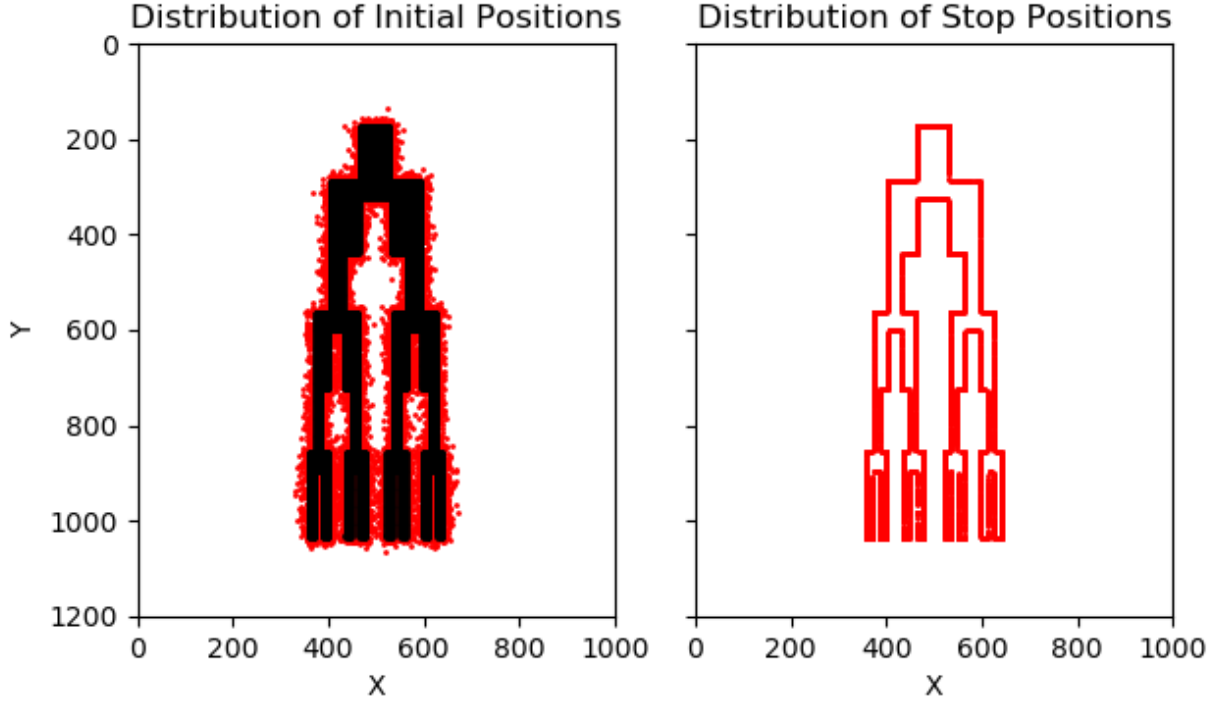
Figure 3.5: 5% of particles in the LRWs coloured by red will be absorbed within 262 steps. In the left subfigure, particles' initial positions are distributed closely surrounding the target branching structure. The in-between space of tiny bottom limbs is filled with particles, while some area between the huge top branches is empty. The right subfigure shows that the red particles characterize the entire boundary of the object.

		p			
		Log-rank	TW	GB	FH
$G_2\ L_3$	$G_2\ L_4$	0.0	0.0	0.0	0.0
	$G_2\ L_5$	0.0	0.0	0.0	0.0
	$G_2\ L_6$	0.0	0.0	0.0	0.0
$G_2\ L_4$	$G_2\ L_5$	0.0016	0.0	0.0	0.0
	$G_2\ L_6$	0.0004	0.0	0.0	0.0
$G_2\ L_5$	$G_2\ L_6$	0.7199	0.0	0.0	0.0

Table 3.3: It is well known that the log-rank test will lose power if the proportional hazard assumption is violated. Except for the log-rank test, other statistical tests indicate that the pairwise survival functions for branching objects in G_2 are statistically different.
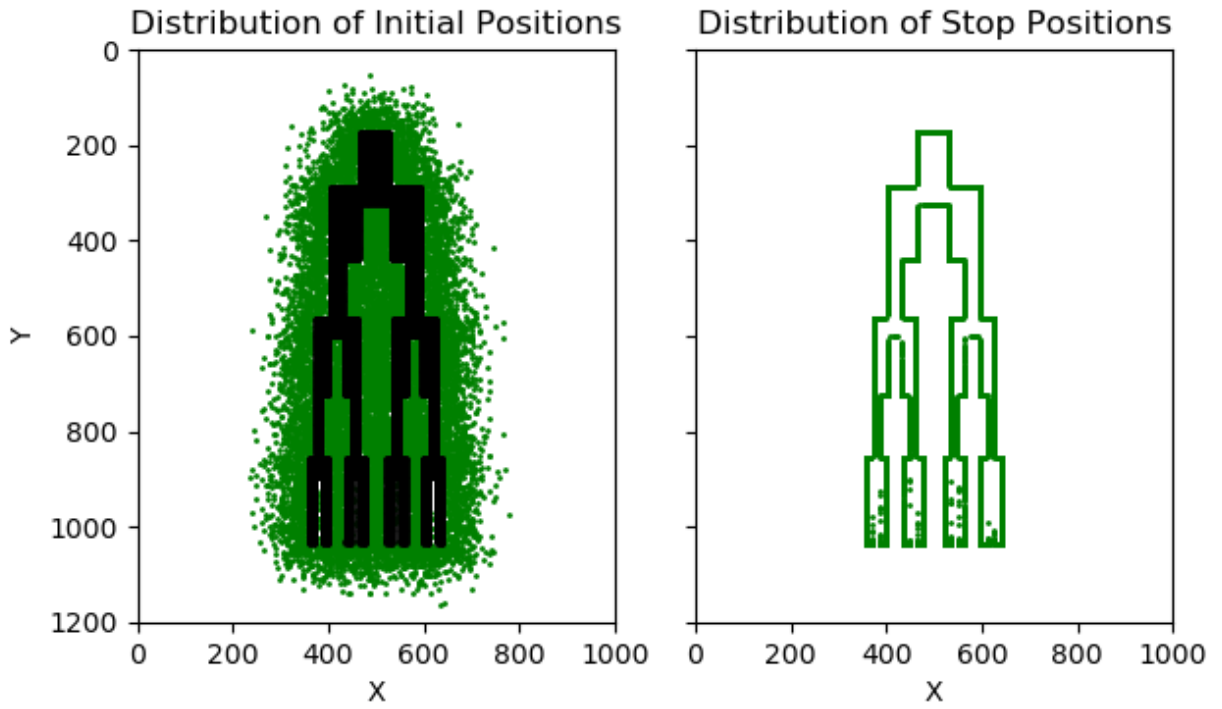
Figure 3.6: Around 10 percent of particles are generated adjacent to the fringe of the object, whose steps ranged from 262 to 4064. Compared with Fig. 3.5, the green points also dispersed thoroughly between the branches. Nevertheless, some in-between regions of the lowest branches are empty, resulting in the object's incompleted edge in the right plot.
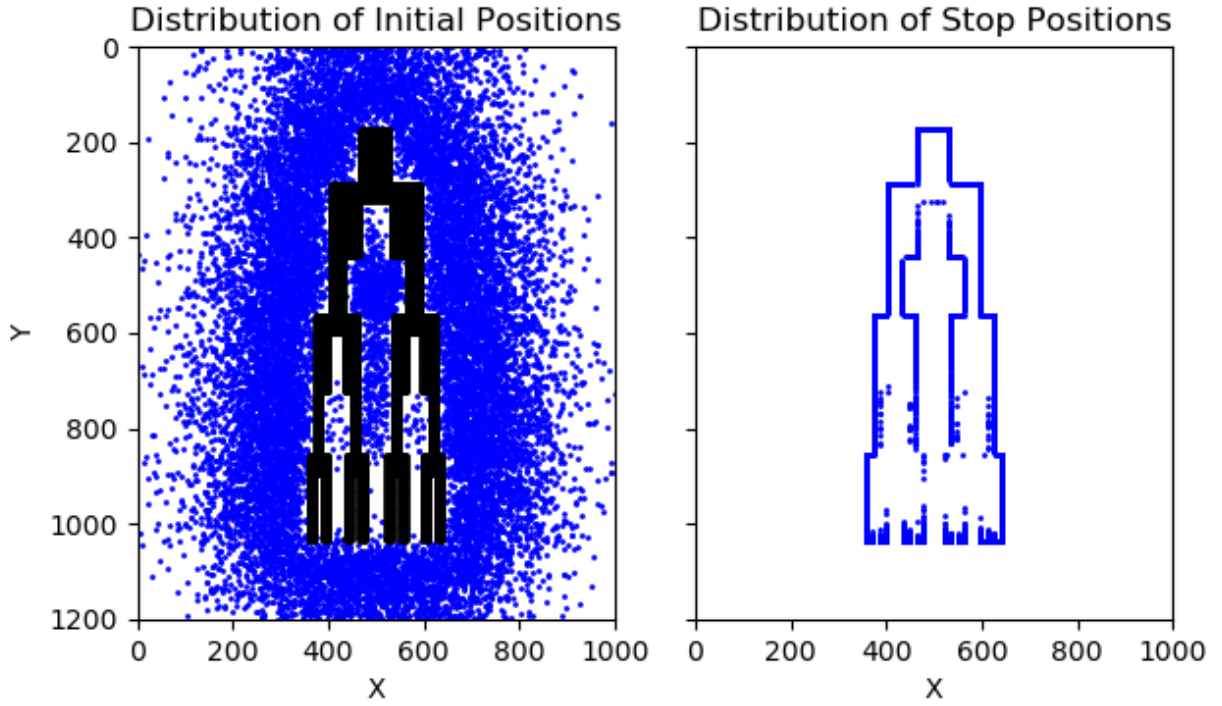
Figure 3.7: Approximate 15 percent of particles originally started LRWs from the region slightly far away from the target object. The number of steps for the blue particles is from 4064 to 32586. Not like the Fig. 3.6, fewer blue particles are located initially between the narrower branches. Moreover, if particles' initial sites are further away from the branching structure's external boundary, they will be more dispersive. The particles in the right subfigure can depict only some top pieces of the internal border.
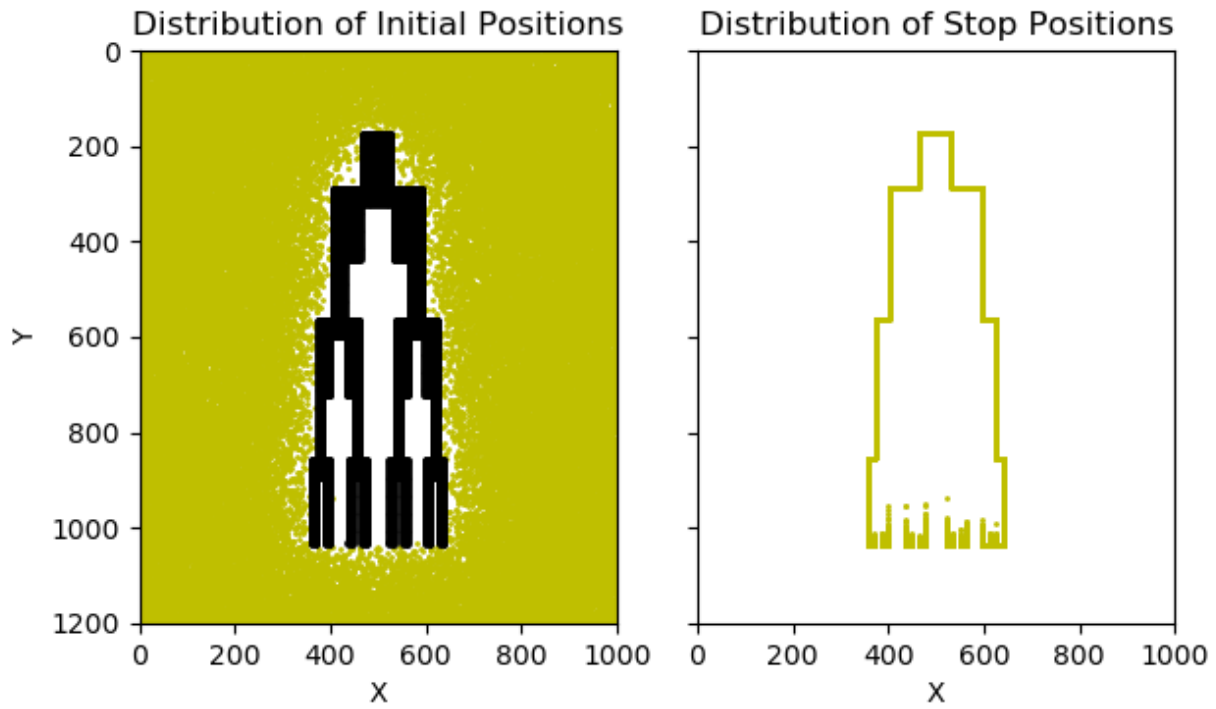
Figure 3.8: The left subplot displays the initial positions of 60% of particles in LRWs distributed uniformly in the empty space of the image but slightly far away from the target object. Moreover, none of them initially started LRWs from the in-between region of the branches. Compared with the Fig. 3.7, the right subplot shows the object's entire external boundary and some internal boundary of its terminal limbs.
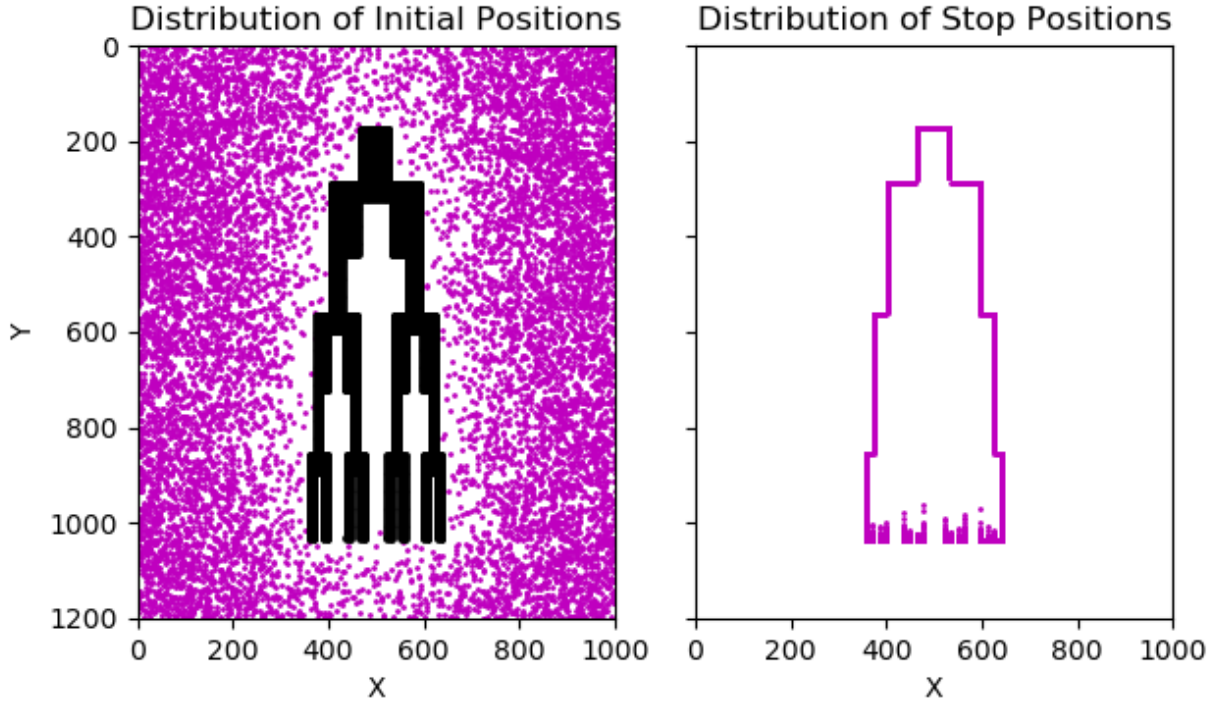
Figure 3.9: Only 10 percent of particles will survive when their number of steps is bigger than 549701. They are distributed initially from a region far away from the object and close to the image's edges. Similar to Fig. 3.8, pink particles cannot delineate too much internal boundary of the branching structure.
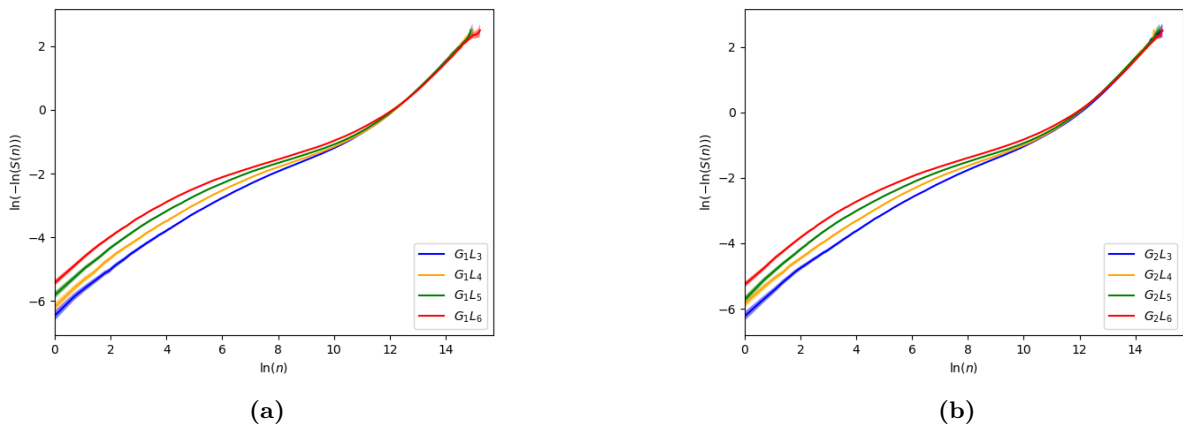

Figure 3.10: (a) and (b) are commonly used graphical techniques to check the proportional hazards (PH) assumption of survival data by finding the parallelism. The survival distributions do not support the PH assumption since the hazard ratio in both G_1 and G_2 is not always constant.
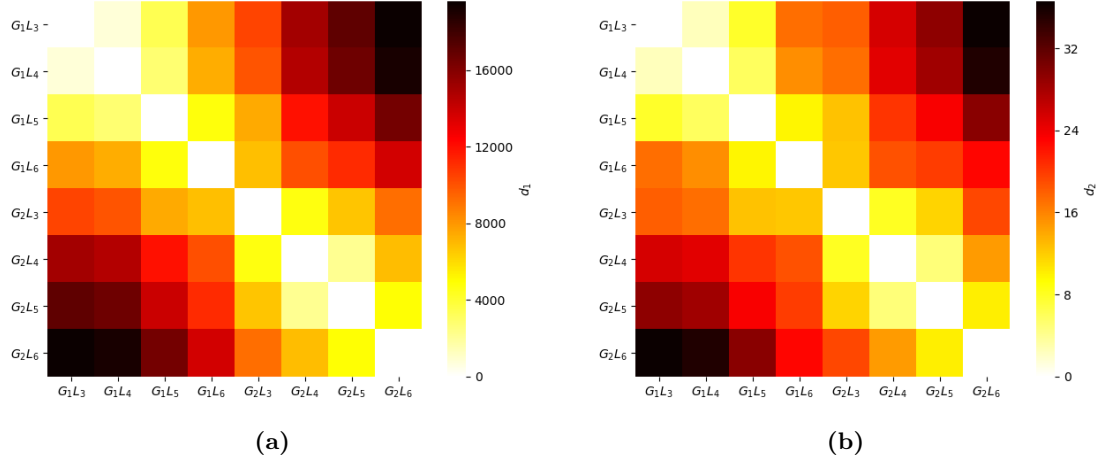
Figure 3.11: The distance matrices of pairwise survival functions for artificial images are visualized by heat maps.

event of interest take place. In the vector space, the distance between $\widehat{S}_1(t)$ and $\widehat{S}_2(t)$ can be defined as the L_p norm of their difference, $1 \leq p \leq \infty$, where

$$
\begin{aligned}
d_p(\widehat{S}_1(t), \widehat{S}_2(t)) &= \|\widehat{S}_1(t) - \widehat{S}_2(t)\|_p \\
&= \Big(\sum_{j=1}^{N} |\widehat{S}_1(\tau_j) - \widehat{S}_2(\tau_j)|^p \Big)^{\frac{1}{p}} \\
&= \Big(\sum_{j=1}^{N} |P(T_1 > \tau_j) - P(T_2 > \tau_j)|^p \Big)^{\frac{1}{p}}
\end{aligned}
\tag{3.1}
$$

Let p be 1 and 2 in Eq. 3.1 respectively, the distance between a pair of survival functions is measured by d_1 and d_2, and is depicted by colors in the heat maps as shown in Fig. 3.11a and Fig. 3.11b. The cells on the main diagonal are white because the distance of an object from itself is zero. Moreover, the off-diagonal cells are symmetric. The darker cells in the heat maps demonstrate more significant dissimilarities between survival functions or curves. In other words, the colour of cells describes the variation in the shapes of branching structures in artificial images.

The top left and bottom right 4×4 cells are the in-group shape comparison of the branching structures. For each column, the cells become darker gradually since the objects in each group are more complicated as j increase. The pale yellow cells adjacent to the diagonal, including (G_1L_3, G_1L_4), (G_1L_4, G_1L_5), (G_1L_5, G_1L_6), (G_2L_3, G_2L_4), (G_2L_4, G_2L_5), and (G_2L_5, G_2L_6), imply that the morphological changes are not dramatically different. Moreover, the top right and bottom left 4×4 cells are the between-group shape comparison. The smaller dissimilarities are closer to the diagonal, which gives a form of clustering the artificial images.

Metric Multidimensional scaling (MDS) [2] is employed to display the distance matrix using a map defined in an abstract Cartesian space. Given a dimension p and a distance matrix $D = (d_{ij})$, MDS is used to find a configuration $\boldsymbol{x_1}, ..., \boldsymbol{x_n} \in \mathbb{R}^p$, which satisfies
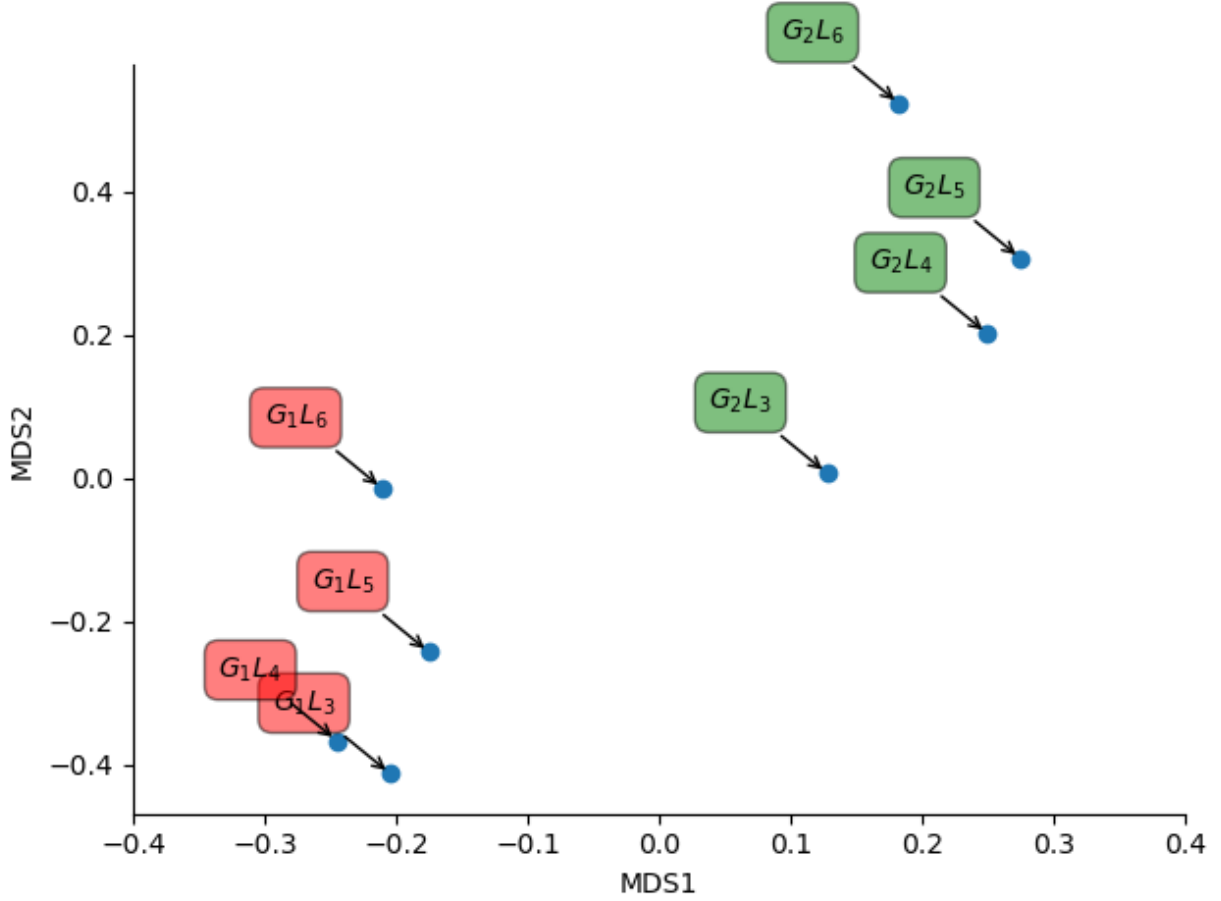
Figure 3.12: It is the MDS of the distance matrix calculated based on d_1, and the stress for the plot is about 0.82%. Points with green labels in the bottom left-hand corner illustrate the G_1 images, while others are G_2 images. It is easy to tell that there are two clusters for the artificial images, which coincide with reality and heat map in Fig. 3.11a.

$$f(d_{ij}) \approx \widehat{d}_{ij} = \|\boldsymbol{x_i} - \boldsymbol{x_j}\|_2 \tag{3.2}$$

as close as possible, and f could be a parametric monotonic function, such as $f(d_{ij}) = \alpha + \beta d_{ij}$.

Metric MDS aims to minimizes the stress denoted by $\mathcal{L}(\widehat{d}_{ij})$ over all $\widehat{d}_{ij}$, α, and β,

$$stress = \mathcal{L}(\widehat{d}_{ij}) = \left(\sum_{i<j} (\widehat{d}_{ij} - f(d_{ij}))^2 / \sum d_{ij}^2 \right)^{\frac{1}{2}} \tag{3.3}$$

In Fig. 3.12 and Fig. 3.13, MDS denotes the artificial images as points in $2-$ dimensional geometric space so that distances between pairs of points match as well as possible the original dissimilarities calculated between the pairwise survival functions. Moreover, the axes in themselves are meaningless. For metric scaling [13], a smaller stress $\mathcal{L}(\widehat{d}_{ij})$ indicates a better fit of the configuration. Generally, a value of the stress of 20% suggests a poor fit, 10% fair, 5% good and 2% an excellent fit.
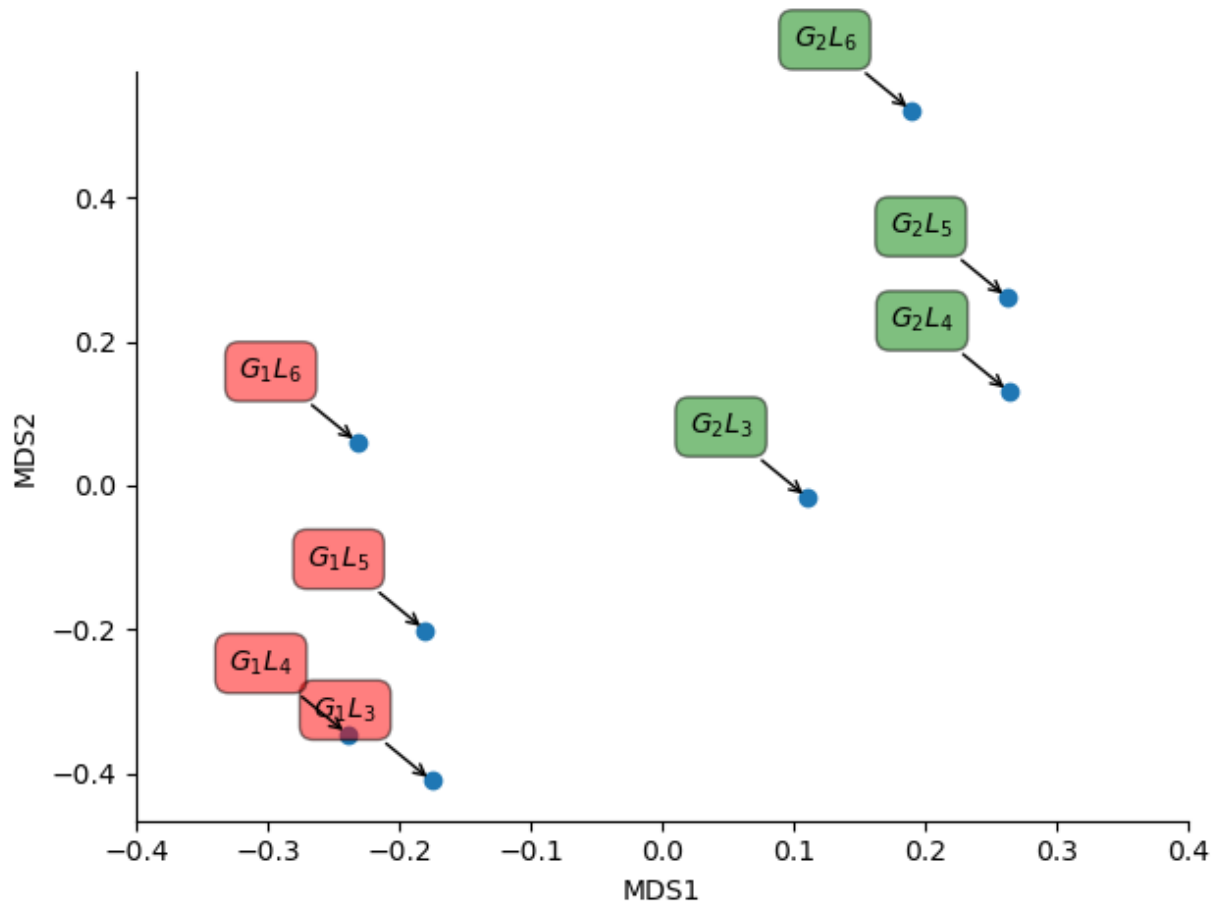
Figure 3.13: It is the MDS plot of the distance matrix calculated by d_2 with the stress 0.4%.
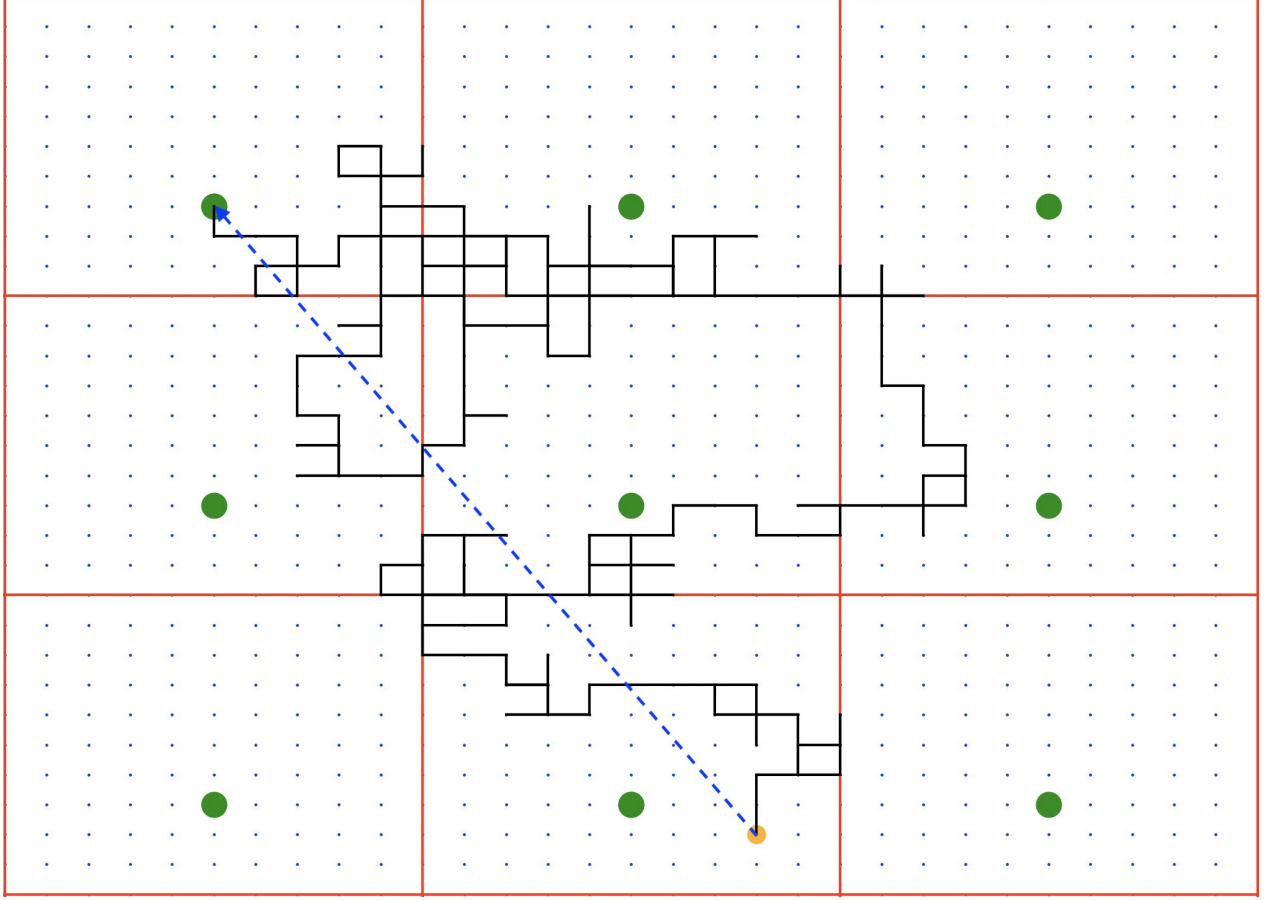
Figure 3.14: This plane is constructed by choosing a primitive cell with four red sides and a green point and replicating it infinitely to tile the whole 2− dimensional space. Moreover, there has no overlaps and voids between copies of the cell. A particle initially started LRWs from the orange site and be absorbed by any of the green points. To ensure smoothness and consistency, if the particle leaves the cell through one edge, it will appear in the adjacent cell with the same velocity. The black line segments show the particle's random trajectories, and the length of the blue dotted arrow is defined as its displacement.

3.2.2 Output Analysis of $S(d)$

3.2.2.1 Periodic Boundary Conditions

In this thesis, periodic boundary conditions (PBCs) are employed to minimize the influence of images' edges. Fig. 3.14 is a simplest example of implementing PBCs in the Euclidean plane E^2 and tracks the trajectory of a particle undergoing LRWs.

3.2.2.2 Relationship between n and d

In this section, the displacement of a particle, d, is the shortest distance from the initial to the stop position in the infinite tiling plane. In theory, the mean square displacement (MSD) of N Brownian particles at $n-$th step in 2−dimensional space is defined as
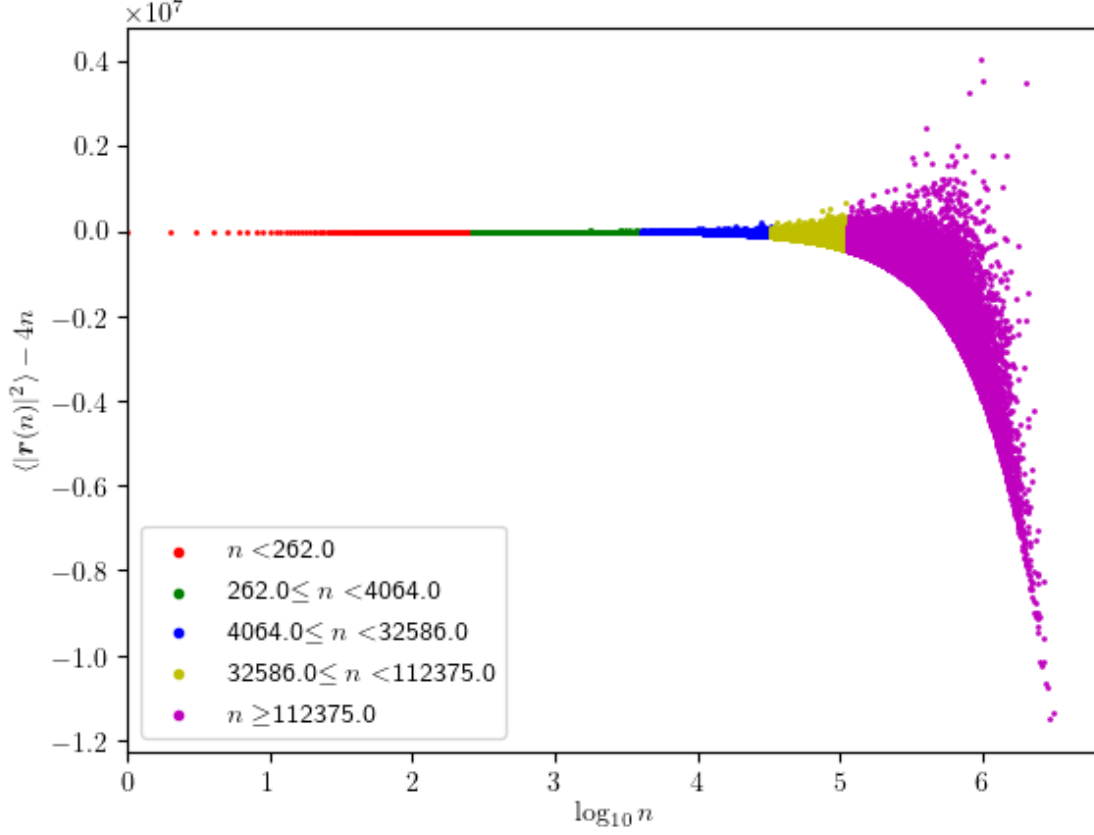
Figure 3.15: Particles in G_1L_3 are divided into several subgroups based on various intervals of steps, and their colours are as identical as the segments in Fig. 3.4.

$$MSD = \langle |\boldsymbol{r}(n)|^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{s}_i(n) - \boldsymbol{s}_i(0))^2 = 4Dn \tag{3.4}$$

where the subscript, i, refers to each particle for which the MSD is calculated. $\boldsymbol{s}_i(n)$ and $\boldsymbol{s}_i(0)$ are the $i-$th particle positions at $n-$th step and at the initial time, respectively. D is diffusion coefficient which is related to the variance of the independent displacements of the particle. In the simulation, D equals 1.

Eq. 3.4 indicates a linear relationship between the mean square displacement of the particle and the number of steps. It is a feature of the normal diffusive behavior. Fig. 3.15 shows how the difference between MSD of the particle and $4n$ varying over $\log_{10} n$ in LRWs. When $n \leq 4064.0$, the variation is not equal to 0 with larger fluctuation, which implies that blue, yellow, and pink particles undergo anomalous diffusion process. In other words,

$$\langle |\boldsymbol{r}(n)|^2 \rangle \propto n^\gamma \tag{3.5}$$

where $\gamma \neq 1$. In Fig. 3.4, negative variation implies $\gamma < 1$ called sub-diffusion process, while positive value
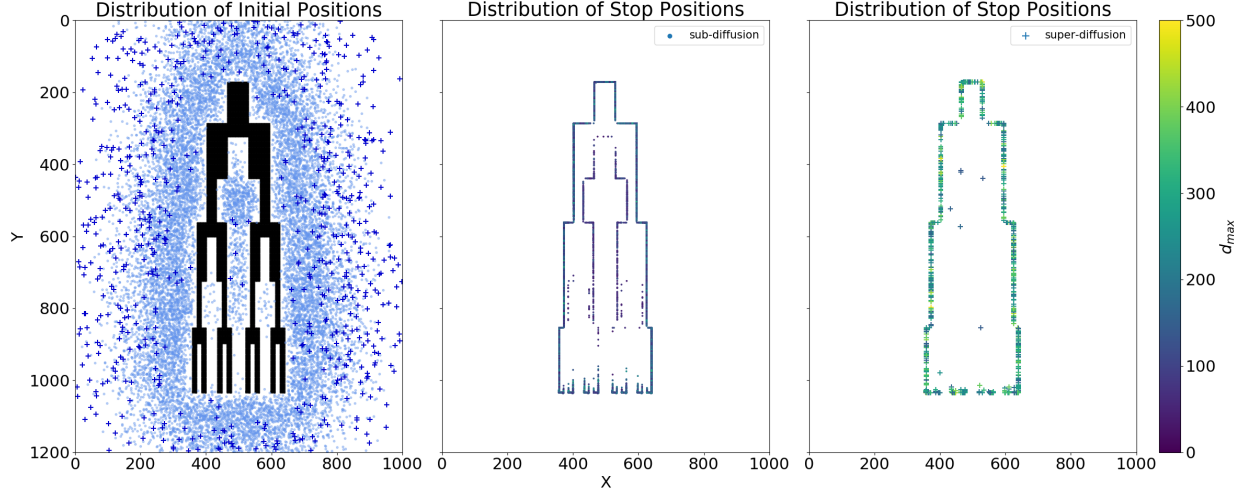
Figure 3.16: In the left figure, 696 dark blue pluses refer to super-diffusion particles, which are distributed closer and more concentrated to the fringe of the branching structure. 15198 pale blue points represent sub-diffusion particles and scatter mainly around the edges of the image. In-between the branches, there has only six super-diffusion particles and an enormous amount of sub-diffusion ones. The middle and right depict the stop positions for super and sub-diffusion particles, respectively, coloured by a perceptually uniform sequential colormap based on their maximum displacement in the tiling space.

denotes $\gamma > 1$ named super-diffusion.

To understand the underlying mechanism of the anomalous-type diffusion process, initial and stop positions of particles in Fig. 3.7, whose steps ranged from 4064 to 32586, are illustrated in Fig. 3.16. Suppose particles are trapped in the narrow space in-between the branches or initially start LRWs near the branching structure. In that case, their movements will be restricted because of the nearby absorbing boundary condition, causing the sub-diffusion phenomenon. As shown in the middle subplot of Fig. 3.16, the maximum displacement of sub-diffusion particles is less than 200. There also have exceptional circumstances that 6 particles, in-between the limbs, undergo super-diffusion since they can explore a large portion of space within a predefined range of steps, as shown in the right subfigure of Fig. 3.16. Generally, particles around the edges of the image will be more likely to pass through the periodic boundary, reappear in the adjacent cell, and continue LRWs with the same velocity until hitting the absorbing boundary, which results in large displacement and the super-diffusion process.

3.2.2.3 Estimated Survival Functions

Similar to Fig. 3.3, it is arduous to detect the variation among the survival curves by eyes. However, their short-term behaviours can be enlarged in insets in Fig. 3.17, and we can conclude that the bigger L_i, $i = 3, ..., 6$, leads to faster decay of survival function. The survival curve is analyzed segment by segment based on particles' initial and stop positions, shown in Fig. 3.19, Fig. 3.20, Fig. 3.21, Fig. 3.22, and Fig. 3.23, for the further understanding.
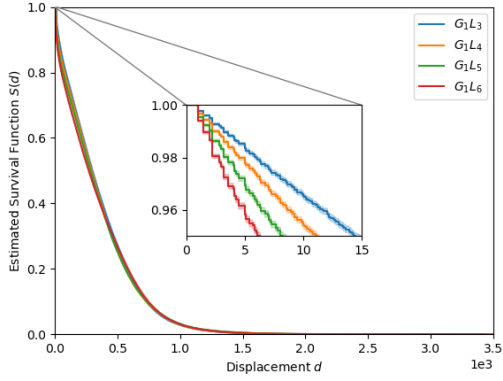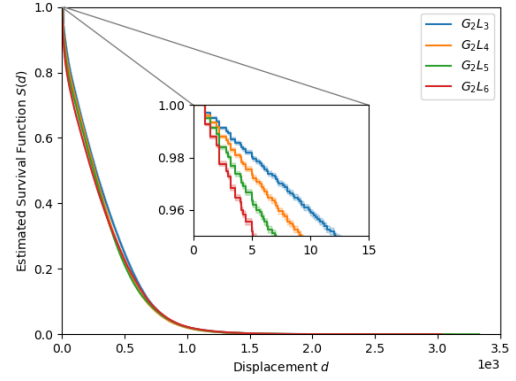
Figure 3.17: (a) and (b) are the estimated survival functions associated with particles' displacement in LRWs in G_1 and G_2, respectively.

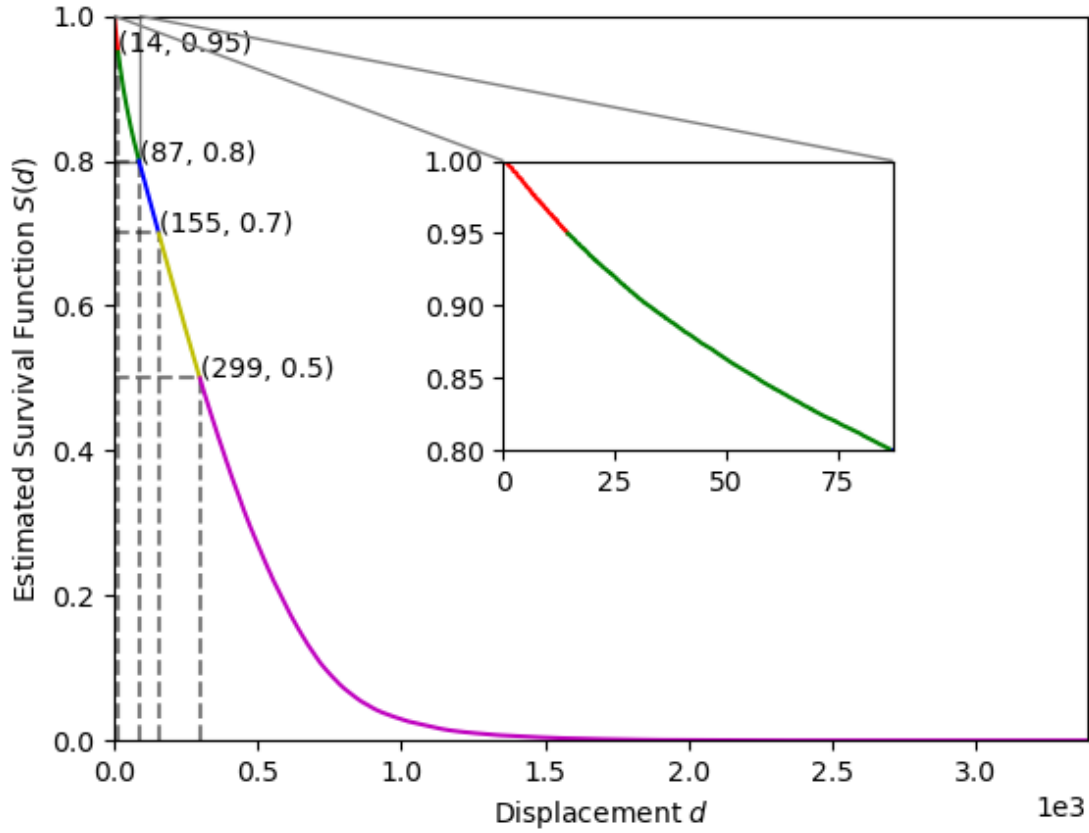**Figure 3.18:** The estimated survival function is separated into five coloured segments based on various intervals of displacement.
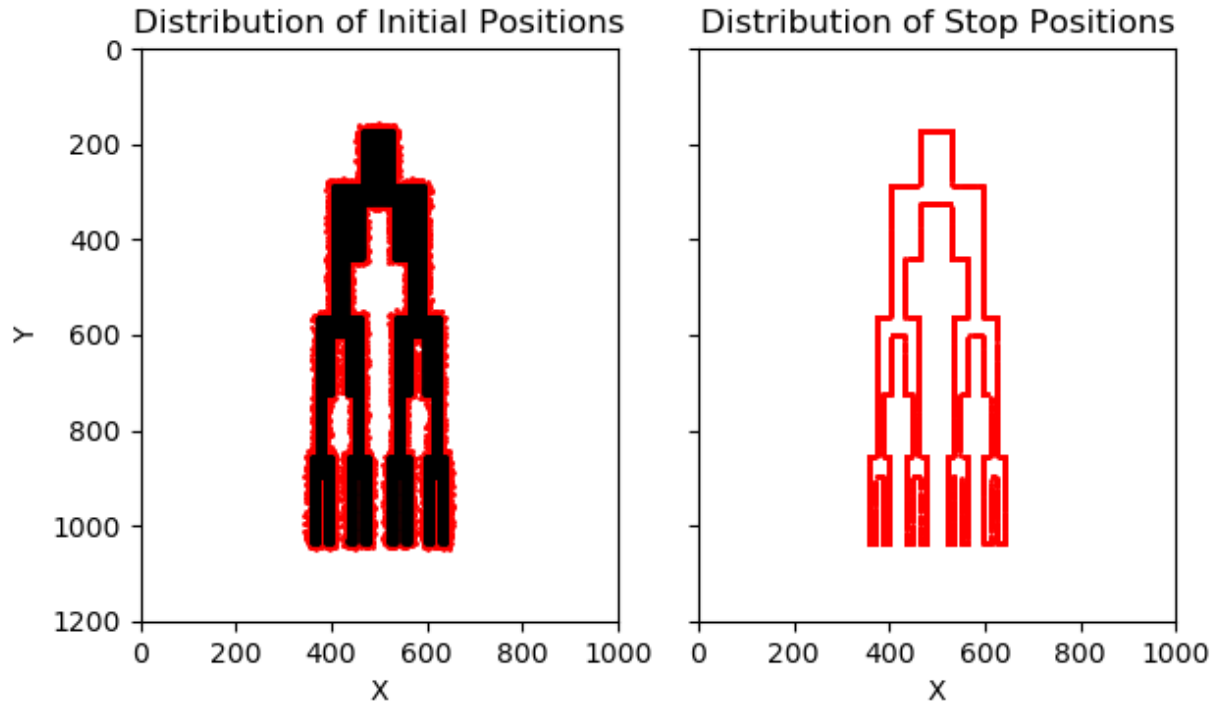
Figure 3.19

		p			
		Logrank	TW	GB	FH
$G_1\ L_3$	$G_1\ L_4$	0.0	0.0	0.0	0.0
	$G_1\ L_5$	0.0	0.0	0.0	0.0
	$G_1\ L_6$	0.0	0.0	0.0	0.0
$G_1\ L_4$	$G_1\ L_5$	0.0072	0.0	0.0	0.0
	$G_1\ L_6$	0.0003	0.0	0.0	0.0
$G_1\ L_5$	$G_1\ L_6$	0.2883	0.0	0.0	0.0

Table 3.4

Distribution of Initial Positions   Distribution of Stop Positions
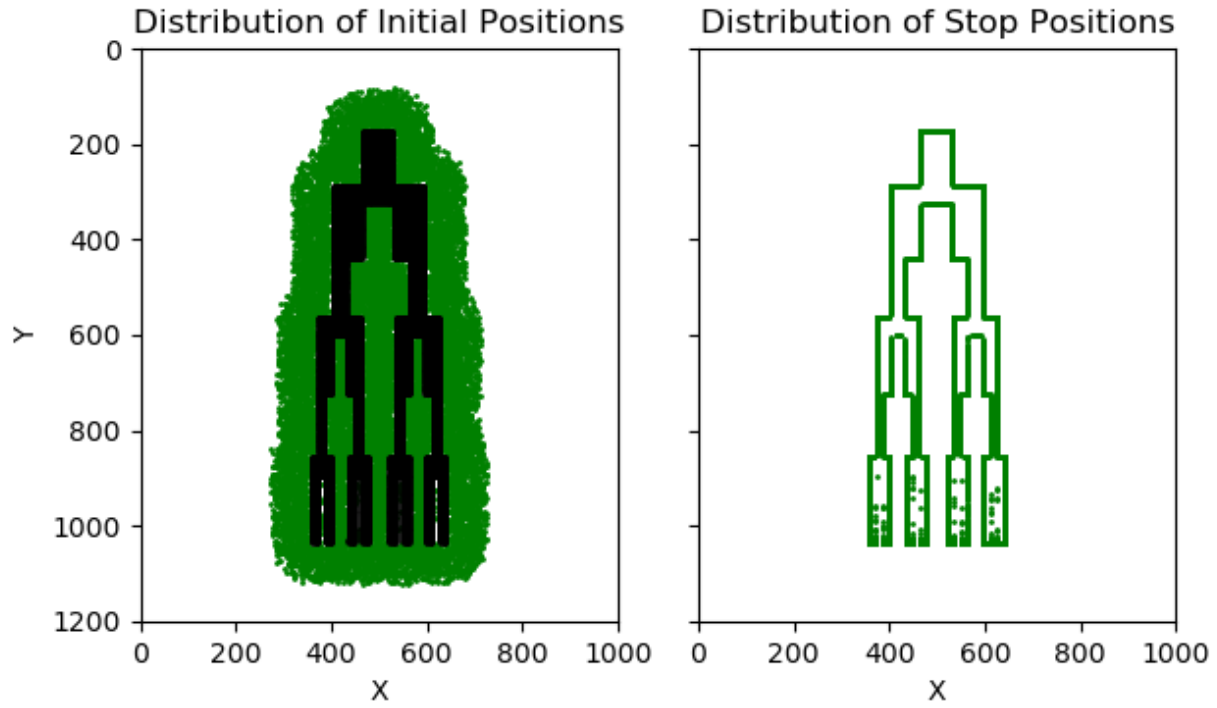
Figure 3.20

		p			
		Logrank	TW	GB	FH
$G_2\ L_3$	$G_2\ L_4$	0.0	0.0	0.0	0.0
	$G_2\ L_5$	0.0	0.0	0.0	0.0
	$G_2\ L_6$	0.0	0.0	0.0	0.0
$G_2\ L_4$	$G_2\ L_5$	0.0001	0.0	0.0	0.0
	$G_2\ L_6$	0.0015	0.0	0.0	0.0
$G_2\ L_5$	$G_2\ L_6$	0.7019	0.0	0.0	0.0

Table 3.5
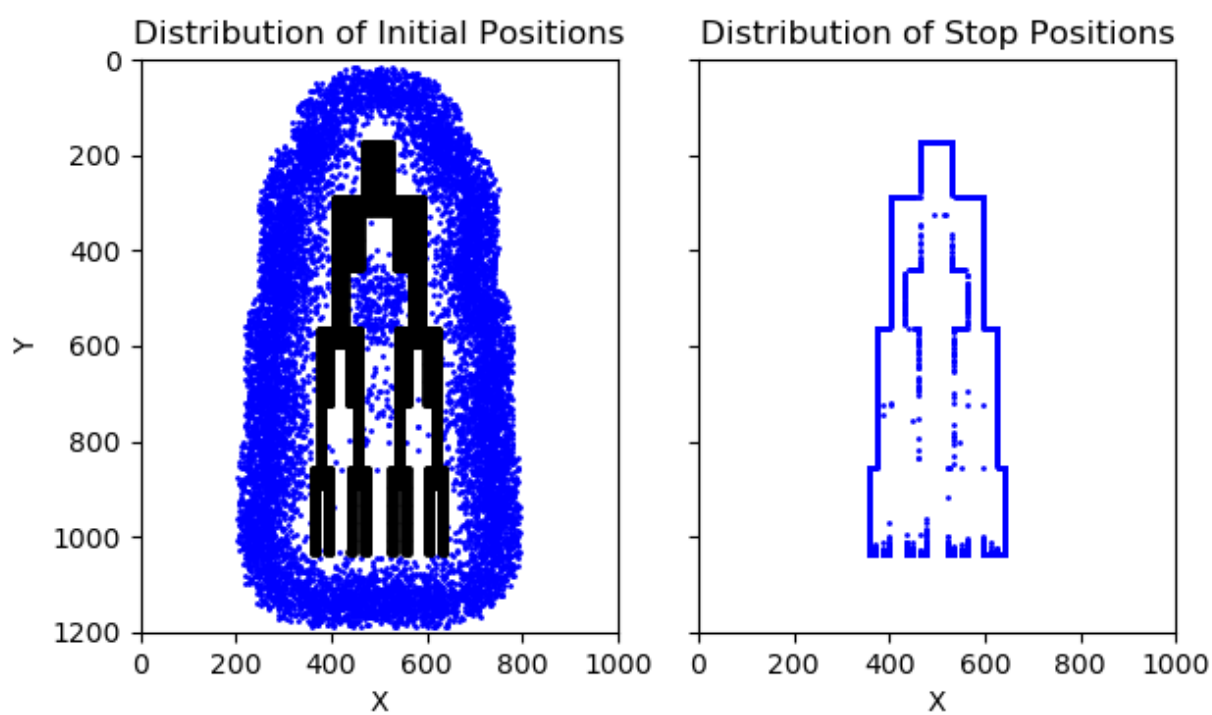
Figure 3.21
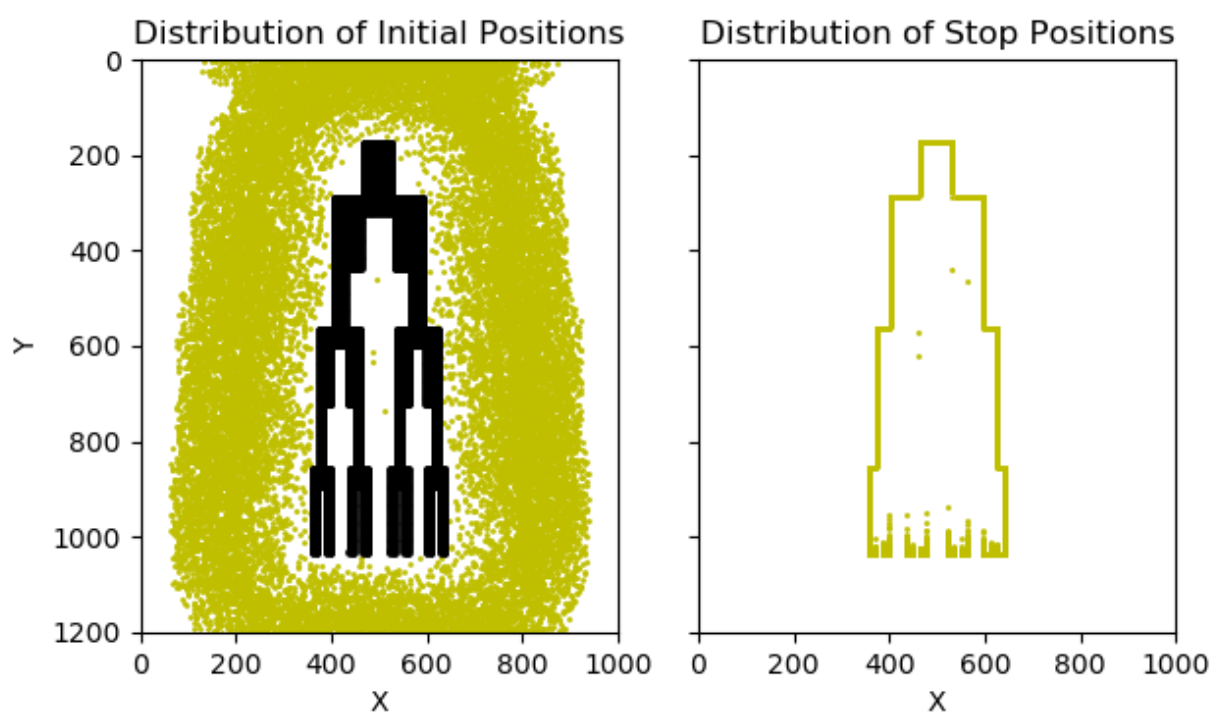
Figure 3.22

Figure 3.23

(a)   (b)

Figure 3.24

(a) **(b)**
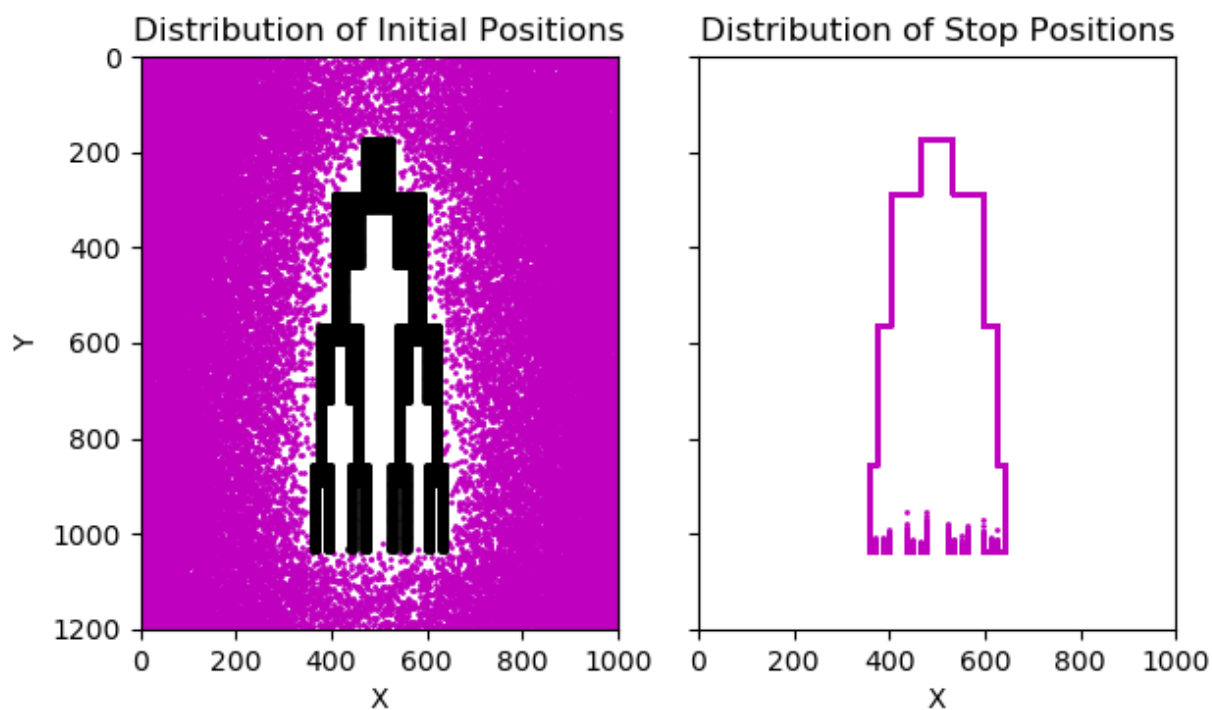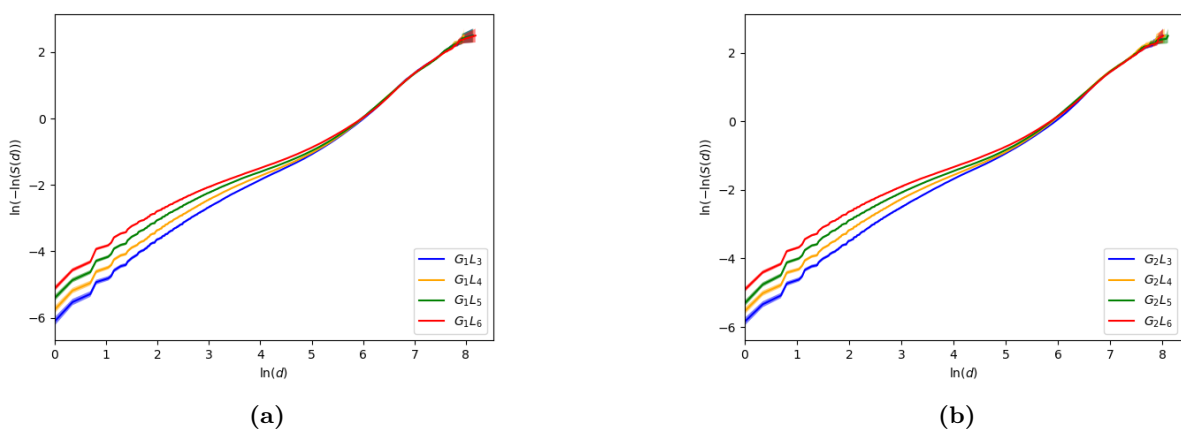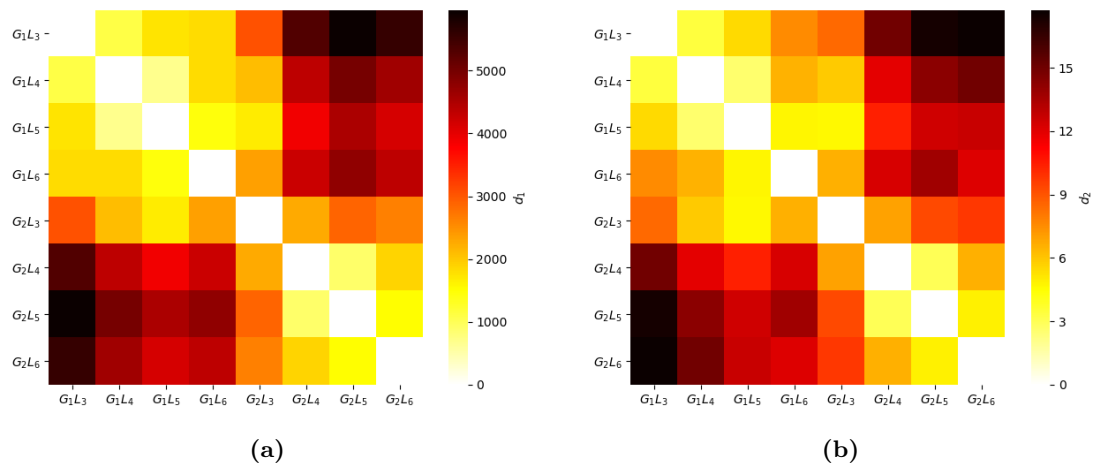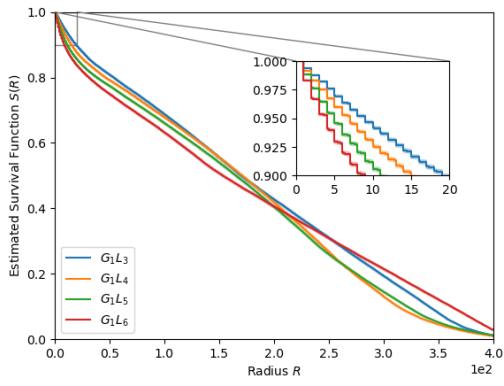
Figure 3.25

(a)
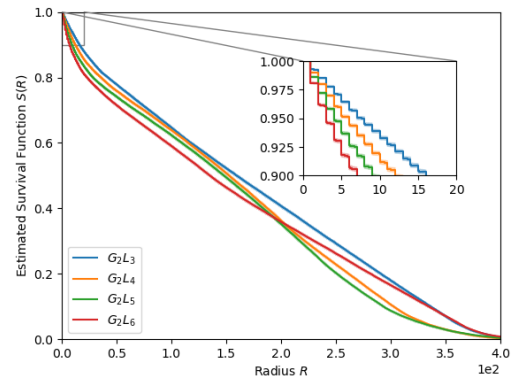
(b)

Figure 3.26

		p			
		Logrank	TW	GB	FH
$G_1\ L_3$	$G_1\ L_4$	0.0	0.0	0.0	0.0
	$G_1\ L_5$	0.0	0.0	0.0	0.0
	$G_1\ L_6$	0.0	0.0	0.0	0.0
$G_1\ L_4$	$G_1\ L_5$	0.1773	0.0	0.0	0.0
	$G_1\ L_6$	0.0	0.0	0.0	0.0
$G_1\ L_5$	$G_1\ L_6$	0.0	0.0	0.0	0.0

Table 3.6

3.2.3   Output Analysis of $S(R)$

		p			
		Logrank	TW	GB	FH
$G_2\ L_3$	$G_2\ L_4$	0.0	0.0	0.0	0.0
	$G_2\ L_5$	0.0	0.0	0.0	0.0
	$G_2\ L_6$	0.0	0.0	0.0	0.0
$G_2\ L_4$	$G_2\ L_5$	0.0	0.0	0.0	0.0
	$G_2\ L_6$	0.0	0.0	0.0	0.0
$G_2\ L_5$	$G_2\ L_6$	0.0	0.0	0.0253	0.0253

Table 3.7

3.3   Conclusion

- In a short time, the survival function of rectangle decays faster than the circle, which conforms to the analytical results.

- The differences of estimated survival functions between circle and rectangle are statistically significant, which coincides with the real shape dissimilarities.

- Within a same group, when t is small, the more branching the object is, the faster the survival function decays.

- Within a same group, the pairwise survival functions are statistically different.

- The corresponding target structures in G_1 and G_3 are invariant shapes under translation since their survival function are not statistically different. In other words, periodic boundary conditions of the image can eliminate the effect of the locations.

- LRWs can describe and classify the geometries, their spatial configurations, and the unoccupied area in the image.

LRWs in Real Root Images

Conclusion

Future Work

Appendix A

Numerical Methods for Solving Parabolic Partial Differential Equations

A.1 Introduction

- Parabolic PDEs: to characterize time-dependent phenomena

- The intrinsically similar features of the traditional computational techniques are mesh discretization in time and space.

A.2 Summary of Commonly Used Numerical Techniques

A.2.1 Finite Difference Method (FDM) [?]

A.2.2 Finite Element Method (FEM) [?]

A.2.3 Other Tranditional Computational Methods

A.3 Limitation in Practice

Appendix B

Method Validation in Annulus

B.1　Analytical Results

B.1.1　Shape Description

- Problem domain Ω: the region bounded by two concentric circles
- Radius of the larger circle: b
- Radius of the smaller circle: a

B.1.2　Solving Initial-Boundary Value Problem (IBVP)

B.1.2.1　Methods

- Dimensional Analysis: non-dimensional variables

    - $\mu = \frac{b}{a}$
    - $\tau = \frac{t}{a^2}$
    - $\hat{r} = \frac{r}{a}$

- Method of separation of variables

B.1.2.2　Mathematical Equations

- Diffusion equation

$$u_\tau = (u_{\hat{r}\hat{r}} + \frac{1}{\hat{r}} u_{\hat{r}} + \frac{1}{\hat{r}^2} u_{\theta\theta}) \tag{B.1}$$

- Uniform initial condition

$$u(\hat{r}, \theta, 0) = \frac{1}{|\Omega|} \tag{B.2}$$

- Homogenous Dirichlet B.C.

$$u(1, \theta, \tau) = 0 \tag{B.3}$$

- Homogenous Neumann B.C.

$$\hat{r}u'(\mu, \theta, \tau) = 0 \tag{B.4}$$

B.1.2.3　Heat Content Calculation

$$S(\tau) = \int_0^{2\pi} d\theta \int_1^\mu \hat{r} d\hat{r} u(\hat{r}, \theta, \tau) \tag{B.5}$$

B.2　Numerical Approximation

B.2.1　Eigenvalues $\lambda_{0,n}$

- Properties

    - $\lambda_{0,n} \in \mathbb{R}^+,\ (n \in \mathbb{N}_+)$
    - Monotonicity and Periodicity

- Estimation

    - $\lambda_{0,n} \in ((n-1)\pi, (n+1)\pi)$ [4]
    - Bisection method [17]

B.2.2 Approximation of $u(\hat{r}, \theta, \tau)$ and $S(\tau)$

- Direct summation

- Series acceleration methods

B.3 Comparison of Numerical and Analytical Results

B.3.1 Sample Size Evaluation

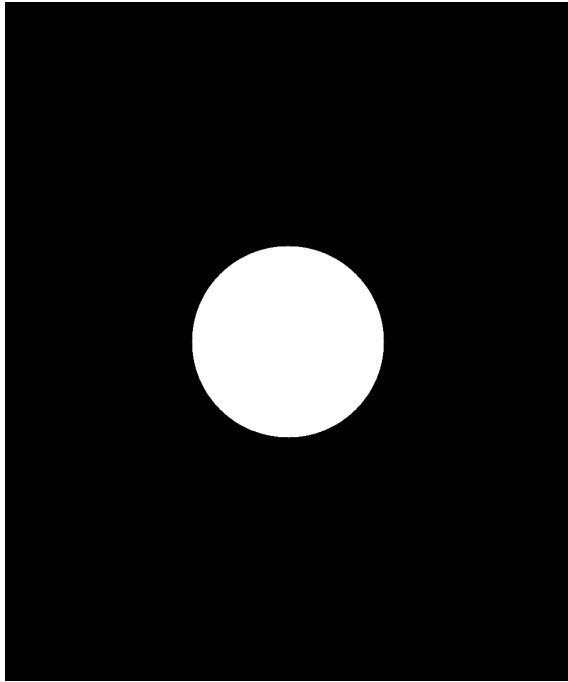### B.3.2 Comparison of $S(\tau)$ and $S(n)$

B.4 Conclusion

- The estimated survival function of LRWs is consistent with the analytical result.

- The number of particles in LRWs determined by DKW inequality is large enough to generate reproducible statistical results.
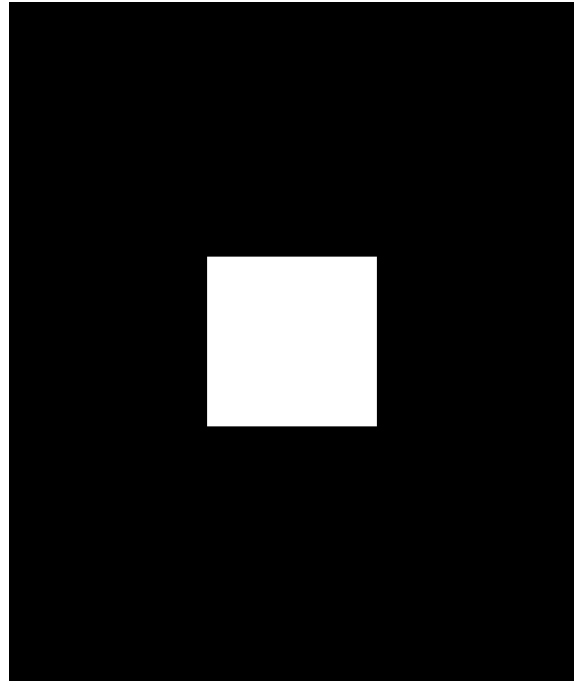
Appendix C

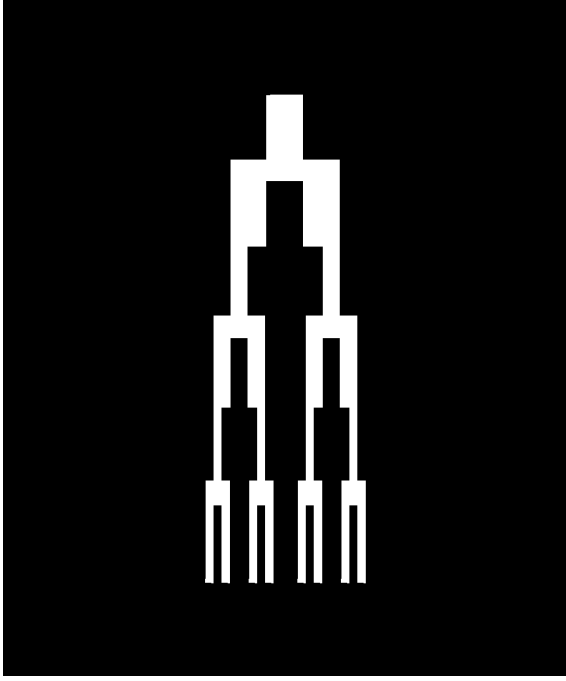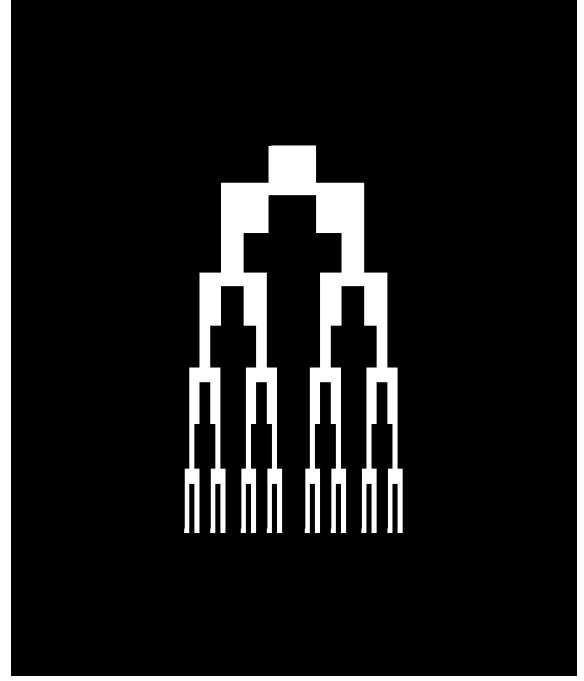# Artificial Images

C.1 Simple Shapes

(a) Circle

(b) Rectangle

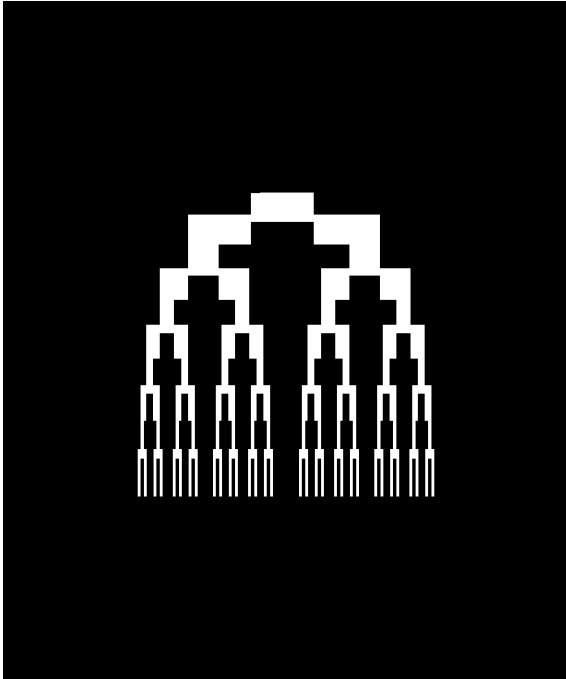**Figure C.1:** Each image size is 1200 by 1000 pixels with 90000 white pixels.

C.2 Complicated Branching Structures
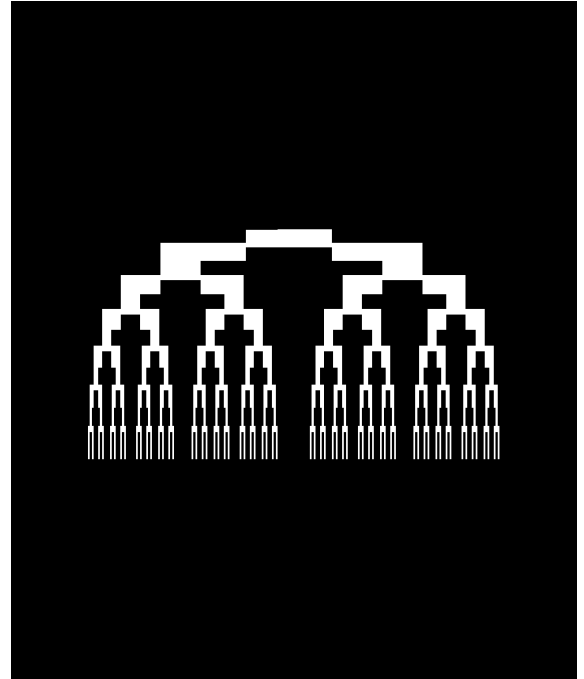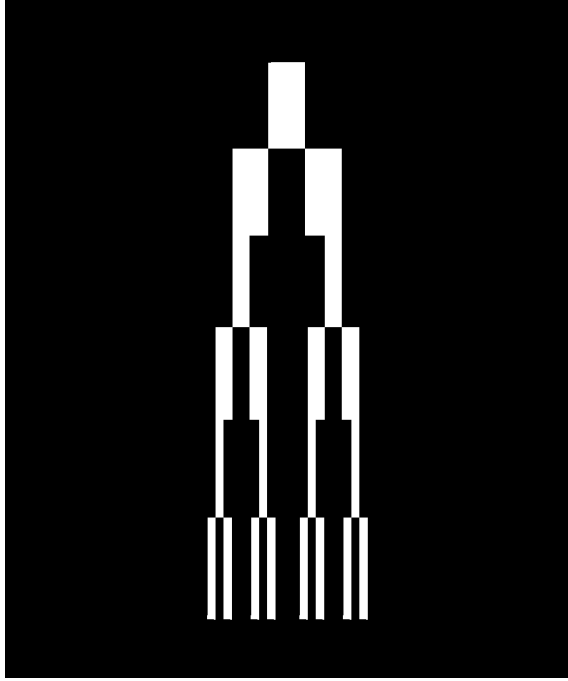
(a) G_1L_3
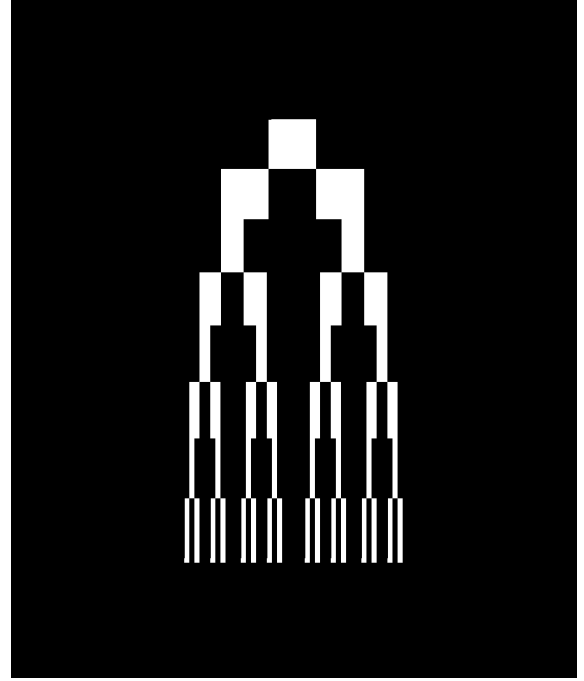

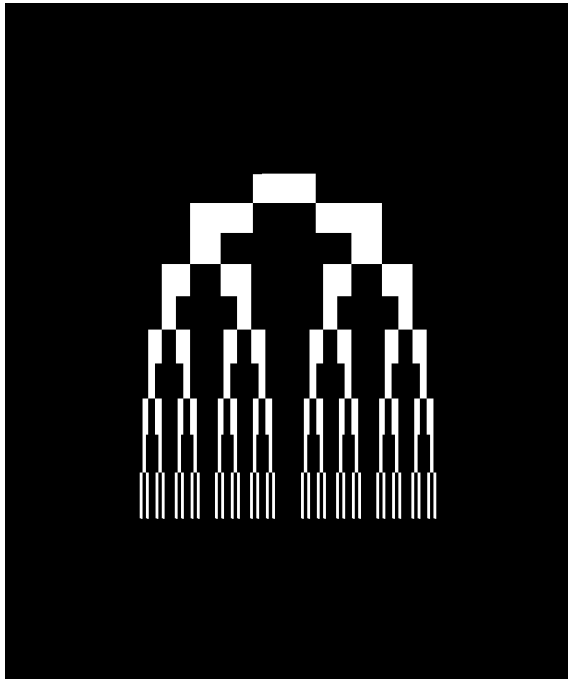**(b)** G_1L_4

(c) G_1L_5

(d) G_1L_6

Figure C.2: In the group one, each image size is 1200 by 1000 pixels with 90000 white pixels.
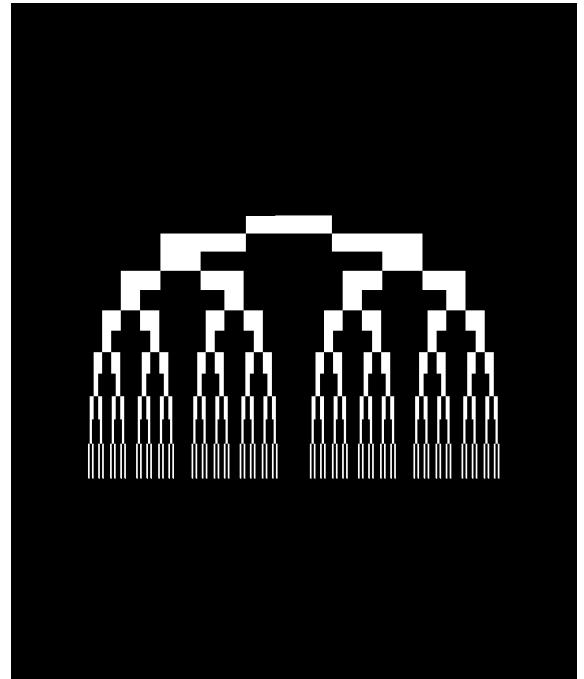
(a) G_2L_3

(b) G_2L_4

(c) G_2L_5

(d) G_2L_6

Figure C.3: In the group two, each image size is 1200 by 1000 pixels with 90000 white pixels.

References

[1] Mathilde Balduzzi, Brad M Binder, Alexander Bucksch, Cynthia Chang, Lilan Hong, Anjali S Iyer-Pascuzzi, Christophe Pradal, and Erin E Sparks. Reshaping plant biology: qualitative and quantitative descriptors for plant morphology. *Frontiers in Plant Science*, 8:117, 2017.

[2] Ingwer Borg and Patrick JF Groenen. *Modern multidimensional scaling: Theory and applications.* Springer Science & Business Media, 2005.

[3] Benjamin M Delory, Mao Li, Christopher N Topp, and Guillaume Lobet. archidart v3. 0: A new data analysis pipeline allowing the topological analysis of plant root systems. *F1000Research*, 7, 2018.

[4] *NIST Digital Library of Mathematical Functions.* http://dlmf.nist.gov/, Release 1.0.26 of 2020-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.

[5] AH Fitter. The topology and geometry of plant root systems: influence of watering rate on root system topology in trifolium pratense. *Annals of Botany*, 58(1):91–101, 1986.

[6] AH Fitter and TR Stickland. Fractal characterization of root system architecture. *Functional Ecology*, pages 632–635, 1992.

[7] HPW Gottlieb. Hearing the shape of an annular drum. *The ANZIAM Journal*, 24(4):435–438, 1983.

[8] HPW Gottlieb. Eigenvalues of the laplacian with neumann boundary conditions. *The ANZIAM Journal*, 26(3):293–309, 1985.

[9] Daniel Grieser and Svenja Maronna. Hearing the shape of a triangle. *Notices of the American Mathematical Society*, 60(11):1440–1447, 2013.

[10] Mark Kac. Can one hear the shape of a drum? *The American Mathematical Monthly*, 73(4P2):1–23, 1966.

[11] Mohamed A Khabou, Lotfi Hermi, and Mohamed Ben Hadj Rhouma. Shape recognition using eigenvalues of the dirichlet laplacian. *Pattern Recognition*, 40(1):141–153, 2007.

[12] Mao Li, Keith Duncan, Christopher N Topp, and Daniel H Chitwood. Persistent homology and the branching topologies of plants. *American Journal of Botany*, 104(3):349–353, 2017.

[13] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[14] BD Sleeman and EME Zayed. Trace formulae for the eigenvalues of the laplacian. *Zeitschrift für angewandte Mathematik und Physik*, 35(1):106–115, 1984.

[15] Jiro Tatsumi, Akira Yamauchi, and Yasuhiro Kono. Fractal analysis of plant root systems. *Annals of Botany*, 64(5):499–503, 1989.

[16] Zoltán Toroczkai. Topological classification of binary trees using the horton-strahler index. *Physical Review E*, 65(1):016130, 2001.

[17] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, CJ Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake Vand erPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1. 0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[18] EME Zayed. Heat equation for an arbitrary doubly-connected region in $\mathbb{R}^2$ with mixed boundary conditions. *Zeitschrift für angewandte Mathematik und Physik*, 40(3):339–355, 1989.