

Maximum Generalised Roundness of Graph Metric Spaces

Author:

De Silva, Raveen

Publication Date:

2020

DOI:

<https://doi.org/10.26190/unsworks/22323>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/70602> in <https://unsworks.unsw.edu.au> on 2024-04-24

MAXIMUM GENERALISED ROUNDNESS OF GRAPH METRIC SPACES

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Raveen de Silva

Supervisor: Associate Professor Ian Doust

School of Mathematics and Statistics,
Faculty of Science,
UNSW Sydney.

October 2020

Thesis/Dissertation Sheet

| | | |
|--|---|---|
| Surname/Family Name | : | de Silva |
| Given Name/s | : | Raveen Daminda |
| Abbreviation for degree as give in the University calendar | : | PhD |
| Faculty | : | Faculty of Science |
| School | : | School of Mathematics and Statistics |
| Thesis Title | : | Maximum Generalised Roundness of Graph Metric Spaces |

Abstract 350 words maximum: (PLEASE TYPE)

This thesis examines the maximum generalised roundness properties of metric spaces arising from graphs. The equivalent concepts of p -negative type and generalised roundness p , introduced by Schoenberg and Enflo respectively, have been used for several decades in the field of distance geometry, primarily to obstruct embeddings of metric spaces isometrically (or with some distortion) into Euclidean space. Given a metric space (X, d) , we are particularly interested in the calculation of its maximum generalised roundness $\wp(X, d)$, and many existing results from various authors are expounded in this thesis. A relatively recent formula of Sánchez provides a method to numerically calculate the quantity of interest, but implementing this in practice involves a number of previously unforeseen difficulties. In this thesis we develop a robust algorithm which enables the efficient numerical calculation of $\wp(X, d)$ for large sets of finite metric spaces.

We focus in particular on metric spaces constructed from graphs, usually with the path metric, in order to relate the maximum generalised roundness value to properties of the underlying graph. Many existing results describe the extremal values of the maximum generalised roundness of trees or connected graphs, but little is known about the values for 'typical' members of these classes. Equipped with the earlier algorithm, we are able to quickly and reliably calculate the maximum generalised roundness of graphs sampled at random from these families, and form hypotheses from the resulting data. We are able to prove that large random trees have maximum generalised roundness arbitrarily close to 1 almost surely, and provide more specific probabilistic results with explicit bounds. We also prove that for large random graphs, the maximum generalised roundness is arbitrarily close to 0 almost surely. In each case, we additionally state stronger heuristic conclusions which are supported by the empirical evidence, but for which we do not yet have rigorous proofs. We also present several partial results, conjectures and unexplained phenomena from our earlier investigations as well as peripheral work on infinite trees and planar graphs.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).

.....
Signature

.....
Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years can be made when submitting the final copies of your thesis to the UNSW Library. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed

Date

COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).'

'For any substantial portions of copyright material used in this thesis, written permission for use has been obtained, or the copyright material is removed from the final public version of the thesis.'

Signed

Date

AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.'

Signed

Date

INCLUSION OF PUBLICATIONS STATEMENT

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The candidate contributed greater than 50% of the content in the publication and is the “primary author”, ie. the candidate was responsible primarily for the planning, execution and preparation of the work for publication
- The candidate has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not:



This thesis contains no publications, either published or submitted for publication
(if this box is checked, you may delete all the material on page 2)



Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement
(if this box is checked, you may delete all the material on page 2)



This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the UNSW Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Candidate's Name

Signature

Date (dd/mm/yy)

Raveen de Silva

Acknowledgements

First of all, I would like to express my sincere gratitude to my supervisor Associate Professor Ian Doust. His guidance was crucial in every stage of the research and in compiling this thesis. I cannot thank him enough for his mentorship over so many years.

I would like to thank Professor Catherine Greenhill, who generously took the time to advise me on several graph theory problems encountered in this project. Dr. Anthony Weston also provided insights which were pivotal to my work.

Thanks also go to Shaymaa Al-Shakarchi, Gavin Robertson and Alan Stoneham, whose assistance and camaraderie are greatly appreciated.

Many thanks to Ramanan Rajkumar and Tracy Jiang for diligently proofreading this document.

Finally, I am deeply grateful for the love and support of my family and friends.

Raveen de Silva, October 19, 2020.

Abstract

This thesis examines the maximum generalised roundness properties of metric spaces arising from graphs. The equivalent concepts of p -negative type and generalised roundness p , introduced by Schoenberg and Enflo respectively, have been used for several decades in the field of distance geometry, primarily to obstruct embeddings of metric spaces isometrically (or with some distortion) into Euclidean space. Given a metric space (X, d) , we are particularly interested in the calculation of its maximum generalised roundness $\wp(X, d)$, and many existing results from various authors are expounded in this thesis. A relatively recent formula of Sánchez provides a method to numerically calculate the quantity of interest, but implementing this in practice involves a number of previously unforeseen difficulties. In this thesis we develop a robust algorithm which enables the efficient numerical calculation of $\wp(X, d)$ for large sets of finite metric spaces.

We focus in particular on metric spaces constructed from graphs, usually with the path metric, in order to relate the maximum generalised roundness value to properties of the underlying graph. Many existing results describe the extremal values of the maximum generalised roundness of trees or connected graphs, but little is known about the values for ‘typical’ members of these classes. Equipped with the earlier algorithm, we are able to quickly and reliably calculate the maximum generalised roundness of graphs sampled at random from these families, and form hypotheses from the resulting data. We are able to prove that large random trees have maximum generalised roundness arbitrarily close to 1 almost surely, and provide more specific probabilistic results with explicit bounds. We also prove that for large random graphs, the maximum generalised roundness is arbitrarily close to 0 almost surely. In each case, we additionally state stronger heuristic conclusions which are supported by the empirical evidence, but for which we do not yet

have rigorous proofs. We also present several partial results, conjectures and unexplained phenomena from our earlier investigations as well as peripheral work on infinite trees and planar graphs.

Contents

| | | |
|------------------|--|-----------|
| Chapter 1 | Introduction | 1 |
| 1.1 | History | 2 |
| 1.1.1 | Euclidean embeddings of metric spaces | 2 |
| 1.1.2 | p -negative type results and associated concepts | 8 |
| 1.1.3 | Generalised roundness p | 11 |
| 1.2 | Aims | 11 |
| 1.3 | Outline | 14 |
| Chapter 2 | Background | 16 |
| 2.1 | p -negative type and generalised roundness p | 16 |
| 2.1.1 | p -negative type | 16 |
| 2.1.2 | Generalised roundness p | 18 |
| 2.1.3 | Equivalence | 19 |
| 2.1.4 | Transforms of the metric | 21 |
| 2.1.5 | Metric subspaces | 21 |
| 2.1.6 | Application to classical spaces | 22 |
| 2.1.7 | The p -negative type gap | 25 |
| 2.2 | Graph metric spaces | 30 |
| 2.2.1 | Fundamentals of graph theory | 30 |
| 2.2.2 | Trees | 32 |
| 2.2.3 | Examples | 33 |
| 2.2.4 | Graph counting | 38 |

| | | |
|-----------|---|----|
| Chapter 3 | Calculation of maximum generalised roundness | 40 |
| 3.1 | Calculation of maximum generalised roundness of individual graph metric spaces | 41 |
| 3.2 | Computational issues | 43 |
| 3.2.1 | Determinant of the p -distance matrix | 43 |
| 3.2.2 | Sum of entries of the inverse of the p -distance matrix | 52 |
| 3.3 | Algorithm | 53 |
| 3.4 | Implementation and Validation | 56 |
| 3.4.1 | Parameters | 57 |
| 3.4.2 | Accuracy | 58 |
| 3.4.3 | Speed | 60 |
| 3.4.4 | Larger graphs | 61 |
| 3.4.5 | Verification | 62 |
| 3.5 | Calculation of the maximum generalised roundness of families of graph metric spaces | 64 |
| 3.6 | Exact calculation of the maximum generalised roundness | 66 |
| Chapter 4 | Maximum generalised roundness of trees | 71 |
| 4.1 | p -negative type properties of trees | 71 |
| 4.2 | The trees on n vertices | 73 |
| 4.3 | The uniform model of random trees | 75 |
| 4.4 | Large random trees | 78 |
| 4.5 | Limiting behaviour | 81 |
| 4.6 | Probabilistic bounds | 83 |
| 4.6.1 | Preliminaries | 85 |
| 4.6.2 | Solution to inequality 4.6.9 | 89 |
| 4.6.3 | Solution to inequality 4.6.8 | 90 |
| 4.6.3.1 | Solution to inequality 4.6.10 | 91 |
| 4.6.3.2 | Solution to inequality 4.6.11 | 92 |

| | |
|---|-----|
| 4.6.4 Conclusion | 93 |
| 4.6.5 An aside: further investigation of inequality 4.6.8 | 95 |
| 4.7 Empirical p -negative type properties of random trees | 100 |
| Chapter 5 Maximum generalised roundness of connected graphs | 107 |
| 5.1 p -negative type properties of graphs | 107 |
| 5.2 The graphs on n vertices | 110 |
| 5.3 Random graphs | 112 |
| 5.4 p -negative type properties of uniform random connected graphs . . | 114 |
| 5.5 Empirical p -negative type properties of uniform random connected graphs | 119 |
| 5.6 p -negative type properties of atypical connected graphs | 125 |
| 5.7 Empirical p -negative type properties of atypical connected graphs . | 129 |
| Chapter 6 Future Work | 132 |
| 6.1 Confluence of determinants for connected uniform random graphs . | 132 |
| 6.2 Weston tree conjecture | 136 |
| 6.3 Sánchez formula types | 137 |
| 6.4 Infinite metric trees | 141 |
| 6.5 Planar graphs | 145 |
| Appendix A Maple programs | 151 |
| Appendix B MATLAB programs | 160 |
| Appendix C Catalogue of trees | 163 |
| Appendix D Catalogue of graphs | 172 |
| References | 175 |

List of Tables

| | | |
|-----|---|-----|
| 3.1 | The first few digits of $\det D_p$ for a nearly singular 100×100 matrix. | 59 |
| 4.1 | The number of labelled and unlabelled trees on n vertices. | 74 |
| 4.2 | The maximum, minimum and average maximum generalised roundness of 1000 trees on n vertices selected uniformly at random. | 79 |
| 4.3 | Values of $\log f_1(\alpha, p)$ for various α, p . | 90 |
| 4.4 | Values of $f_2(\beta)$ for various $\beta \in (0, 1)$. | 92 |
| 4.5 | Values of $M_{\beta, 0.001}$ for various $\beta \in (0, 1)$. | 94 |
| 5.1 | The number of labelled and unlabelled connected graphs on n vertices. | 111 |
| 5.2 | The maximum, minimum and average maximum generalised roundness of 1000 graphs on n vertices from the connected uniform model. | 116 |
| 5.3 | Values of $\wp(L_{2m})$ for various m , to four decimal places. | 127 |
| 6.1 | Values of $\wp(G_m)$ for various m , to four decimal places. | 149 |
| C.1 | All trees on four unlabelled vertices and their maximum generalised roundness values. | 163 |
| C.2 | All trees on five unlabelled vertices and their maximum generalised roundness values. | 163 |
| C.3 | All trees on six unlabelled vertices and their maximum generalised roundness values. | 164 |
| C.4 | All trees on seven unlabelled vertices and their maximum generalised roundness values. | 165 |
| C.5 | All trees on eight unlabelled vertices and their maximum generalised roundness values. | 166 |
| C.5 | All trees on eight unlabelled vertices and their maximum generalised roundness values (continued). | 167 |

| | |
|---|-----|
| C.6 All trees on nine unlabelled vertices and their maximum generalised | |
| roundness values. | 168 |
| C.6 All trees on nine unlabelled vertices and their maximum generalised | |
| roundness values (continued). | 169 |
| C.6 All trees on nine unlabelled vertices and their maximum generalised | |
| roundness values (continued). | 170 |
| C.6 All trees on nine unlabelled vertices and their maximum generalised | |
| roundness values (continued). | 171 |
| D.1 All connected graphs on four unlabelled vertices and their maximum | |
| generalised roundness values. | 172 |
| D.2 All connected graphs on five unlabelled vertices and their maximum | |
| generalised roundness values. | 173 |
| D.2 All connected graphs on five unlabelled vertices and their maximum | |
| generalised roundness values (continued). | 174 |

List of Figures

| | | |
|-----|--|----|
| 1.1 | Two examples of metric spaces on four points. | 2 |
| 1.2 | Greater Capital City Statistical Areas of Australia by industry of employment, after multidimensional scaling. | 3 |
| 1.3 | A four-point metric space. | 6 |
| 1.4 | An embedding of the metric transform into three-dimensional Euclidean space. | 7 |
| 2.1 | A four-point metric space. | 28 |
| 2.2 | The seven point metric space C_7 . | 30 |
| 2.3 | A graph on three vertices. | 34 |
| 2.4 | A graph on three vertices. | 36 |
| 2.5 | The path graph on 5 vertices, P_5 . | 37 |
| 2.6 | The star graph on 6 vertices, S_6 . | 37 |
| 3.1 | A graph on seven vertices. | 44 |
| 3.2 | A plot of $\det D_p$ for the graph depicted in 3.1. | 45 |
| 3.3 | Symbolic computation of $\det D_p$ is much slower than repeatedly calculating D_p and its determinant numerically. | 46 |
| 3.4 | Some configurations with more than one zero within an interval. | 48 |
| 3.5 | For $G = C_5$, $\det D_p$ has a zero without a change in sign. | 49 |
| 3.6 | A graph on six vertices. | 50 |
| 3.7 | If one stationary point occurs in the interval, we may detect a zero without a change in sign. | 50 |
| 3.8 | If two or more stationary points occur in the interval, a zero may be missed. | 51 |

| | | |
|------|--|-----|
| 3.9 | Average seconds for Maple to calculate the determinant of 70×70 matrices numerically at different values of the precision variable d (on a standard PC). | 60 |
| 3.10 | Average seconds for MATLAB to calculate the determinant of 70×70 matrices numerically at different values of the precision variable d (on a standard PC). | 61 |
| 3.11 | $\det D_p$ for the graph $G(100, 0.5, 0)$. | 62 |
| 3.12 | $\det D_p$ for the graph $G(100, 0.5, 0)$ near the first zero. | 62 |
| 3.13 | $q(2^p)$ for the graph $G(100, 0.5, 0)$ near the first zero. | 64 |
| 3.14 | A graph on five vertices. | 66 |
| 4.1 | Log-scaled histogram of maximum generalised roundness values of all 262144 trees on 8 labelled vertices | 77 |
| 4.2 | The two trees on four vertices, up to isomorphism, P_4 and S_4 . | 78 |
| 4.3 | Histograms of maximum generalised roundness values of 1000 trees on n vertices selected uniformly at random, $20 \leq n \leq 90$ | 80 |
| 4.4 | The star graph on four vertices, S_4 . | 81 |
| 4.5 | Plots of the left-hand side of inequality (4.6.7) with various expressions for k in terms of n . | 86 |
| 4.6 | The function $f(t) = te^t$ and its inverse relation, the multi-valued Lambert W function. | 88 |
| 4.7 | A plot of the sample mean, minimum and maximum of the maximum generalised roundness values of $N = 1000$ trees on n vertices generated uniformly at random. | 101 |
| 4.8 | Curve fitting for the sample mean and standard deviation of one thousand trees on n vertices, $10 \leq n \leq 100$. | 103 |
| 4.9 | A plot of the z -scores of our observed values of $m_{n,1000}$ and $M_{n,1000}$, relative to $N(\bar{\varphi}_{n,1000}, s_{n,1000}^2)$. | 105 |
| 4.10 | A graph on 100 vertices with unusually large maximum generalised roundness. | 106 |
| 5.1 | Three graphs on four vertices. | 110 |

| | | |
|------|---|-----|
| 5.2 | Histograms of maximum generalised roundness values of 1000 graphs on n vertices from the connected uniform model, $20 \leq n \leq 90$ | 115 |
| 5.3 | A plot of the sample mean, maximum and minimum of the maximum generalised roundness values of $N = 1000$ graphs on n vertices generated by the connected uniform model. | 121 |
| 5.4 | Curve fitting for the sample mean and standard deviation of one thousand connected graphs on n vertices, $10 \leq n \leq 100$. | 123 |
| 5.5 | A plot of the z -scores of our observed values of $m_{n,1000}$ and $M_{n,1000}$, relative to $N(\bar{\varphi}_{n,1000}, s_{n,1000}^2)$. | 124 |
| 5.6 | The graph L_{10} . | 126 |
| 5.7 | A plot of $\varphi(G)$ against $ E(G) $ for $G = G^*(100, q)$, where $q \in [0.05, 0.95]$ uniformly at random. | 130 |
| 6.1 | $\det D_p$ for the graph $G(100, 0.5, 0)$. | 132 |
| 6.2 | A plot of $\det D_p$ for $N = 1000$ graphs on $n = 100$ vertices generated by the connected uniform model. | 133 |
| 6.3 | More detail of the ‘pinching’ effect from Figure 6.2. | 134 |
| 6.3 | More detail of the ‘pinching’ effect from Figure 6.2. (continued) | 135 |
| 6.4 | Two non-isomorphic six-vertex graphs with equal maximum generalised roundness. | 136 |
| 6.5 | Two non-isomorphic five-vertex graphs with equal maximum generalised roundness. | 137 |
| 6.6 | A tree on six vertices. | 138 |
| 6.7 | The determinant of the p -distance matrix of T from Example 6.3.6. | 139 |
| 6.8 | The number of Type II graphs among one thousand connected uniform random graphs on n vertices. | 139 |
| 6.9 | Histograms of maximum generalised roundness values among a total of one thousand graphs from the connected Erdős-Rényi model $G^*(50, q)$, where $q \in (0.05, 0.95)$ uniformly at random. | 140 |
| 6.10 | An example of a comb graph. | 141 |
| 6.11 | The forked path graph. | 142 |
| 6.12 | A truncation of the forked path graph. | 142 |

| | |
|---|-----|
| 6.13 A truncation of the forked path graph with added edges. | 142 |
| 6.14 A plot of $\det D_p$ for the cycle graph on five vertices. | 144 |
| 6.15 All connected graphs on six unlabelled vertices where the maximum generalised roundness occurs at a maximum turning point of $\det D_p$. | 145 |
| 6.16 The complete bipartite graph $K_{2,n-2}$ | 147 |
| 6.17 A planar graph on ten vertices | 148 |

CHAPTER 1

Introduction

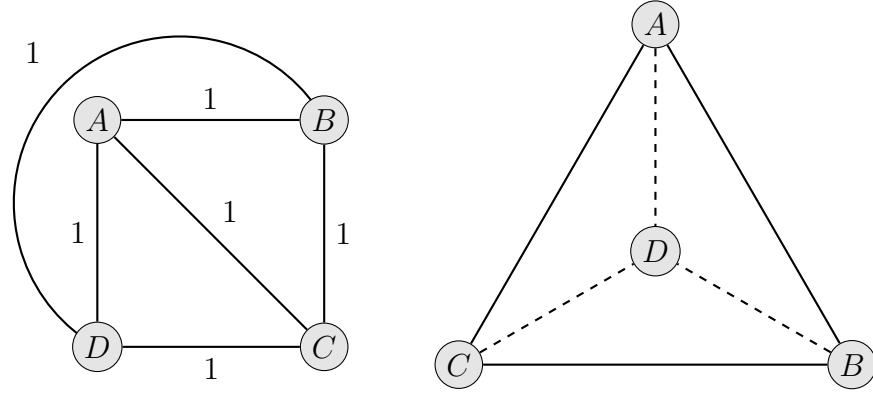
Distance geometry is the study of sets of points given only the pairwise distances between them. Metric spaces (and more generally semi-metric spaces, without the triangle inequality) are examples of settings where the distances between points are known, but there is no obvious sense of how the points are positioned.

A classical problem in distance geometry asks whether a particular metric space can be embedded isometrically into Euclidean space. For example, the discrete metric space on four points can be identified with the points of a regular tetrahedron in $(\mathbb{R}^3, \|\cdot\|_2)$ so that all distances are preserved, but similar embeddings cannot be achieved for some other four-point spaces, as shown in Figure [1.1](#).

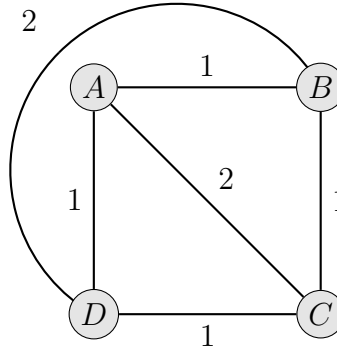
In the 1930s Schoenberg attacked this problem by introducing the concept of the p -negative type of the metric space. Much later it was shown that this was equivalent to a different concept, generalised roundness p , which originated from Enflo's solution to some quite different embedding problems. The central concept studied in this thesis is a constant $\wp(X, d)$ called the maximal generalised roundness, or supremal p -negative type, of a metric space (X, d) .

In Section [1.1](#) we give an overview of the history of these ideas and give some simple examples to illustrate the concepts involved.

Section [1.2](#) then describes the aims of the work in the later chapters of this thesis. This includes a more thorough understanding of the distribution of maximum generalised roundness among certain families of metric spaces, as well as an efficient algorithm for the numerical computation of the maximum generalised roundness of a finite metric space. Our main focus in the thesis will be on metric spaces which arise from graphs equipped with the path metric. At the end of this chapter we outline the structure of the remainder of the thesis.



(a) The discrete metric space on four points can be embedded into \mathbb{R}^3 .



(b) Some other metric spaces cannot be embedded into \mathbb{R}^n for any n .

Figure 1.1: Two examples of metric spaces on four points.

1.1 History

1.1.1 Euclidean embeddings of metric spaces

Many problems in fields such as sensor networks, molecular modelling and navigation require us to draw inferences about the arrangement of the points with some given distance structure within some vector space with well understood structure. Such information can allow us to visualise the points in order to use clustering or similar techniques, or to apply a wider set of tools such as those from linear algebra.

Multidimensional scaling is the process of representing dissimilarity data, which naturally correspond to the distances of a metric space, by finding a set of coordinates in some N -dimensional space which preserve these distances (often allowing some amount of distortion either of the distances or in their relative order). The following quote from Alon et al. [2] illustrates the utility of embedding results in this field.

‘The problem of multidimensional scaling is that of mapping points with some measured pairwise distances into some target metric space. Originally, the MDS community considered embeddings into an ℓ_p space, with the goal of aiding in visualization, compression, clustering, or nearest-neighbor searching; thus, low-dimensional embeddings were sought. An isometric embedding preserves all distances, while more generally, metric embeddings trade off the dimension with the fidelity of the embeddings.’

Example 1.1.1. Consider the fifteen largest Greater Capital City Statistical Areas of Australia, namely the greater metropolitan area of each state capital, the rest of each state, the Australian Capital Territory, Greater Darwin and the rest of the Northern Territory. The 2016 Australian Census [3] counted the number of residents of each area employed in each of twenty industries. After normalising, this data can be represented by fifteen points lying on a hyperplane in \mathbb{R}^{20} . Applying MATLAB’s `cmdscale` allows us to embed these points in Euclidean space of much lower dimension, at the cost of some distortion. In particular, Figure 1.2 shows an embedding in \mathbb{R}^2 where the relative error in the distances is up to 27%.

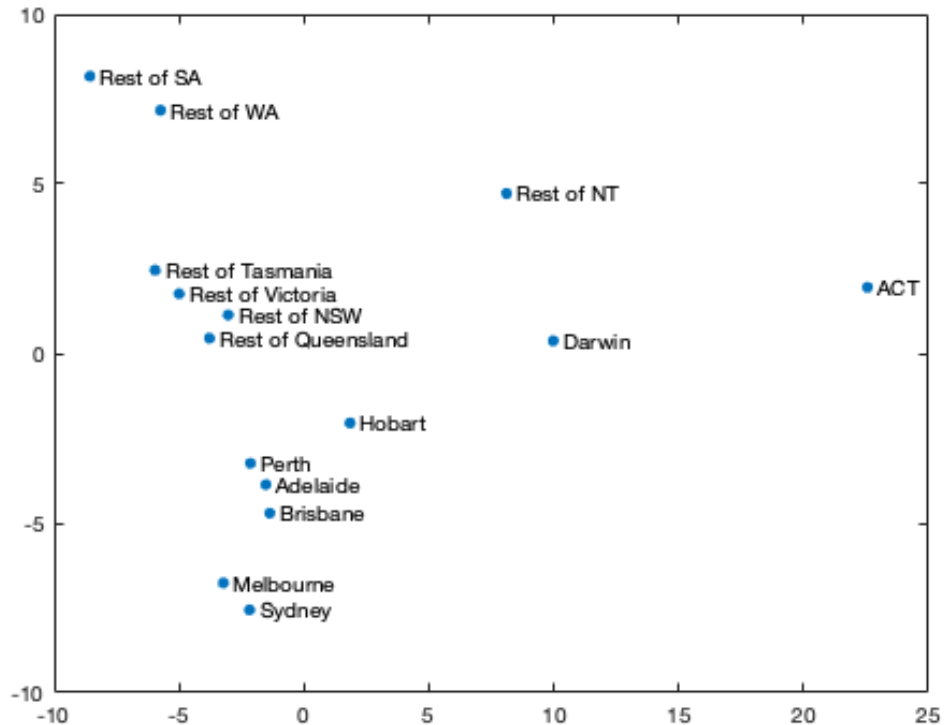


Figure 1.2: Greater Capital City Statistical Areas of Australia by industry of employment, after multidimensional scaling.

Even so, several conclusions can be immediately seen from this visualisation. The employment profiles are similar in the largest two capitals, Sydney and Melbourne, whereas the regional and rural areas of South Australia and Western Australia are quite distinct from those of the other states. Notably, the ACT is particularly dissimilar to any other area, primarily because of its large share of employees in Public Administration and Safety. \square

Another common application is in the construction of phylogenetic trees [1, 35, 75, 77], where the aim is to infer ancestry relationships among a group of organisms by analysing the pairwise differences between their DNA sequences in terms of the Hamming metric.

One approach to answering such questions follows from Cayley [18] and Menger [57, 58], who developed the Cayley-Menger determinant. This technique relates the $(n-1)$ -dimensional volume of an n -point simplex to the determinant of a particular $(n+1) \times (n+1)$ matrix, whose entries are purely in terms of the distances between the points. The minimum dimension of Euclidean space that permits such an embedding can also be found by analysing these determinants.

We will study another approach, first proposed by Schoenberg [72], who introduced the following inequality criterion and proved that it completely determines whether a finite metric space can be embedded into Euclidean space as well as the dimension required for such an embedding.

Theorem 1.1.2 ([72, Theorem 1]). *Let (X, d) be a finite metric space, where $X = \{x_0, x_1, \dots, x_n\}$. Then (X, d) embeds isometrically into Euclidean space if and only if*

$$\sum_{i,j=1}^n (d(x_0, x_i)^2 + d(x_0, x_j)^2 - d(x_i, x_j)^2) \omega_i \omega_j \geq 0$$

for all $\omega_1, \dots, \omega_n \in \mathbb{R}$.

Furthermore, the rank of the matrix

$$M = (d(x_0, x_i)^2 + d(x_0, x_j)^2 - d(x_i, x_j)^2)_{i,j=1}^n$$

is the minimum embedding dimension, that is, if $\text{rank}(M) = r$, then (X, d) embeds into \mathbb{R}^r but not into \mathbb{R}^{r-1} .

Schoenberg [74] used an algebraic manipulation to simplify this condition into a more symmetrical form.

Definition 1.1.3 ([74]). Let (X, d) be a finite metric space, where $X = \{x_1, \dots, x_n\}$. Then (X, d) has 2-negative type if

$$\sum_{i,j=1}^n \alpha_i \alpha_j d(x_i, x_j)^2 \leq 0$$

for all $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ such that $\alpha_1 + \dots + \alpha_n = 0$.

Theorem 1.1.4 ([74]). *A finite metric space (X, d) embeds isometrically into Euclidean space if and only if it has 2-negative type.*

The proof relied on the work of Mathias [54] and Bochner [8], who defined a real continuous even function $f(\mathbf{x})$ as positive definite if

$$\sum_{i,j=1}^m f(\mathbf{x}^{(i)} - \mathbf{x}^{(j)}) \rho_i \rho_j \geq 0$$

for any real ρ_i and points $\mathbf{x}^{(i)}$. Schoenberg equated both of the properties in Theorem 1.1.4 to the question of whether the function $e^{-|\mathbf{x}|^2}$ is positive definite, and in fact proved the more general statement that $e^{-|\mathbf{x}|^\alpha}$ is positive definite if and only if $0 < \alpha \leq 2$. In [73] he considered the effect of replacing a metric $d(x_i, x_j)$ with the new metric $d_c(x_i, x_j) = d(x_i, x_j)^c$ where $c \in (0, 1)$, and applied his results to deduce the embedding properties of the new metric space (X, d_c) . This is an example of what is now called a metric transform of (X, d) .

More generally, let F be a continuous increasing concave function such that $F(0) = 0$, guaranteeing that $F \circ d$ is still a metric on X . Denote the corresponding metric space $(X, F \circ d)$ as simply $F(X)$. A natural example is the metric space X^c given by the power function $F(t) = t^c$, where $0 < c < 1$. We observe that Euclidean embeddings of these metric transforms directly correspond to a generalisation of the 2-negative type property above.

Definition 1.1.5. Let (X, d) be a finite metric space, where $X = \{x_1, \dots, x_n\}$, and let $p \geq 0$. Then (X, d) has p -negative type if

$$\sum_{i,j=1}^n \alpha_i \alpha_j d(x_i, x_j)^p \leq 0$$

for all $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ such that $\alpha_1 + \dots + \alpha_n = 0$.

Furthermore, if (X, d) has p -negative type and the inequality is strict unless all α_i are zero, (X, d) is said to have strict p -negative type.

This definition can be naturally extended to the infinite case.

Definition 1.1.6. An infinite metric space is said to have (strict) p -negative type if every finite metric subspace has (strict) p -negative type.

The analogue to Theorem 1.1.4 is not exact. Instructive examples include infinite discrete metric spaces, in which all distances between discrete points are 1. Here, if the point set is countably infinite, the metric space embeds into ℓ_2 but not into Euclidean space. If instead the cardinality is larger than the continuum, it cannot be embedded into ℓ_2 .

Theorem 1.1.7. *An infinite metric space embeds isometrically into a (possibly non-separable) Hilbert space if and only if it has 2-negative type.*

We see that X^c is Euclidean, that is, has 2-negative type, if and only if X has $2c$ -negative type. We can then ask: for a given metric space (X, d) , for which values of c does X^c embed isometrically into Euclidean space? The following example illustrates a simple embedding problem of this form, which will motivate our future work.

Example 1.1.8. Let $X = \{x_1, x_2, x_3, x_4\}$, with

$$d(x_i, x_j) = \begin{cases} \frac{1}{2} & i = 4 \text{ or } j = 4 \\ 1 & \text{otherwise} \end{cases},$$

as depicted in Figure 1.3.

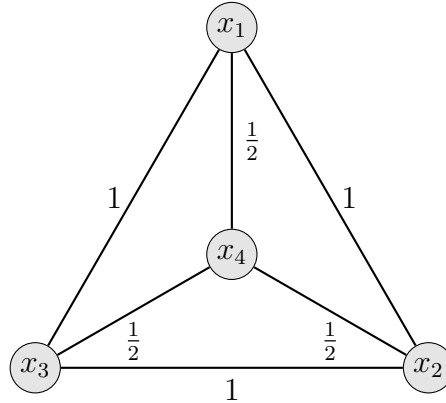


Figure 1.3: A four-point metric space.

We will first use a naïve geometrical approach, and then use Theorem 1.1.4 to present an alternative approach in terms of 2-negative type, which will generalise more easily to other metric spaces.

It is straightforward to see that this metric space is not embeddable into Euclidean space, as x_4 would have to be the midpoint of each interval joining two of the other points.

However, with some deformation, the embedding may be possible. Consider instead the metric transform X^c , by replacing each distance of $\frac{1}{2}$ with $\gamma = 2^{-c}$. Again, a geometric argument suffices here. Suppose an embedding exists. As there are only four points, the minimum embedding dimension is at most three, that is, X^c embeds into three-dimensional Euclidean space, as shown in Figure 1.4.

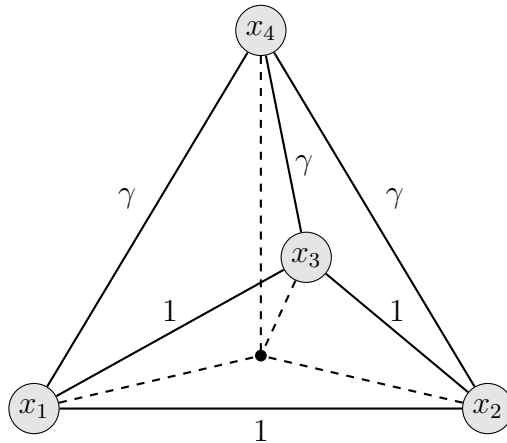


Figure 1.4: An embedding of the metric transform into three-dimensional Euclidean space.

The points x_1 , x_2 and x_3 form an equilateral triangle, and by symmetry x_4 must lie on a line through the centre of this triangle, orthogonal to the plane in which it lies. It follows that the minimum value of γ is realised when the four points are coplanar, with elementary trigonometry showing that this value is $\gamma = 1/\sqrt{3}$. Therefore X^c is embeddable if and only if $2^{-c} \geq 3^{-1/2}$, that is, $c \leq \frac{1}{2} \log_2 3$, and thus (X, d) has p -negative type for $p \in [0, \log_2 3]$.

We will now confirm our results in terms of the 2-negative type criterion from Theorem 1.1.4.

First, we test whether the original metric space X has 2-negative type. The desired inequality is

$$(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3) + \frac{1}{4}(\alpha_1 + \alpha_2 + \alpha_3)\alpha_4 \leq 0,$$

which after eliminating α_4 reduces to

$$(\alpha_1 + \alpha_2 + \alpha_3)^2 \geq 4(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3).$$

We must now ask whether this holds for all $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. It is clear that this fails; consider for example $\alpha_1 = \alpha_2 = \alpha_3 = 1$.

However, for X^c to have 2-negative type, we require

$$(\alpha_1 + \alpha_2 + \alpha_3)^2 \geq 2^{2c}(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3).$$

Again considering $\alpha_1 = \alpha_2 = \alpha_3 = 1$, we see that c can be no greater than $\frac{1}{2} \log_2 3$. In fact, we can prove that this value is sharp by manipulating

$$(\alpha_1 - \alpha_2)^2 + (\alpha_2 - \alpha_3)^2 + (\alpha_3 - \alpha_1)^2 \geq 0$$

to give

$$(\alpha_1 + \alpha_2 + \alpha_3)^2 \geq 3(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3).$$

Thus X^c has 2-negative type (and therefore embeds isometrically in \mathbb{R}^3) for all $c \in [0, \frac{1}{2} \log_2 3]$. Once again it follows that X has p -negative type for $p \in [0, \log_2 3]$, confirming our earlier result. \square

It is not hard to see that when more points are involved, the geometric analysis becomes increasingly intricate and eventually infeasible. As a result, p -negative type presents an attractive option for these embedding problems.

The embedding properties of metric transforms were further investigated by Deza and Maehara [23], who proved that if X is an n -point metric space, then X^c embeds in a Euclidean space for

$$0 < c \leq \frac{1}{2n \log 2} = \frac{0.7213 \dots}{n}.$$

They made a conjecture¹ as to how much this bound could be improved, and proved that their proposed bound is sharp for a particular ‘truncated distance’ metric.

1.1.2 p -negative type results and associated concepts

Given the results above, a natural direction is to describe the set of p for which a given metric space has p -negative type. Let

$$\mathcal{P}(X, d) = \{p \geq 0 : (X, d) \text{ has } p\text{-negative type}\}.$$

¹This conjecture will be presented in Chapter 5.

An easy calculation (see Proposition 2.1.3) confirms that all metric spaces have 0-negative type. Schoenberg [74] proved that if (X, d) has p -negative type, then it has q -negative type for all $0 \leq q < p$, so the set $\mathcal{P}(X, d)$ is either just $\{0\}$ or it is an interval. Further, this interval is closed, as seen by fixing the α_i in Definition 1.1.5 and taking the limit as the exponent p approaches the right endpoint from below. Thus the set $\mathcal{P}(X, d)$ is always of the form $[0, \wp]$ or $[0, \infty)$. While the quantity \wp is a maximum, it is often referred to in the literature simply as a supremum, and we use the same terminology here.

Definition 1.1.9. Let (X, d) be a metric space. Define the supremal p -negative type of (X, d) as

$$\wp(X, d) = \sup\{p \geq 0 : (X, d) \text{ has } p\text{-negative type}\},$$

with $\wp(X, d) = \infty$ if the supremum does not exist.

The supremal p -negative type can be explicitly computed for several classical Banach spaces. The first such result was established by Schoenberg [74], who proved that for $q \leq 2$, $L_q([0, 1])$ has p -negative type for all $0 < p \leq q$. Note that for $q < 1$, the corresponding L_q space is only a quasimetric space, with the triangle inequality relaxed by a constant factor of $2^{\frac{1}{q}-1}$.

Schoenberg also proposed the question of the p -negative type properties of $L_q([0, 1])$ for $q > 2$. Yost [84] proved that all real normed spaces of dimension at most two embed linearly and isometrically into $L_1[0, 1]$, so $\wp \geq 1$ in these cases. It was proven that the finite dimensional sequence spaces $\ell_q^{(n)}$ where $n \geq 3$ fail to have p -negative type for any $p > 0$, as a consequence of results from Koldobsky [49] for $2 < q < \infty$ and Misiewicz [59] for $q = \infty$, showing a marked difference from the $q \leq 2$ case. These results were generalised by Lennard, Tonge and Weston [51] to any $L_q(\mu)$ space of dimension at least three.

Lennard, Tonge and Weston [51] noted that as a consequence of this result, the Schatten class \mathcal{C}_q , where $q > 2$, also fails to have p -negative type for any $p > 0$. A recent result of Dahma and Lennard [21] extended the same result to $q \in (0, 2)$.

These supremal p -negative type values often present an obstruction to isometric embeddings, as a metric (or quasimetric) space cannot embed into another of larger supremal p -negative type. Proving the existence of an embedding from p -negative type information is far more difficult, and requires additional structure. Bretagnolle, Dacunha-Castelle and Krivine [11] discovered the surprising result that for $0 < q \leq$

2, a real quasinormed space (in which the triangle inequality associated with a norm is instead relaxed by a constant factor) can be linearly embedded into L_q space if and only if it has q -negative type. A notable consequence of this result is the nesting of real L_q spaces: if $0 < q_1 < q_2 \leq 2$ then there is a linear isometric embedding of L_{q_2} into L_{q_1} . More recent work has aimed to discover analogous results about metric spaces which do not have the structure of a vector space, such as the work of Graham [41, 42], Graham and Winkler [43, 44] and Roth and Winkler [70] on embeddings of graphs, including into Cartesian products of other graphs.

A relaxation of the graph embedding problem asks whether a graph $G = (V, E)$ is a unit-distance graph in \mathbb{R}^d . This criterion is satisfied if there exists an injective function $f : V \rightarrow \mathbb{R}^d$ such that if $\{v, w\} \in E$, then $|f(v) - f(w)| = 1$. Thus neighbouring vertices can be mapped to pairs of points one unit apart, but unlike the isometric embedding problems discussed in this thesis, there is no requirement to also preserve distances greater than one within the graph. Erdős, Harary and Tutte [32] defined the Euclidean dimension $\dim G$ as the smallest d for which G is a unit-distance graph in \mathbb{R}^d . For forty years, the best available result was due to Erdős and Simonovits [34], who proved that $\dim G \leq \Delta(G) + 2$, where $\Delta(G)$ is the maximum degree of the graph. A recent paper of Frankl, Kupavskii and Swanepoel [37] improved this bound to $\dim G \leq \Delta(G)$, with the exception of cases where $\Delta(G) = 3$ and G contains $K_{3,3}$, and proved that if $|E(G)| \leq \binom{d+2}{2}$ then $\dim G \leq d$, as well as further results on the related concept of spherical dimension.

Other related work relates to ultrametric spaces, in which the triangle inequality is further strengthened to

$$d(x_1, x_3) \leq \max(d(x_1, x_2), d(x_2, x_3)).$$

These correspond exactly to certain dendrograms, as shown by Carlsson and Mémoli [17, Theorem 9], and therefore find applications in clustering techniques for data analysis among other fields. Faver et al. [36] proved several key results for these spaces, most notably that ultrametric spaces are precisely those metric spaces with strict p -negative type for all $p \geq 0$. This work was extended by Doust, Sánchez and Weston [27] in their analysis of the p -negative type gap of ultrametric spaces, providing an explicit combinatorial formula for the asymptotic negative type constant.

1.1.3 Generalised roundness p

In 1964 Lindenstrauss [53] conjectured that no two Lebesgue spaces $L_q[0, 1]$ (where $q \in [1, 2]$) are uniformly homeomorphic. This conjecture was proven by Enflo [30] in 1968. The method employed by Enflo involves the definition of a property of metric spaces now known as² roundness p , where p is again a non-negative real constant. The criterion for a given metric space to have roundness p involves an inequality condition reminiscent of the parallelogram identity of Euclidean space. Enflo proved that $L_q[0, 1]$ has roundness q , and fails to have roundness p for any $p > q$, and used this to show that $L_{q_1}[0, 1]$ and $L_{q_2}[0, 1]$ are not uniformly homeomorphic for $1 \leq q_1, q_2 \leq 2$ where $q_1 \neq q_2$.

A related question was asked by the topologist Yuri M. Smirnov concerning universal uniform embedding spaces. A metric space M is a universal uniform embedding space if every separable metric space is uniformly homeomorphic to a subset of M . Since every separable metric space is isometric to a subset of $C[0, 1]$, this Banach space is certainly a universal uniform embedding space. It was already known that every separable metric space is homeomorphic to a subset of $L_2[0, 1]$, but not necessarily uniformly homeomorphic. Smirnov's question was whether this Hilbert space is a universal uniform embedding space.

Enflo [31] introduced a new property called³ generalised roundness p , generalising his earlier concept of roundness. He proved that all Hilbert spaces have generalised roundness 2 while universal uniform embedding spaces cannot have generalised roundness p for any $p > 0$, and deduced that no Hilbert space could be a universal uniform embedding space.

Research on p -negative type and generalised roundness p was undertaken independently for several decades, until Lennard, Tonge and Weston [51, Theorem 2.4] proved that the two conditions are in fact equivalent. We can therefore refer to the constant $\wp(X, d)$ as either the supremal p -negative type or the maximum generalised roundness; we will use the latter term. This constant will be the key object of study in this thesis.

1.2 Aims

Calculating the maximum generalised roundness of any particular metric space has historically been very challenging, and until recently there were rather few spaces

²Enflo's original terminology is slightly different from the current standard.

³Again, the formal definition can be found in Section 2.1.

for which this value was known explicitly, even among finite metric spaces with relatively few points. Certain upper and lower bounds were known for particular classes of metric spaces, but it has not been clear whether these bounds are sharp or not, and it has been particularly unclear whether the spaces for which the maximum generalised roundness can be computed are ‘typical’ or ‘exceptional’ within these families.

Examples of some of the open problems in this area include

Question 1: For each $n \geq 2$ the quantity

$$c_n = \inf\{\wp(X, d) : (X, d) \text{ is an } n \text{ point metric space}\}$$

is strictly positive. Only the values of c_2, c_3, c_4 and c_6 are known, from [23].

What are the values of c_5, c_7, c_8, \dots ? What is the limiting behaviour of c_n ?

Are these infima achieved, and if so, for which metric spaces?

Question 2: For each $n \geq 2$, the quantity

$$d_n = \inf\{\wp(X, d) : (X, d) \text{ is an } n \text{ point metric tree}\}$$

is strictly greater than 1. What are the values of this sequence, and how does it behave as $n \rightarrow \infty$? Are the infima achieved, and if so, for which trees?

In Chapter 5, we present a general conjecture for c_n , which Deza and Maehara [23] proved to be a lower bound, and is widely believed to give the true value. Until now, however, there has not been a great deal of empirical evidence to support the conjecture.

Previous research has made more progress in revealing generalised roundness properties of metric trees. The first major result in this area was the proof by Hjorth et al. [47] that finite metric trees are of strict 1-negative type. This was extended by Doust and Weston [28], who introduced the p -negative type gap, a measure of strictness of the p -negative type inequality. These authors provided a concise formula to evaluate the 1-negative type gap of metric trees, which led to a lower bound for d_n as proven by Li and Weston [52]. Again, there is currently no basis to assess whether this bound is at all sharp.

A result of Sánchez [71], which we present in Chapter 3, gave a formula for the maximum generalised roundness of a finite metric space in terms of certain equations associated to its distance matrix. For some special spaces, these equations can be

solved exactly to obtain an explicit expression for $\wp(X, d)$, but in general they can only be solved numerically. Nevertheless, this has opened up the possibility to collect experimental data within natural classes of finite metric spaces. An important class of such spaces is the class of path metric graphs, that is, metric spaces which arise from the unweighted path metric of a connected graph. These metric spaces are of course important in many areas, but they have the added advantage that much is known about the general properties of this class, and there are finitely many of them on a specified number of points.

Much of the work in this thesis aims to establish results about the distribution of the maximal generalised roundness within various classes of such metric spaces. A starting point for this project was to empirically answer questions such as

Question 3: What can be said about the maximal generalised roundness of a randomly chosen n vertex metric graph? How do the average and other distribution features vary with n ?

Question 4: What can be said about the distribution of maximum generalised roundness of metric trees on n vertices?

Question 5: What can be said about the distribution of maximum generalised roundness of planar graphs on n vertices?

The results of this experimental work could then be used to form reasonable conjectures which might be proven analytically.

Since it involves calculation of the determinant and the inverse of an $n \times n$ matrix, applying Sánchez's formula for numerical computation presents a number of challenges, especially as the size of the metric space gets large. In Chapter 3 we present several previously unexplored properties of this p -distance matrix and discuss the practical issues that arise. A significant outcome of the thesis is a robust and reliable algorithm for the computation of the maximum generalised roundness of a finite metric space with up to 100 points. These metric spaces are large enough that general patterns can be seen, and convincing conjectures formed.

A further challenge is that the size of these families of metric spaces grows very rapidly in n , and so it quickly becomes impossible to evaluate $\wp(X, d)$ for every space in the family. Therefore we must decide on a suitable random graph model to work with, and a way of generating elements via this model. Some of our results will turn out to depend crucially on the model chosen.

One quickly sees from the data that with high probability, the maximal generalised roundness of an n point metric graph goes to zero as $n \rightarrow \infty$. To make this precise one needs to prove a statement of the following form.

Conjecture: For all $\epsilon, \delta > 0$ there exists N such that for all $n \geq N$,

$$\mathbb{P}(\wp(X, d) < \epsilon \mid (X, d) \text{ is an } n \text{ point metric graph}) > 1 - \delta.$$

It is actually not hard to give a proof of this. Much more challenging is to quantify this more tightly by determining how N depends on ϵ and δ . How large does n need to be so that a randomly chosen n point metric graph (X, d) has $\wp(X, d) < 0.1$ with probability 0.9999? Similar phenomena arise in our study of trees, planar graphs and so on.

The thesis is then partly experimental and partly theoretical, and one may classify the various results within it as analytically proven, experimentally verified or ‘computationally dependent’. This last category is perhaps best explained by an example.

Example 1.2.1. Consider the sequence $(a_k)_{k=1}^{10000}$, where $a_k = \sin k + \sin(k^2)$. It is difficult to provide an elegant analytical proof that the largest element of this sequence is a_{6046} . Instead, one might compute each term numerically, and select the largest value. However, these values are necessarily inexact, so one must have confidence in how they are obtained. In particular, one must ensure that a_{6046} is definitively the largest term, even accounting for the error bounds on all terms. One might need to perform a more careful analysis of the values which are near to optimal in order to be convinced that a_{6046} is indeed the largest. \square

In the context of this thesis, the development of a robust algorithm for computing $\wp(X, d)$ allows us to similarly identify the extreme values of the maximum generalised roundness in some families of metric graphs.

1.3 Outline

In Chapter [2](#), we formally introduce the basic definitions and results for p -negative type and generalised roundness p , as well as the graph theory used in later chapters.

In Chapter [3](#), we present a formula due to Sánchez [\[71\]](#) which is our primary method for calculating the maximum generalised roundness of finite metric spaces

in terms of the p -distance matrix. We then design and verify an algorithm implementing this theorem, and explore the complications involved.

In Chapter [4](#), we explore the maximum generalised roundness properties of path metric trees. We first detail existing procedures to generate random trees with either uniform or non-uniform probability distributions. Using the algorithm from Chapter [3](#) on random samples of trees, we can then approximate the distribution of maximum generalised roundness values of trees. This data leads us to hypothesise that the maximum generalised roundness of a random tree on n vertices, under appropriate models, converges in distribution to 1. This result is proven by identifying metrically embedded star graphs, drawing upon the work of Moon [\[61\]](#) on the vertex degrees of random trees. We then provide probabilistic results with explicit bounds for n , and present empirical data to support stronger conclusions.

In Chapter [5](#), we apply a similar approach to path metric graphs. We generate random graphs on n vertices in order to calculate their maximum generalised roundness. In this case we show that the distribution converges to 0, making use of the work of Palka [\[66\]](#) on complete bipartite induced subgraphs. The bounds obtained are again far from sharp, as indicated by our test data. We also discuss the relationship between the density of a graph and its maximum generalised roundness.

In Chapter [6](#), we lay out some directions for future work, including partial results concerning infinite trees and planar graphs.

CHAPTER 2

Background

In this chapter we introduce the metric space properties of p -negative type and generalised roundness p and associated concepts which will be drawn upon in later chapters. Proofs will be included where they are helpful in understanding the later material. We also present some fundamentals of graph theory, including the construction of metric spaces from unweighted and weighted graphs, and outline the application of p -negative type and generalised roundness p ideas to graph metric spaces.

2.1 p -negative type and generalised roundness p

2.1.1 p -negative type

We begin with a formal exposition of Schoenberg's p -negative type criterion, first introduced in 1937 [74].

Definition 2.1.1. Let $p \geq 0$. A metric space (X, d) has p -negative type if

$$\sum_{i,j=1}^k \alpha_i \alpha_j d(x_i, x_j)^p \leq 0 \quad (2.1.1)$$

for all $k \in \mathbb{N}$, $x_1, \dots, x_k \in X$ and $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ with $\sum_{\ell=1}^k \alpha_\ell = 0$. Note that $0^0 := 0$ here, and the x_i need not be distinct.

Note that this formulation applies equally to finite and infinite metric spaces, as discussed in Remark 1.1.6.

The stronger condition of strict p -negative type proves to be useful in later results.

Definition 2.1.2. Let $p \geq 0$, and let (X, d) be a metric space of p -negative type. (X, d) has strict p -negative type if the equality case of (2.1.1) occurs only when all $\alpha_\ell = 0$.

Questions of embeddings, and indeed inequality (2.1.1), are somewhat meaningless for a one point metric space. For a two point space, (2.1.1) reduces to the question of whether $-\alpha^2 d(x_1, x_2)^p \leq 0$ for all $\alpha \in \mathbb{R}$, which is obviously true for all $p \geq 0$. For this reason, unless specified otherwise we will consider only nontrivial metric spaces on at least three points and likewise graphs on at least three vertices in this thesis.

Given a metric space (X, d) , we would like to know for which values of p it has (strict) p -negative type.

Proposition 2.1.3. *Any metric space (X, d) has strict 0-negative type.*

Proof. Let x_1, \dots, x_k be distinct points of X . Setting $p = 0$, so $d(x_i, x_j)^p = 1$ for distinct i, j , we have

$$\sum_{i,j=1}^k \alpha_i \alpha_j d(x_i, x_j)^p = \sum_{\substack{i,j=1 \\ i \neq j}}^k \alpha_i \alpha_j = \sum_{i=1}^k \alpha_i \sum_{j=1}^k \alpha_j - \left(\sum_{\ell=1}^k \alpha_\ell^2 \right) = - \sum_{\ell=1}^k \alpha_\ell^2 \leq 0,$$

with equality if and only if all $\alpha_\ell = 0$. □

Schoenberg [73, Theorem 2] proved that a finite metric space of 2-negative type also has q -negative type for all $0 \leq q < 2$. This was generalised in 2009 by Li and Weston as follows.

Theorem 2.1.4 ([52, Theorem 5.4]). *Let $p \geq 0$, and let (X, d) be a metric space of p -negative type. Then (X, d) has strict q -negative type for all $0 \leq q < p$.*

Corollary 2.1.5 ([52, Corollary 5.11]). *Let (X, d) be a metric space of p -negative type, but not strict p -negative type. Then $\wp(X, d) = p$.*

Proof. If (X, d) has p -negative type, then clearly $\wp(X, d) \geq p$. If $\wp(X, d) = q > p$, then Theorem 2.1.4 implies that (X, d) has strict p -negative type. Thus $\wp(X, d) = p$ as required. □

It follows that a metric space (X, d) has p -negative type for all values p in a (possibly degenerate) interval with left endpoint 0 included. The right endpoint therefore determines the set almost completely.

Definition 2.1.6. Let (X, d) be a metric space. The supremal p -negative type is

$$\wp(X, d) = \sup\{p \geq 0 : (X, d) \text{ has } p\text{-negative type}\},$$

or ∞ if the supremum does not exist.

Remark 2.1.7. Considering the left-hand side of (2.1.1), we observe that

$$\sup_{\substack{x_1, \dots, x_k \in X \\ \alpha_1 + \dots + \alpha_k = 0}} \sum_{i,j=1}^k \alpha_i \alpha_j d(x_i, x_j)^p$$

is a continuous function of p . This ensures that if $\wp(X, d)$ is finite, then (X, d) has $\wp(X, d)$ -negative type. Thus the supremum is in fact a maximum if it is finite.

2.1.2 Generalised roundness p

We now present the notion of generalised roundness p , from Enflo [30, 31].

Definition 2.1.8. Let $p \geq 0$. A metric space (X, d) has generalised roundness p if

$$\sum_{\substack{i,j=1 \\ i < j}}^n d(a_i, a_j)^p + \sum_{\substack{i,j=1 \\ i < j}}^n d(b_i, b_j)^p \leq \sum_{i,j=1}^n d(a_i, b_j)^p \quad (2.1.2)$$

for all $n \in \mathbb{N}$ and $a_1, \dots, a_n, b_1, \dots, b_n \in X$, where the a_i and b_j are not necessarily distinct.

Note that Definition 2.1.8 allows a point $x \in X$ to appear among both the a_i and the b_j simultaneously. However, Lennard, Tonge and Weston [51] showed that if $a_1 = b_1$, the terms involving this point cancel, so the inequality reduces to

$$\sum_{\substack{i,j=2 \\ i < j}}^n d(a_i, a_j)^p + \sum_{\substack{i,j=2 \\ i < j}}^n d(b_i, b_j)^p \leq \sum_{i,j=2}^n d(a_i, b_j)^p.$$

It therefore suffices to consider only cases where $\{a_i\}_{i=1}^n$ and $\{b_j\}_{j=1}^n$ are disjoint.

Although we may assume that no point appears among both $\{a_i\}_{i=1}^n$ and $\{b_j\}_{j=1}^n$, we must still allow repetition within each set. Consequently, even for a finite metric space there is no natural upper bound on the size of these sets. In other words, one cannot reduce the question of whether a finite metric space has generalised roundness p to an exercise of checking (2.1.2) for some finite collection of sets $\{a_i\}$ and $\{b_j\}$.

A useful alternative was developed by Lennard, Tonge and Weston [51], who noted that any choice of multiplicities of a_i and b_j can be simulated by assigning a weight to each point.

Definition 2.1.9. A (q, t) -simplex in X is a vector

$$S = [a_i; b_j]_{q,t} = (a_1, \dots, a_q, b_1, \dots, b_t)$$

where $a_1, \dots, a_q, b_1, \dots, b_t$ are distinct elements of X .

Definition 2.1.10. A load vector for a (q, t) -simplex S is an vector

$$\vec{\omega} = (m_1, \dots, m_q, n_1, \dots, n_t)$$

where $m_1, \dots, m_q, n_1, \dots, n_t \in \mathbb{R}^+$. Together, a (q, t) -simplex S and an associated load vector $\vec{\omega}$ form a loaded (q, t) -simplex denoted $S(\vec{\omega}) = [a_i(m_i); b_j(n_j)]_{q,t}$. We call $S(\vec{\omega})$ a normalised (q, t) -simplex if $\vec{\omega}$ also satisfies

$$m_1 + \dots + m_q = 1 = n_1 + \dots + n_t.$$

Theorem 2.1.11 ([51, Theorem 2.2]). *A metric space (X, d) has generalised roundness p if and only if for all $q, t \in \mathbb{N}$ and all normalised (q, t) -simplices $S(\vec{\omega}) = [a_i(m_i); b_j(n_j)]_{q,t}$ in X we have*

$$\sum_{1 \leq i_1 < i_2 \leq q} m_{i_1} m_{i_2} d(a_{i_1}, a_{i_2})^p + \sum_{1 \leq j_1 < j_2 \leq t} n_{j_1} n_{j_2} d(b_{j_1}, b_{j_2})^p \leq \sum_{i,j=1}^{q,t} m_i n_j d(a_i, b_j)^p. \quad (2.1.3)$$

The reverse direction requires only straightforward algebraic manipulation, while the forward direction relies on a density argument.

2.1.3 Equivalence

In 1995, almost three decades after Enflo defined generalised roundness p , it was proven by Lennard, Tonge and Weston that p -negative type and generalised roundness p are in fact equivalent.

Theorem 2.1.12 ([51, Theorem 2.4]). *A metric space (X, d) has p -negative type if and only if it has generalised roundness p .*

Proof. We shall only prove the forward direction. Assume that (X, d) has p -negative type.

Let $n \in \mathbb{N}$ and let $a_1, \dots, a_n, b_1, \dots, b_n \in X$. Set

$$x_1 = a_1, x_2 = a_2, \dots, x_n = a_n$$

and

$$x_{n+1} = b_1, x_{n+2} = b_2, \dots, x_{2n} = b_n,$$

and define

$$\alpha_k = \begin{cases} 1 & \text{if } 1 \leq k \leq n \\ -1 & \text{if } n+1 \leq k \leq 2n \end{cases}$$

so that $\sum_{k=1}^{2n} \alpha_k = 0$. Then the p -negative type inequality (2.1.1) guarantees that

$$\sum_{i,j=1}^{2n} \alpha_i \alpha_j d(x_i, x_j)^p \leq 0.$$

The left hand side can be decomposed further, so we have

$$\sum_{i,j=1}^n d(a_i, a_j)^p - \sum_{i,j=1}^n d(a_i, b_j)^p - \sum_{i,j=1}^n d(b_i, a_j)^p + \sum_{i,j=1}^n d(b_i, b_j)^p \leq 0$$

and upon rearranging and dividing by two, we have

$$\sum_{\substack{i,j=1 \\ i < j}}^n d(a_i, a_j)^p + \sum_{\substack{i,j=1 \\ i < j}}^n d(b_i, b_j)^p \leq \sum_{i,j=1}^n d(a_i, b_j)^p,$$

so (X, d) has generalised roundness p .

The proof of the reverse direction is much longer. It relies on an algebraic manipulation of the generalised roundness p inequality followed by a density argument, so it is omitted here. \square

Recalling the supremal p -negative type $\wp(X, d)$ from Definition 2.1.6, we can therefore assign the same notation to the analogue in terms of generalised roundness.

Definition 2.1.13. Let (X, d) be a metric space. The maximum generalised roundness of (X, d) is

$$\wp(X, d) = \sup\{p \geq 0 : (X, d) \text{ has generalised roundness } p\},$$

or ∞ if the supremum does not exist.

We will refer to $\wp(X, d)$ primarily as the maximum generalised roundness throughout this thesis. Again, Remark [2.1.7](#) guarantees that the supremum is achieved if it is finite, so it is indeed a maximum.

2.1.4 Transforms of the metric

Suppose that (X, d) is a metric space. It is clear that any positive multiple of d also gives a metric on X , and it is a standard fact that d^c is also a metric for any $c \in [0, 1]$. A simple application of the definition of p -negative type shows the following results which will be needed later in the thesis.

Theorem 2.1.14. *Let (X, d) be a metric space. Then $\wp(X, cd) = \wp(X, d)$ for any positive c .*

Proof. Applying the p -negative type inequality [\(2.1.1\)](#) to (X, cd) , a common factor of c^p appears in every term of the sum, and dividing this out recovers the original p -negative type inequality. Thus p -negative type is invariant under scaling, and so is the maximum generalised roundness. \square

Theorem 2.1.15. *Let (X, d) be a metric space. Then $\wp(X, d^c) = \frac{\wp(X, d)}{c}$ for any $c \in (0, 1]$.*

Proof. Again examining inequality [\(2.1.1\)](#), we see that (X, d^c) has p -negative type if and only if (X, d) has (cp) -negative type, and the result follows. \square

2.1.5 Metric subspaces

Definition 2.1.16. Let (X, d) be a metric space and let Y be a nonempty subset of X . Then $(Y, d|_{Y \times Y})$ is said to be a metric subspace of (X, d)

The following easy fact relates the maximum generalised roundness of a metric space to that of its metric subspaces.

Theorem 2.1.17. *Let (X, d) be a metric space and let $(Y, d|_{Y \times Y})$ be a metric subspace. Then $\wp(X, d) \leq \wp(Y, d|_{Y \times Y})$.*

Proof. If the generalised roundness p inequality [\(2.1.2\)](#) holds for any choice of points in X , then it remains true if we restrict our choice of $\{a_i\}_{i=1}^n$ and $\{b_j\}_{j=1}^n$ to only points in Y . It then follows that the maximum generalised roundness of (X, d) is no greater than that of $(Y, d|_{Y \times Y})$ as required. \square

Remark 2.1.18. It is immediate from the definition of maximum generalised roundness that if there is an isometric bijection between two metric spaces (X_1, d_1) and (X_2, d_2) then $\wp(X_1, d_1) = \wp(X_2, d_2)$. We will sometimes abuse the above terminology slightly by saying that (Y, d') is itself a metric subspace of (X, d) if it is isometric to a metric subspace of X . In particular, in that situation one has

$$\wp(X, d) \leq \wp(Y, d').$$

This result shows in particular that it is impossible to isometrically embed a metric space into one with larger maximum generalised roundness. We will use this result repeatedly in later chapters to give upper bounds on the value of $\wp(X, d)$ by showing the existence of suitable metric subspaces Y .

2.1.6 Application to classical spaces

One might start by determining the maximum generalised roundness of the real line \mathbb{R} . This requires only a small computation and was known to Enflo and Schoenberg in terms of generalised roundness and p -negative type respectively.

Theorem 2.1.19. *The maximum generalised roundness of the real line with the usual metric is $\wp(\mathbb{R}, |\cdot|) = 2$.*

Proof. This proof is reproduced from [51, Corollary 2.6].

Choosing $n = 2$, we let $a_1 = 0$ and $a_2 = 2$, and let $b_1 = b_2 = 1$ to obtain

$$\sum_{\substack{i,j=1 \\ i < j}}^2 |a_i - a_j|^p + \sum_{\substack{i,j=1 \\ i < j}}^2 |b_i - b_j|^p = 2^p$$

and

$$\sum_{i,j=1}^2 |a_i - b_j|^p = 4.$$

Then (2.1.3) fails for $p > 2$ and hence $\wp(\mathbb{R}, |\cdot|) \leq 2$.

For the other direction, we demonstrate that $(\mathbb{R}, |\cdot|)$ has generalised roundness 2. Simplifying (2.1.2), we must prove that

$$\sum_{\substack{i,j=1 \\ i < j}}^n (a_i - a_j)^2 + \sum_{\substack{i,j=1 \\ i < j}}^n (b_i - b_j)^2 \leq \sum_{i,j=1}^n (a_i - b_j)^2$$

holds for arbitrary $\{a_i\}_{i=1}^n$ and $\{b_j\}_{j=1}^n$ in \mathbb{R} . Rearranging, we have

$$\begin{aligned}
& \sum_{i,j=1}^n (a_i - b_j)^2 - \left[\sum_{\substack{i,j=1 \\ i < j}}^n (a_i - a_j)^2 + \sum_{\substack{i,j=1 \\ i < j}}^n (b_i - b_j)^2 \right] \\
&= \left[n \sum_{i=1}^n (a_i^2 + b_i^2) - 2 \left(\sum_{i,j=1}^n a_i b_j \right) \right] - \left[(n-1) \sum_{i=1}^n (a_i^2 + b_i^2) - 2 \sum_{\substack{i,j=1 \\ i < j}}^n (a_i a_j + b_i b_j) \right] \\
&= \sum_{i=1}^n (a_i^2 + b_i^2) + \sum_{\substack{i,j=1 \\ i \neq j}}^n (a_i a_j + b_i b_j) - 2 \left(\sum_{i,j=1}^n a_i b_j \right) \\
&= \left(\sum_{i=1}^n a_i \right)^2 + \left(\sum_{i=1}^n b_i \right)^2 - 2 \left(\sum_{i=1}^n a_i \right) \left(\sum_{i=1}^n b_i \right) \\
&\geq 0
\end{aligned}$$

as required, so $\wp(\mathbb{R}, |\cdot|) = 2$. □

Remark 2.1.20. It is worth noting that the first part of this proof together with Theorem [2.1.17](#) imply that if (X, d) is any metric space in which there are three points $x, y, z \in X$ where y is the midpoint between x and z , then $\wp(X, d) \leq 2$. For example, this bound applies to any normed space. It will similarly apply to most of the metric spaces arising from graphs which we will study later.

A second generalisation of the proof of the above theorem can be seen in the following result from Enflo, with complex numbers in place of the reals.

Theorem 2.1.21 ([\[31\]](#) Theorem 1]). *Let μ be a positive measure. Then the maximum generalised roundness of $L_2(\mu)$ is 2.*

Proof. As noted above in Remark [2.1.20](#), we must have $\wp(L_2(\mu), |\cdot|) \leq 2$.

Suppose $\{a_i\}_{i=1}^n, \{b_j\}_{j=1}^n \in L_2(\mu)$. Then using the previous result pointwise, we have

$$\begin{aligned} \sum_{\substack{i,j=1 \\ i < j}}^n d(a_i, a_j)^2 + \sum_{\substack{i,j=1 \\ i < j}}^n d(b_i, b_j)^2 &= \sum_{\substack{i,j=1 \\ i < j}}^n \int |a_i(\omega) - a_j(\omega)|^2 d\mu + \sum_{\substack{i,j=1 \\ i < j}}^n \int |b_i(\omega) - b_j(\omega)|^2 d\mu \\ &\leq \sum_{i,j=1}^n \int |a_i(\omega) - b_j(\omega)|^2 d\mu \\ &= \sum_{i,j=1}^n d(a_i, b_j)^2. \end{aligned}$$

Thus $\wp(L_2(\mu), |\cdot|) \geq 2$ also, completing the proof. \square

Lennard, Tonge and Weston [51] credited the generalisation to L_q spaces where $1 \leq q < 2$ to lectures also delivered by Enflo, and proved the extension to $0 < q < 1$. Note that for $q < 1$, the usual distance $d_q(f, g) = \|f - g\|_q$ does not satisfy the triangle inequality, but rather the weaker property

$$d_q(f, g) \leq 2^{\frac{1}{q}-1} (d_q(f, h) + d_q(h, g)).$$

Thus L_q is only a quasimetric space. The earlier definitions of p -negative type and generalised roundness p require no further amendment to cover quasimetric spaces, and Theorem 2.1.17 extends naturally.

The same result was earlier proven in terms of p -negative type by Schoenberg [74].

Theorem 2.1.22 ([51, Corollary 2.6]). *Let μ be a positive measure, and let $0 < q \leq 2$. Then*

$$\wp(L_q(\mu)) = q.$$

A surprising result of [51] is that this behaviour changes drastically beyond $q = 2$.

Theorem 2.1.23 ([51, Theorem 2.8]). *If $L_q(\mu)$ has dimension at least three then*

$$\wp(L_q(\mu)) = 0$$

for $2 < q \leq \infty$.

Remark 2.1.24. Yost [84] provided a simple proof that all two dimensional normed spaces embed isometrically in L_1 , so they have generalised roundness 1 by Theorem 2.1.17.

Remark 2.1.25. The results of Theorem 2.1.23 for finite dimensional sequence spaces $\ell_q^{(n)}$, where again $n \geq 3$, were previously known, with the $q = \infty$ case as a result of Misiewicz [59] and the $2 < q < \infty$ case from Koldobsky [49].

Remark 2.1.26. The original proof of Theorem 2.1.23 is rather indirect. It does not proceed by finding suitable families of points which must violate the generalised roundness inequalities or the p -negative type inequalities, nor does the proof give any insight as to what such sets might look like. Doust, Sánchez and Weston [26] provided a constructive proof in the case of three-dimensional ℓ_∞ space, that is, $\wp(\ell_\infty^{(3)}) = 0$, and noted that this conclusion applies also to metric spaces containing an isometric copy of $\ell_\infty^{(3)}$, such as c_0 and $C[0, 1]$.

These results have been used together with Theorem 2.1.17 to disprove embeddings. It is clear from Theorem 2.1.22 that a metric space (X, d) must have q -negative type in order to embed isometrically into L_q space ($0 < q \leq 2$).

Bretagnolle, Dacunha-Castelle and Krivine [11] proved that this condition is in fact sufficient for real normed spaces, and even quasinormed spaces.

Definition 2.1.27. Let X be a real vector space. A function $\|\cdot\| : X \rightarrow [0, \infty)$ is a quasinorm if

- $\|x\| \geq 0$ for all $x \in X$,
- $\|\lambda x\| = |\lambda| \|x\|$ for all $x \in X$, $\lambda \in \mathbb{R}$, and
- there exists $K \geq 1$ such that $\|x + y\| \leq K (\|x\| + \|y\|)$ for all $x, y \in X$.

Then $(X, \|\cdot\|)$ is a real quasinormed space.

Theorem 2.1.28 ([11, Theorem 2]). *Let $(X, \|\cdot\|)$ be a real quasinormed space. Then $(X, \|\cdot\|)$ is linearly isometric to a subspace of some L_q space if and only if it has q -negative type.*

Corollary 2.1.29 ([11]). *Let $0 < q_1 \leq q_2 \leq 2$. Then there is a linear isometric embedding of L_{q_2} into L_{q_1} .*

2.1.7 The p -negative type gap

The above results show that one may give an upper bound for $\wp(X, d)$ by choosing a suitable subset Y with known maximum generalised roundness. Proving good lower bound results for the maximum generalised roundness is significantly more difficult.

Hjorth et al. [47] showed that certain classes of metric spaces, including all metric trees, must have maximum generalised roundness at least 1. Doust and Weston [28] improved this lower bound using a concept called the p -negative type gap, which quantifies the sharpness of the generalized roundness inequality. Importantly these works gave a link between graph theoretic structures of a metric space X , and the geometric quantity $\wp(X)$.

Definition 2.1.30. Let (X, d) be a metric space of p -negative type. Let $S(\vec{\omega})$ be a normalised (q, t) -simplex in X , and define

$$\gamma_S^p(\vec{\omega}) = \sum_{i,j=1}^{q,t} m_i n_j d(a_i, b_j)^p - \sum_{1 \leq i_1 < i_2 \leq q} m_{i_1} m_{i_2} d(a_{i_1}, a_{i_2})^p - \sum_{1 \leq j_1 < j_2 \leq t} n_{j_1} n_{j_2} d(b_{j_1}, b_{j_2})^p.$$

The p -negative type gap of (X, d) is given by

$$\Gamma_X^p = \inf_{S(\vec{\omega})} \gamma_S^p(\vec{\omega}).$$

Clearly, (X, d) has generalized roundness p if and only if $\Gamma_X^p \geq 0$.

Doust and Weston also proved an equivalent formulation in terms of p -negative type.

Theorem 2.1.31 ([28, Remark 4.17]). *Let (X, d) be a metric space of p -negative type. The p -negative type gap Γ_X^p is the largest non-negative constant such that*

$$\frac{\Gamma_X^p}{2} \left(\sum_{\ell=1}^k |\alpha_\ell| \right)^2 + \sum_{1 \leq i, j \leq k} d(x_i, x_j)^p \alpha_i \alpha_j \leq 0$$

for all $k \geq 2$, $x_1, \dots, x_k \in X$ and $\alpha_1, \dots, \alpha_k \in \mathbb{R}$ with $\sum_{\ell=1}^k \alpha_\ell = 0$.

Geometrically, the 2-negative type gap for a subset of Euclidean space measures the minimum distance between the weighted barycentres of two disjoint finite collections of points, as observed by Deza and Maehara [24].

A natural dichotomy emerges between values of p for which $\Gamma_X^p = 0$, so the inequalities of interest are already sharp, and those where $\Gamma_X^p > 0$. It is clear that if the gap is positive, then (X, d) has strict p -negative type. The converse is not trivial, and was proven to hold for finite spaces by Li and Weston.

Theorem 2.1.32 ([52, Theorem 4.1]). *Let (X, d) be a finite metric space of p -negative type. Then (X, d) has strict p -negative type if and only if $\Gamma_X^p > 0$.*

Remark 2.1.33. An infinite metric space (X, d) may have strict p -negative type even if $\Gamma_X^p = 0$. Doust and Weston [28, Theorem 5.7] provide an example of an ‘infinite necklace’ exhibiting this property for $p = 1$.

Li and Weston also proved that the gap can be directly used to extend the range of p for which (X, d) is known to have (strict) p -negative type.

Theorem 2.1.34 ([52, Theorem 3.3]). *Let (X, d) be a finite metric space with $n = |X| \geq 3$, of strict p -negative type, with diameter*

$$\text{diam}(X) = \max_{x, y \in X} d(x, y)$$

and scaled diameter

$$\mathcal{D}(X) = \frac{\text{diam}(X)}{\min_{\substack{x, y \in X \\ x \neq y}} d(x, y)}.$$

Then (X, d) is also of strict q -negative for all $q \in [p, p + \xi)$ where

$$\xi = \frac{\log \left(1 + \frac{\Gamma_X^p}{\text{diam}(X)^p \gamma(n)} \right)}{\log \mathcal{D}(X)},$$

and

$$\gamma(n) = 1 - \frac{1}{2} \left(\frac{1}{\lfloor n/2 \rfloor} + \frac{1}{\lceil n/2 \rceil} \right).$$

Corollary 2.1.35 ([52, Corollary 4.3]). *Let (X, d) be a finite metric space with maximum generalised roundness p . Then (X, d) does not have strict p -negative type.*

Remark 2.1.36. This result does not extend to infinite metric spaces, as exemplified by the counterexample discussed in Remark 2.1.33, where $\wp(X) = 1$.

Weston used the method of Lagrange multipliers to compute the 0-negative type gap of an n -point metric space, which is clearly independent of the metric.

Theorem 2.1.37 ([78, Theorem 3.2]). *Let (X, d) be a metric space on $n \geq 2$ points. The 0-negative type gap of X is given by*

$$\Gamma_X^0 = \frac{1}{2} \left(\frac{1}{\lfloor n/2 \rfloor} + \frac{1}{\lceil n/2 \rceil} \right).$$

Remark 2.1.38. For the discrete metric space on n points, the distances within the metric transform X^p are independent of p , and so is the p -negative type gap. Thus

$$\Gamma_X^p = \frac{1}{2} \left(\frac{1}{\lfloor n/2 \rfloor} + \frac{1}{\lceil n/2 \rceil} \right)$$

for all $p \geq 0$.

Remark 2.1.39 ([78, Corollary 3.4]). Letting $n \rightarrow \infty$ in Theorem 2.1.37, we see that the 0-negative type gap of any infinite metric space is zero, in contrast to finite metric spaces where $\Gamma_X^0 > 0$.

More generally, Wolf provided a formula to compute the p -negative type gap of a finite metric space, and demonstrates several applications to well-known spaces.

Theorem 2.1.40 ([82, Theorem 3.5]). *Let (X, d) be a finite metric space of strict p -negative type. Let $n = |X|$, and define $D_p = (d(x_i, x_j)^p)_{i,j=1}^n$ and $\mathbb{1} = (1, 1, \dots, 1)$ in \mathbb{R}^n . The p -negative type gap of X is given by*

$$\Gamma_X^p = \frac{2}{\beta},$$

where

$$\beta = \max_{x \in [-1, 1]^n} \langle Bx, x \rangle = 4 \max_{x \in [0, 1]^n} \langle Bx, x \rangle$$

and

$$B = \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle^{-1} (D_p^{-1} \mathbb{1}) (D_p^{-1} \mathbb{1})^T - D_p^{-1}.$$

Example 2.1.41 ([82, Example 3.6]). Recall from Example 1.1.8 that the metric space X depicted in Figure 2.1 has p -negative type for $p \leq \log_2 3$. It follows from Theorem 2.1.4 that the strict p -negative type condition holds for $p < \log_2 3$.

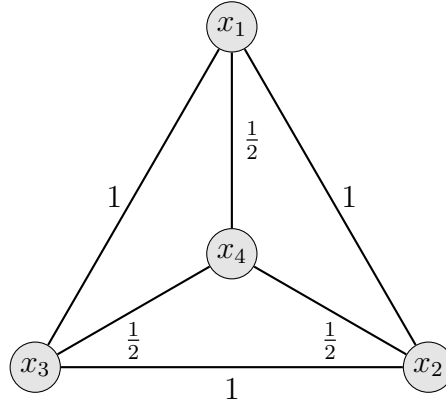


Figure 2.1: A four-point metric space.

Letting $\alpha = 2^{-p}$, the p -distance matrix is

$$D_p = \begin{pmatrix} 0 & 1 & 1 & \alpha \\ 1 & 0 & 1 & \alpha \\ 1 & 1 & 0 & \alpha \\ \alpha & \alpha & \alpha & 0 \end{pmatrix},$$

so the inverse is

$$D_p^{-1} = \frac{1}{3\alpha^2} \begin{pmatrix} -2\alpha^2 & \alpha^2 & \alpha^2 & \alpha \\ \alpha^2 & -2\alpha^2 & \alpha^2 & \alpha \\ \alpha^2 & \alpha^2 & -2\alpha^2 & \alpha \\ \alpha & \alpha & \alpha & -2 \end{pmatrix}.$$

Routine calculations show that

$$\begin{aligned} B &= \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle^{-1} (D_p^{-1} \mathbb{1}) (D_p^{-1} \mathbb{1})^T - D_p^{-1} \\ &= \frac{1}{2(3\alpha - 1)} \begin{pmatrix} 4\alpha - 1 & 1 - 2\alpha & 1 - 2\alpha & -1 \\ 1 - 2\alpha & 4\alpha - 1 & 1 - 2\alpha & -1 \\ 1 - 2\alpha & 1 - 2\alpha & 4\alpha - 1 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix}. \end{aligned}$$

The value of

$$\beta = \max_{x \in [-1,1]^4} \langle Bx, x \rangle$$

must occur at one of the “corners” of the 4-cube, so one can simply evaluate $\langle Bx, x \rangle$ at a small number of points to show that

$$\Gamma_X^p = \begin{cases} \frac{3}{4} - \frac{1}{4}2^{-p} & p \leq 2 - \log_2 3 \\ 2^{-p} - \frac{1}{3} & 2 - \log_2 3 \leq p < \log_2 3. \end{cases}$$

□

Example 2.1.42 ([82, Corollary 4.1]). Let $n \geq 3$, $C_n = \{x_1, \dots, x_n\}$ and

$$d(x_i, x_j) = \min(|i - j|, n - |i - j|),$$

representing the smallest number of steps between vertices i and j of an n -sided polygon. The example of C_7 is depicted in Figure 2.2.

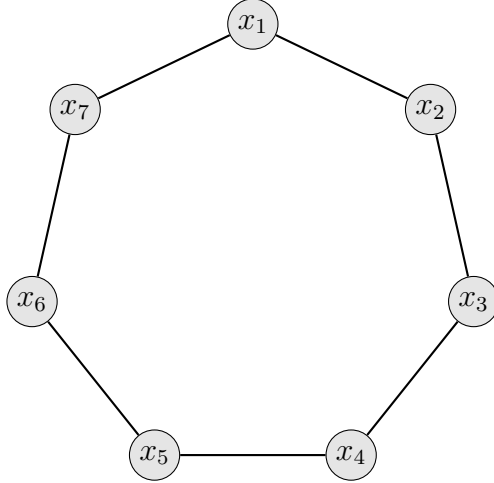


Figure 2.2: The seven point metric space C_7 .

Wolf proved that the 1-negative type gap is

$$\Gamma_X^1 = \begin{cases} 0 & \text{if } n \text{ is even} \\ \frac{1}{2} \frac{n}{n^2 - 2n - 1} & \text{if } n \text{ is odd.} \end{cases}$$

□

Using this method, computation of the p -negative type gap for a given n -point metric space requires us to consider 2^{n-1} assignments of -1 and 1 to the components of x , which is computationally difficult as n becomes large.

2.2 Graph metric spaces

A large and important family of metric spaces are those which, like Example [2.1.42](#), are associated with connected graphs. The work of Hjorth et al. [\[47\]](#) and Doust and Weston [\[28\]](#) indicate that there are some unexpected connections between purely graph theoretic properties, and the geometric constants considered in the previous section. A starting point for the research presented here was to examine whether there are other similar connections to be uncovered. In particular we will be looking at the typical geometric properties of metric spaces constructed from random graphs of various types.

2.2.1 Fundamentals of graph theory

We will first formalise the main ideas and terminology from graph theory that we will be using. The reader is encouraged to consult Diestel [\[25\]](#) for further details.

Definition 2.2.1. A graph is a pair of sets $G = (V, E)$, where V is a nonempty set whose elements are called vertices and E is a set of unordered pairs of distinct vertices, whose elements are called edges.

Remark 2.2.2. Throughout this thesis, all graphs are undirected. In addition, no duplicate edges or loops are permitted, so our graphs are simple.

Remark 2.2.3. A graph is said to be finite if its vertex set is finite, and infinite otherwise. In this thesis, all graphs will be finite except in Chapter [6](#).

Definition 2.2.4. Let $G = (V, E)$ be a graph. The order of G is the number of vertices $|V|$, which we will denote by $|G|$.

Definition 2.2.5. Let $G = (V, E)$ be a graph, and let $v \in V$ be a vertex. The degree of v is the number of edges incident to v , that is,

$$\deg v = |\{e \in E : v \in e\}|.$$

Within a graph, we can form a subgraph by selecting some or all of the vertices and a subset of the corresponding edges. Subgraphs will be vital in our understanding of the structure of a graph.

Definition 2.2.6. Let $G = (V, E)$ be a graph. A subgraph of G is a graph $G' = (V', E')$ where $V' \subseteq V$ and $E' \subseteq E$.

Traversing several edges in succession forms a path, allowing us to describe indirect connections between pairs of vertices.

Definition 2.2.7. A path $P = v_0v_1v_2 \dots v_{k-1}v_k$ is comprised of a sequence of edges $\{v_0, v_1\}, \{v_1, v_2\}, \dots, \{v_{k-1}, v_k\}$, where the intermediate vertices v_1, \dots, v_{k-1} are distinct. We can denote subpaths of P as follows:

$$\begin{aligned} Pv_j &= v_0v_1 \dots v_{j-1}v_j \\ v_iP &= v_iv_{i+1} \dots v_{k-1}v_k \\ v_iPv_j &= v_iv_{i+1} \dots v_{j-1}v_j. \end{aligned}$$

Definition 2.2.8. A cycle is a path $v_0v_1v_2 \dots v_{k-1}v_0$ consisting of three or more edges.

Definition 2.2.9. A graph is acyclic if it has no cycles.

Definition 2.2.10. A graph is connected if there is a path joining each pair of distinct vertices.

We can now define path lengths within the graph, by taking the shortest path available between each pair of vertices.

Definition 2.2.11. Let $G = (V, E)$ be a connected graph. Define the path metric $d : V \times V \rightarrow \mathbb{N}$ as follows: for $u, v \in V$, let $d(u, v)$ be the minimum number of edges in a path between u and v .

Note that the graph must be connected to ensure that $d(u, v)$ is defined for all pairs of vertices.

Proposition 2.2.12. *The path metric satisfies the definition of a metric, so (G, d) is a metric space.*

Notation 2.2.13. We denote the metric space (G, d) simply as G , and refer to it as a path metric graph.

The path metric is not the only metric with which the vertices of a graph can be endowed. Other notions of distance can be introduced by adding a weight to each edge.

Definition 2.2.14. Let $G = (V, E)$ be a graph. Let $w : E \rightarrow (0, \infty)$ be a function that assigns a weight to each edge. We refer to (G, w) as a weighted graph, and often write $|e|$ for the weight of the edge.

Proposition 2.2.15. *In a connected weighted graph (G, w) , let $d_w(u, v)$ be the minimum total weight of a path between u and v . This also satisfies the definition of a metric.*

We may therefore refer to (G, w) as a metric graph. Note that if all edge weights are set to 1, we recover the path metric d .

Remark 2.2.16. Observe that any metric space can be represented by an appropriately weighted complete graph.

2.2.2 Trees

As with many graph problems, the more restricted setting of trees allows us to solve problems more easily due to the additional structure. Several key results on the maximum generalised roundness properties of trees are well known in the field, from the work of Doust and Weston [28] among others. We will present and extend these results in Chapter 4, and both the methods and outcomes will motivate further study among general graphs in Chapters 5 and 6.

We begin our preliminaries with one of many equivalent definitions of a tree.

Definition 2.2.17. A tree is a connected acyclic graph.

Doust and Weston [28] made important deductions by partitioning the vertices of a tree about a fixed edge, and analysing the distances both across the partition and on each side. The following fundamental result is crucial to this approach.

Proposition 2.2.18. *Let T be a tree and $v, w \in V(T)$ be two distinct vertices of T . There is a unique path between v and w .*

Proof. As T is connected, we are guaranteed the existence of one such path by Definition 2.2.10.

Suppose that two paths exist, say $P = vu_1u_2 \dots u_kw$ and $P' = vu'_1u'_2 \dots u'_\ell w$. We will prove the presence of a cycle. First, we remove duplicated edges from the start of both paths, by traversing both paths together until the last common vertex $u_i = u'_i = x$. We then have subpaths xP and xP' , noting that the second vertices of the two paths are $u_{i+1} \neq u'_{i+1}$. Next, let y be the first vertex of $u_{i+1}P$ to appear in $u'_{i+1}P'$ also. Note that some such y must exist as w is common to both paths. This leaves two subpaths xPy and $xP'y$, which cannot have any intermediate vertices in common by the definition of y . Thus the path formed by xPy followed by the reverse of $xP'y$ is a cycle in T , contradicting the definition of a tree, so exactly one path exists between v and w . \square

It is well known that a finite tree has exactly one fewer edge than it has vertices.

Proposition 2.2.19. *A tree on n vertices has exactly $n - 1$ edges.*

A proof can be found in [25, Corollary 1.5.3]

The connected subgraphs of a tree are themselves acyclic, so they are also trees.

Definition 2.2.20. Let T be a tree. A subtree is a connected subgraph of T .

2.2.3 Examples

For some applications it is desirable to be able to embed a metric graph isometrically, or with some distortion, in a Euclidean space \mathbb{R}^n . One would therefore want to know its maximum generalised roundness. We begin with some of the most elementary graphs.

Example 2.2.21. Let K_n be the unweighted complete graph on n vertices. Then for $v, w \in K_n$,

$$d(v, w) = \begin{cases} 1 & \text{if } v \neq w \\ 0 & \text{if } v = w, \end{cases}$$

so $d(v, w)^p$ is independent of p . Since K_n has generalised roundness 0 by Corollary 2.1.37, it must also have generalised roundness p for all $p > 0$, so $\wp(K_n) = \infty$. \square

This graph metric space corresponds exactly to the discrete metric space on n points. This example is in fact atypical, and other graphs exhibit very different maximum generalised roundness properties, as alluded to in Remark 2.1.20.

Theorem 2.2.22. *Let G be an unweighted connected graph on at least three vertices which is not complete. Then $\wp(G) \leq 2$.*

Proof. Suppose that G has generalised roundness p . As G is connected but not complete, there must exist three vertices $v_1, v_2, v_3 \in V$ so that $\{v_1, v_2\}, \{v_2, v_3\} \in E$ but $\{v_1, v_3\} \notin E$, so $d(v_1, v_3) = 2$. Applying Definition 2.1.8 with $n = 2$, and choosing $a_1 = v_1$, $a_2 = v_3$ and $b_1 = b_2 = v_2$, the generalised roundness p inequality (2.1.3) gives $2^p \leq 4$ and hence $p \leq 2$. \square

A similar result also applies to weighted graphs. The following example of a three vertex graph with two weighted edges is crucial in establishing the more general bound.

Example 2.2.23. Let $a \in (0, 1)$ be fixed, and let (G, w) be a weighted graph on three vertices as in Figure 2.3. We claim that $\wp(G, w) = 2$.

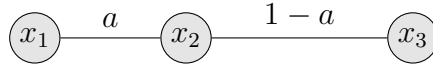


Figure 2.3: A graph on three vertices.

We begin by constructing the corresponding metric space (X, d) , with point set $X = \{x_1, x_2, x_3\}$ and distances

$$\begin{aligned} d(x_1, x_2) &= a, \\ d(x_2, x_3) &= 1 - a, \\ d(x_1, x_3) &= 1. \end{aligned}$$

Note that Example 2.2.21 implies that any two-point subspace has generalised roundness p for any $p \in [0, \infty)$, so it remains to test the simplices involving all three points.

Consider the normalised $(2, 1)$ -simplex $S = [a_i(m_i); b_j(n_j)]_{2,1}$ given by

$$m_1 = t, m_2 = 1 - t, n_1 = 1 \quad \text{and} \quad a_1 = x_1, \quad a_2 = x_3, \quad b_1 = x_2,$$

where $t \in (0, 1)$. For this simplex, the generalised roundness p inequality (2.1.3) reduces to

$$t(1 - t) \leq ta^p + (1 - t)(1 - a)^p. \quad (2.2.1)$$

By symmetry, the only other non-trivial $(2, 1)$ -simplex to consider is given by

$$m_1 = t, m_2 = 1 - t, n_1 = 1 \quad \text{and} \quad a_1 = x_1, \quad a_2 = x_2, \quad b_1 = x_3,$$

where $t \in (0, 1)$. Here, (2.1.3) reduces to

$$t(1 - t)a^p \leq t + (1 - t)(1 - a)^p. \quad (2.2.2)$$

Comparing to (2.2.1), the left-hand side has been reduced and the right-hand side increased since $a, t \in (0, 1)$. Hence if some p satisfies (2.2.1) for all $t \in (0, 1)$, then it also satisfies (2.2.2) for all $t \in (0, 1)$. Thus we need only consider the first $(2, 1)$ -simplex to determine the values of p for which (X, d) has generalised roundness p .

First, we observe that substituting $p = 2$ reduces (2.2.1) to $(a + t - 1)^2 \geq 0$, which holds for all $t \in (0, 1)$ with equality for $t = 1 - a$. Therefore (X, d) has 2-negative type, but does not have strict 2-negative type, so by Corollary 2.1.5 we have $\wp(X, d) = 2$. We will therefore write $\wp(G, w) = 2$ also. \square

It follows that whenever the graph G discussed in Example 2.2.23 occurs as a metric subspace of a larger metric space H , we can bound $\wp(H)$ using Theorem 2.1.17.

Proposition 2.2.24. *For any metric space (X, d) with $|X| \geq 3$, in which three points $x_1, x_2, x_3 \in X$ satisfy*

$$d(x_1, x_3) = d(x_1, x_2) + d(x_2, x_3),$$

we have $\wp(X, d) \leq 2$.

Such a triple occurs in every weighted graph, unless the graph is complete and has no edge equal in weight to another path joining its endpoints. Note however

that this condition cannot be removed; in fact we can easily construct a metric space with $\wp(X, d) = p$ for any $2 < p \leq \infty$.

Remark 2.2.25. Let (G, w) be the weighted graph depicted in Figure 2.4, where $a \in [1, 2)$ is fixed.

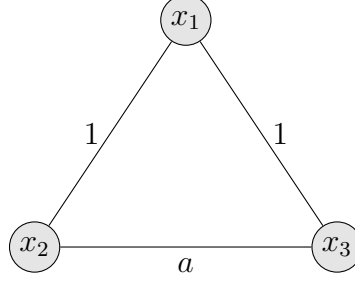


Figure 2.4: A graph on three vertices.

Let (X, d) be the corresponding metric space. Again, we need only consider two nontrivial normalised $(2, 1)$ -simplices.

The first is $S_1 = [a_i(m_i); b_j(n_j)]_{2,1}$ given by

$$m_1 = t, m_2 = 1 - t, n_1 = 1 \quad \text{and} \quad a_1 = x_2, \quad a_2 = x_3, \quad b_1 = x_1,$$

where $t \in (0, 1)$. For this simplex, the generalised roundness p inequality (2.1.3) reduces to

$$t(1 - t)a^p \leq t + (1 - t) = 1. \quad (2.2.3)$$

The other simplex is $S_2 = [a_i(m_i); b_j(n_j)]_{2,1}$ given by

$$m_1 = t, m_2 = 1 - t, n_1 = 1 \quad \text{and} \quad a_1 = x_1, \quad a_2 = x_2, \quad b_1 = x_3,$$

where $t \in (0, 1)$. Here, (2.1.3) simplifies to

$$t(1 - t) \leq t + (1 - t)a^p. \quad (2.2.4)$$

Again comparing with (2.2.3), since $a > 1$, the left-hand side has been reduced and the right-hand side increased. Therefore we need only consider whether a value of p satisfies (2.2.3).

Rearranging (2.2.3), we have

$$p \leq \log_a \frac{1}{t(1 - t)} \leq \log_a 4,$$

so

$$\wp(X, d) = \wp(G, w) = \log_a 4,$$

which can take any value in $(2, \infty]$ as a ranges over $[1, 2)$.

For some families of graphs with very rigid structure, we can explicitly find the maximum generalised roundness. Two examples are detailed below.

Example 2.2.26. Let $n \geq 3$, and let P_n be a graph with vertex set $\{v_1, v_2, \dots, v_n\}$ and edge set $\{\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{n-1}, v_n\}\}$. We refer to P_n as the path graph on n vertices. Figure 2.5 depicts P_5 .

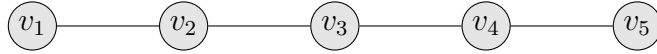


Figure 2.5: The path graph on 5 vertices, P_5 .

Recall from Example 2.2.23 that $\wp(P_3) = 2$. Noting that P_3 is a metric subspace of P_n for all $n \geq 3$, we have $\wp(P_n) \leq 2$ by Theorem 2.1.17. This result can alternatively be seen as an application of Remark 2.2.22.

However, P_n itself is a metric subspace of \mathbb{R} with the usual metric. Recall from Theorem 2.1.19 that $\wp(\mathbb{R}) = 2$, so using Theorem 2.1.17 again, we are able to deduce that $\wp(P_n) \geq 2$.

Combining the two results, we have $\wp(P_n) = 2$ for all $n \geq 3$. □

Example 2.2.27 ([28, Theorem 5.6]). Let $n \geq 3$, and let S_n be a graph with vertex set $\{v_0, v_1, \dots, v_{n-1}\}$ and edge set $\{\{v_0, v_1\}, \{v_0, v_2\}, \dots, \{v_0, v_{n-1}\}\}$. We refer to S_n as the star graph on n vertices. Figure 2.6 depicts S_6 .

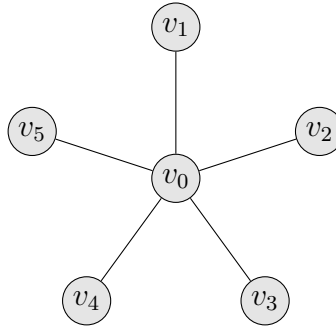


Figure 2.6: The star graph on 6 vertices, S_6 .

To find the maximum generalised roundness, Doust and Weston consider a normalised (q, t) -simplex $S = [a_i(m_i); b_j(n_j)]_{q,t}$ in S_n .

First, if v_0 does not appear in S , that is, $v_0 \neq a_i, b_j$, then the generalised roundness p inequality (2.1.3) holds trivially for all $p \geq 0$ as all distances are 2. Thus we proceed with the assumption that v_0 is one of the vertices of the simplex, say b_1 .

Next, we prove that it suffices to consider only simplices where $t = 1$. A simple calculation shows that if $t > 1$, S satisfies (2.1.3) only if the same is true of the $(q, t - 1)$ simplex formed by omitting b_2 from the simplex and adding its weight n_2 to that of b_1 . Repeating this transformation, the problem can be restricted to simplices where $t = 1$.

In this case, (2.1.3) reduces to

$$\left(\sum_{1 \leq i_1 < i_2 \leq q} m_{i_1} m_{i_2} \right) 2^p \leq \left(\sum_i^q m_i \right) 1^p = 1.$$

After manipulation, it can be seen that the coefficient of 2^p is at most $\frac{1}{2} \left(1 - \frac{1}{q} \right)$, which is maximised when $q = n - 1$, and hence

$$p \leq 1 + \log_2 \left(1 + \frac{1}{n-2} \right).$$

Thus we have

$$\wp = 1 + \log_2 \left(1 + \frac{1}{n-2} \right).$$

□

Note that the graphs in Examples 2.2.26 and 2.2.27 are examples of trees. Both have maximum generalised roundness strictly greater than 1, a special property of trees which will be explored in Chapter 4. In Chapter 5, we will demonstrate that such behaviour is rare among graphs in general.

For larger and more complex graphs, it is typically not possible to compute the maximum generalised roundness directly from the definitions and theorems presented earlier in this chapter. Chapter 3 details an alternative approach for these computations.

2.2.4 Graph counting

In Chapters 4 and 5, we will discuss the typical maximum generalised roundness properties of trees and graphs. Our focus will be on unweighted graphs, as there

are only finitely many of these on n vertices. There are two natural schemes for enumeration of these graphs, depending on the question of vertex labelling and graph isomorphisms.

Definition 2.2.28. Two graphs $G = (V, E)$ and $G' = (V', E')$ are said to be isomorphic if there is an edge-preserving bijection $f : V \rightarrow V'$, that is, $(v, w) \in E$ if and only if $(f(v), f(w)) \in E'$. We denote this relation $G \simeq H$.

Proposition 2.2.29. *Graph isomorphism (\simeq) is an equivalence relation on graphs on n vertices, and on trees on n vertices.*

One natural approach to count graphs is to label the vertices v_1, \dots, v_n , counting isomorphic graphs with multiplicity. The other method is to count graphs up to isomorphism, with unlabelled vertices, yielding the smaller number of isomorphism classes only. We primarily concern ourselves with the labelled case for two main reasons. Graph isomorphism is not a natural lens to apply in our study of maximum generalised roundness, as these properties are often determined entirely by a subgraph as in Theorem [2.1.17](#). Furthermore, it is simple enough to list all the graphs for small n , but the number of graphs in these families grows superexponentially. Instead we use established methods of randomly sampling from these families, which count graphs on labelled vertices.

CHAPTER 3

Calculation of maximum generalised roundness

For many years, calculating the maximum generalised roundness of a particular finite metric space was a difficult endeavour. Even in spaces with few points, it is difficult to directly apply the generalised roundness p inequality from Definition [2.1.8](#),

$$\sum_{\substack{i,j=1 \\ i < j}}^n d(a_i, a_j)^p + \sum_{\substack{i,j=1 \\ i < j}}^n d(b_i, b_j)^p \leq \sum_{i,j=1}^n d(a_i, b_j)^p,$$

as the points a_i and b_j may be repeated. Although the weighted version presented in Theorem [2.1.11](#) does not allow repetition, it introduces an additional real-valued variable for each point of the metric space, again making the problem of maximising p intractable in most cases. One encounters the same difficulty with the definition of p -negative type from Definition [2.1.1](#). However, an alternative approach was developed by Sánchez [\[71\]](#), who analysed the p -distance matrix, the elementwise p th power of the distance matrix.

In Section [3.1](#) we shall introduce Sánchez's formula, and discuss some of the issues involved in using this formula in practice. For most finite metric spaces $p(X, d)$ is a zero of a rather complicated function, and so does not have an expression in a closed form. Rather, Sánchez's formula allows one to find $p(X, d)$ numerically. For an individual small metric space, this is not too difficult to do using some human judgement. In order to calculate $p(X, d)$ for thousands of larger spaces, a robust algorithm is needed, and so most of this chapter is devoted to a discussion of the development and verification of such an algorithm.

At the end of the chapter we will look at some special structures where Sánchez's formula can be used to give an exact value of $p(X, d)$.

3.1 Calculation of maximum generalised roundness of individual graph metric spaces

We first introduce some notation used in Theorem [3.1.3](#)

Definition 3.1.1. Let (X, d) be a finite metric space, with points labelled x_1, \dots, x_n . For $p \geq 0$, define the p -distance matrix by

$$D_p = (d(x_i, x_j)^p)_{i,j}.$$

Notation 3.1.2. Denote by $\mathbb{1}$ a vector whose entries are all 1, so that the sum of the entries of a matrix A may be written as $\langle A\mathbb{1}, \mathbb{1} \rangle$.

Sánchez provided the following characterisation of strict p -negative type in terms of the p -distance matrix. Importantly, this criterion is much easier to verify than the definitions of p -negative type or generalised roundness p that were presented in Chapter [2](#)

Theorem 3.1.3 ([\[71\]](#), Theorem 2.3]. *Let (X, d) be a finite metric space of p -negative type, with p -distance matrix D_p . Then (X, d) has strict p -negative type if and only if*

$$\det D_p \neq 0 \quad \text{and} \quad \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle \neq 0.$$

Proof. Recall that the p -negative type condition in Definition [2.1.1](#) selects n points $\{x_1, \dots, x_n\}$ of X and assigns weights α_i to these points. For finite metric spaces, we can assign weight zero to the remaining points of X to obtain a weight vector $\alpha \in \mathbb{R}^{|X|}$ whose entries sum to zero. We denote the set of such vectors

$$\Pi_0 = \{\alpha \in \mathbb{R}^{|X|} : \langle \alpha, \mathbb{1} \rangle = 0\},$$

so (X, d) has p -negative type if and only if

$$\langle D_p \alpha, \alpha \rangle \leq 0$$

for all $\alpha \in \Pi_0$, with strict p -negative type if equality occurs when $\alpha = 0$ only. We similarly define $\Pi_1 = \{\alpha \in \mathbb{R}^{|X|} : \langle \alpha, \mathbb{1} \rangle = 1\}$.

Suppose (X, d) has strict p -negative type. We first show that D_p is not singular, that is, $D_p \alpha = 0$ has no non-trivial solutions. Suppose $\alpha \notin \Pi_0$. Then

$$v = \frac{\alpha}{\langle \alpha, \mathbb{1} \rangle} \in \Pi_1,$$

as is

$$w = \left(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0 \right),$$

so $w - v \in \Pi_0$. As (X, d) is of negative type, we have

$$\langle D_p(w - v), w - v \rangle \leq 0.$$

Since D_p is symmetric, we have $\langle D_p w, v \rangle = \langle D_p v, w \rangle$, and $D_p \alpha = 0$ implies that $\langle D_p v, v \rangle = 0$ also, so after expanding and simplifying we have

$$\langle D_p w, w \rangle \leq 0.$$

However,

$$\langle D_p w, w \rangle = \frac{1}{2} d(x_1, x_2)^p > 0,$$

so we have a contradiction and thus $\alpha \in \Pi_0$. Now, $\langle D_p \alpha, \alpha \rangle = \langle 0, \alpha \rangle = 0$, so for strict p -negative type, $\alpha = 0$ only, proving that $\det D_p = 0$.

Next, we show that $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle \neq 0$. Let $\beta = D_p^{-1} \mathbb{1}$, so we aim to prove $\beta \notin \Pi_0$. Assuming the contrary, we have $\langle D_p \beta, \beta \rangle = \langle \mathbb{1}, \beta \rangle = 0$. Again using the strict p -negative type condition, this can only be solved by $\beta = 0$, which is impossible, completing the proof.

Finally, suppose that (X, d) is of p -negative type with $\det D_p$ and $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ nonzero. We wish to prove that (X, d) is of strict p -negative type by showing that $\langle D_p \alpha, \alpha \rangle = 0$ only when $\alpha = 0$. As (X, d) is of p -negative type, the symmetric bilinear form $(-\langle D_p x, y \rangle)$ is a semi-inner product on Π_0 , so by the Cauchy-Schwarz inequality we have

$$|\langle D_p \alpha, \beta \rangle|^2 \leq |\langle D_p \alpha, \alpha \rangle| |\langle D_p \beta, \beta \rangle| = 0$$

for any $\beta \in \Pi_0$. We therefore have $\langle D_p \alpha, \beta \rangle = 0$, which can only hold for all such β if $D_p \alpha = \lambda \mathbb{1}$. Since D_p is non-singular, $\alpha = \lambda D_p^{-1} \mathbb{1}$, so $0 = \langle \alpha, \mathbb{1} \rangle = \lambda \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$, and thus $\lambda = 0$. It follows that $\alpha = 0$ as required, completing the proof. \square

Remark 3.1.4. Note in particular that the choice of labelling in Definition [3.1.1](#) does not influence the two quantities of interest in Theorem [3.1.3](#).

We can therefore formulate a one-variable expression for the maximum generalised roundness. This presents a great improvement from its earlier definition as the solution of an optimisation problem in $|X| + 1$ variables.

Corollary 3.1.5 ([71, Corollary 2.4]). *Let (X, d) be a finite metric space with p -distance matrix D_p . Then*

$$\wp(X, d) = \min\{p \geq 0 : \det D_p = 0 \text{ or } \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = 0\}.$$

Proof. The proof follows immediately from Theorem 3.1.3 and Theorem 2.1.35. \square

In principle, we can now calculate the maximum generalised roundness of any finite metric space, by solving for the first non-negative roots of the functions $\det D_p$ and $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$. In practice however, it is not at all straightforward to perform this calculation accurately and efficiently, especially for large metric spaces. In the following subsections, we will see that only numerical computation of these functions is feasible, and that several complications arise in our efforts to isolate and approximate the smallest zero of each function.

3.2 Computational issues

In this section we shall look more closely at the quantities that appear in Sánchez's formula and examine some of the problems that may be encountered in the application of this formula to metric spaces on n points.

3.2.1 Determinant of the p -distance matrix

Since the entries of D_p are of the form $(d(x_i, x_j))^p$, upon expansion we see that

$$\begin{aligned} \det D_p &= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n (d(x_i, x_{\sigma(i)})^p) \\ &= \sum_k c_k d_k^p, \end{aligned}$$

where each d_k is a product of distances between points of X . Evidently, solving $\det D_p = 0$ analytically is often infeasible or impossible.

Example 3.2.1. Consider the graph depicted in Figure 3.1.

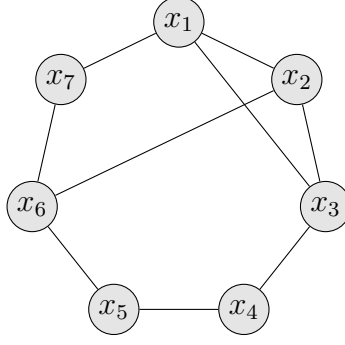


Figure 3.1: A graph on seven vertices.

The p -distance matrix is

$$D_p = \begin{pmatrix} 0 & 1 & 1 & 2^p & 3^p & 2^p & 1 \\ 1 & 0 & 1 & 2^p & 2^p & 1 & 2^p \\ 1 & 1 & 0 & 1 & 2^p & 2^p & 2^p \\ 2^p & 2^p & 1 & 0 & 1 & 2^p & 3^p \\ 3^p & 2^p & 2^p & 1 & 0 & 1 & 2^p \\ 2^p & 1 & 2^p & 2^p & 1 & 0 & 1 \\ 1 & 2^p & 2^p & 3^p & 2^p & 1 & 0 \end{pmatrix}.$$

Using Maple, we can evaluate the determinant

$$\begin{aligned} \det D_p = & 8 \times 9^p - 2 \times 32^p + 14 \times 2^p - 4 \times 3^p - 4 \times 27^p - 14 \times 6^p + 4 \times 36^p \\ & + 30 \times 12^p - 6 \times 18^p - 4 \times 72^p + 12 \times 24^p - 42 \times 48^p + 22 \times 96^p + 20 \times 144^p \\ & - 4 \times 192^p - 6 \times 288^p - 26 \times 64^p + 8 \times 128^p - 26 \times 4^p - 20 \times 8^p + 52 \times 16^p \\ & + 2 \times 162^p - 4 \times 216^p + 6 \times 54^p + 2 \times 432^p - 12 \times 108^p, \end{aligned}$$

which does not factorise further, so $\det D_p = 0$ cannot be solved analytically. \square

We are therefore required to solve numerically for the zeroes. Observe that our function $\det D_p$ is smooth, not only in this example but in general, so one can certainly use standard computer packages to find the zeros numerically.

Recall that Corollary [3.1.5](#) requires us to find the smallest positive zero of $\det D_p$. Numerical packages are usually very good at finding a zero of a function, but they are typically unable to find the smallest zero, or all zeros.^{[1](#)} For example, Maple's `fsolve` employs a combination of algorithms [\[46\]](#) including Newton's method to find

¹A notable exception is that there are well established methods to find all roots of a polynomial, easily accessible in computing packages such as Maple, which we will later make use of.

a zero within a specified interval, so in cases where we can symbolically calculate $\det D_p$, it is tempting to use the Maple command

$$\text{fsolve}(\text{Determinant}(D_p)=0, p=0..2); \quad (3.2.1)$$

Unfortunately, this frequently gives a zero which is not the smallest one.

Example 3.2.2. In Example 3.2.1, applying `fsolve` on $[0, 2]$ gives an answer of $1.3016\dots$, which is only the second smallest zero. To ensure correctness, it is crucial that we select a search interval containing only one zero, by first closely examining the graph of $\det D_p$.

```
with(GraphTheory):
with(LinearAlgebra):
V := [1,2,3,4,5,6,7]:
E := {{1,2},{2,3},{3,4},{4,5},{5,6},{6,7},{7,1},{1,3},{2,6}}:
G := Graph(V,E):
n := NumberOfVertices(G):
Dm := AllPairsDistance(G):
Dp := Matrix(n,n,(i,j) -> Dm[i,j]^p):
plot(Determinant(Dp),p=0..1);
```

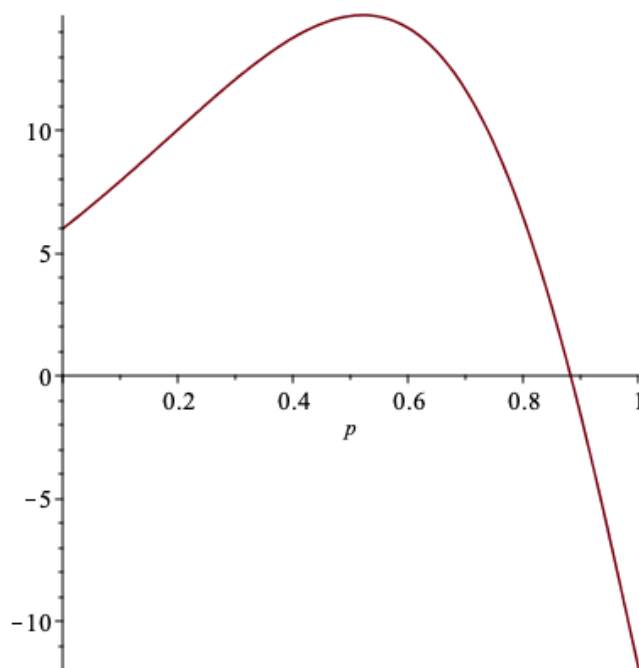


Figure 3.2: A plot of $\det D_p$ for the graph depicted in 3.1.

Assessing Figure 3.2 we can easily isolate the smallest zero within the interval $(0.8, 0.9)$, and we can obtain an approximate value for this zero using the command

```
fsolve(Determinant(Dp)=0,p=0.8..0.9);
```

to find $p \approx 0.8816$. □

For many of the problems that we wish to attack later in this thesis however, this ‘hands-on’ method of numerically calculating the smallest zero is not feasible.

First of all, it is computationally expensive to even compute the determinant symbolically, making the calculation prohibitively slow for larger metric spaces, as the number of terms may grow factorially in n in the worst case. For example, symbolic calculation of the determinant in Maple on a desktop PC typically takes less than one-tenth of a second for $n = 10$, a few seconds for $n = 15$, a few minutes for $n = 20$ and several hours for $n = 30$.

More importantly however, in later applications, we wish to perform this calculation for many metric spaces, in order to gather statistical data describing the maximum generalised roundness properties of various families of metric spaces. The solution outlined thus far is not suitable, as plotting and examining the graph of $\det D_p$ for each such metric space would be incredibly time consuming. We instead aim for a fast, robust procedure that can be applied without human intervention, allowing us to iterate through a family of metric spaces, or a random sample thereof, and calculate the maximum generalised roundness of each space in turn.

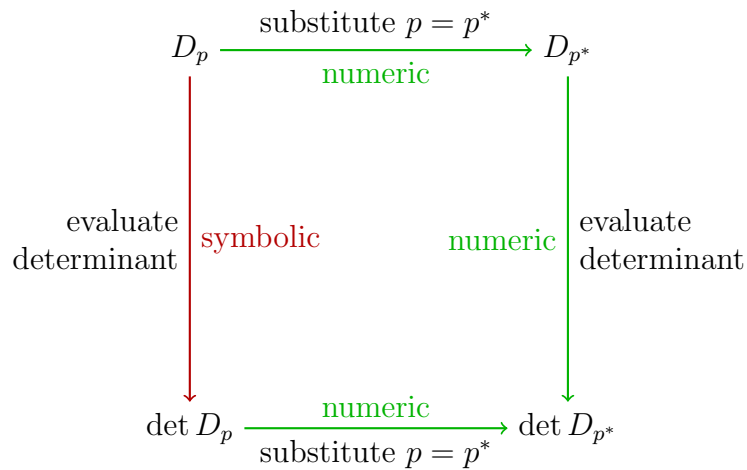


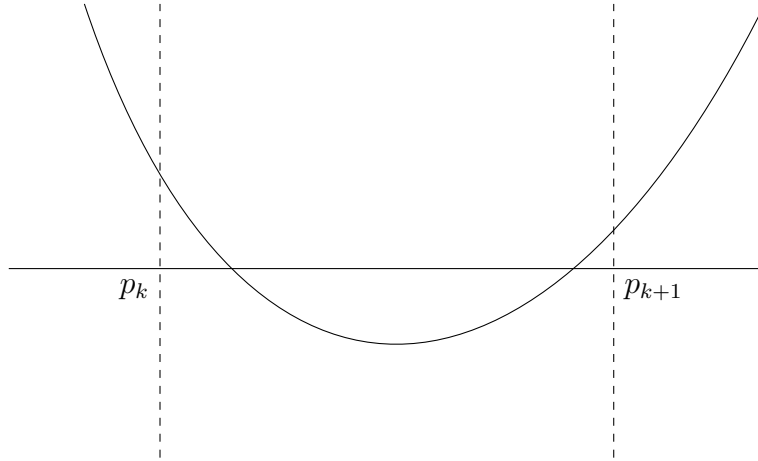
Figure 3.3: Symbolic computation of $\det D_p$ is much slower than repeatedly calculating D_p and its determinant numerically.

A much faster approach is to use numerical computations throughout, as seen in Figure 3.3. We consider $\det D_p$ as a function of p , and sample values of this function. For each value of p , the p -distance matrix is recalculated, and its determinant can be found by standard methods, such as the LU factorisation. In this way, even for $n = 100$ we can sample $\det D_p$ for one thousand values of p in about one minute in Maple, and faster still in MATLAB.

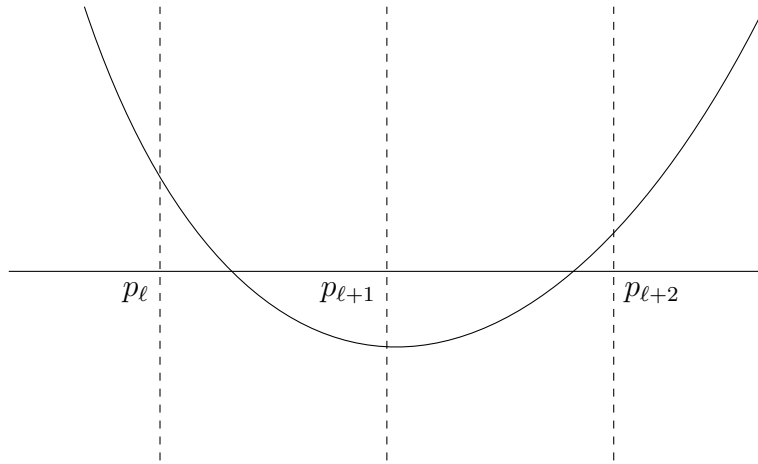
In Example 3.2.2, a human operator was required to visually identify an interval to search for the smallest zero. To eliminate this role, our algorithm must isolate this smallest zero. Recall that unless G is a complete graph, then $p(G) \leq 2$ by Proposition 2.2.22, so we are only concerned with values of p in the finite interval $[0, 2]$. Therefore we compute $\det D_p$ for several values $0 = p_0 < p_1 < \dots < p_m$ where $p_m > 2$, choosing $p_k = k\epsilon$ for simplicity. When the first change in sign of $\det D_p$ is found, we can search an interval of width ϵ for a zero.

However, this relies on our ability to programmatically identify a change in sign either side of the smallest zero. If two zeros lie close together, we may fail to find either of them, as displayed in Figure 3.4a. This can of course be rectified by sampling more frequently from the curve, as in Figure 3.4b. Unfortunately, it is hard to know a priori just how fine the sampling needs to be. There is also a tradeoff between speed and accuracy, as a smaller value of ϵ may require more computations before a change in sign can be found.

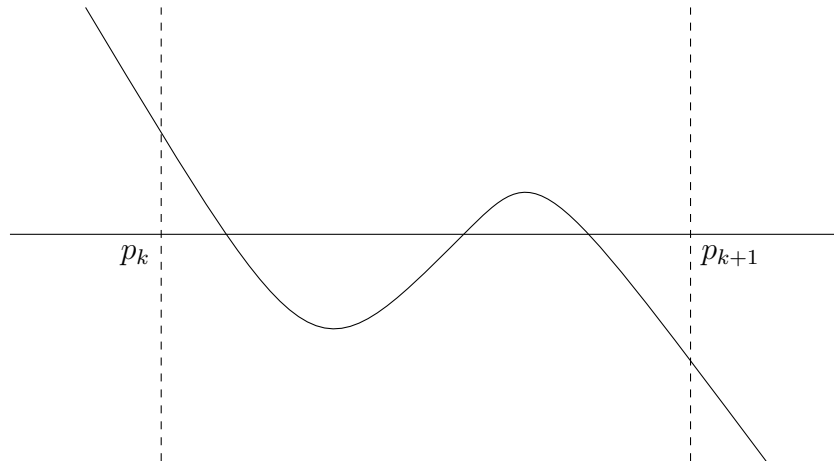
Similarly, if three or more zeros occur between two values of p , the rootfinding algorithm may find the wrong zero. Figure 3.4c illustrates that as long as a change of sign still occurs, the resulting error can only be an overestimate of at most ϵ .



(a) A change in sign may not be observed across an interval containing two zeros.



(b) Reducing the step size ϵ ensures that a change in sign is detected.



(c) If three zeros occur in the interval, the smallest zero may be overlooked.

Figure 3.4: Some configurations with more than one zero within an interval.

A particularly difficult case is C_5 , the cycle graph on five vertices, where we have

$$\det D_p = 2(1 + 2^p)(2^{2p} - 3 \times 2^p + 1)^2.$$

This quantity is non-negative throughout the neighbourhood of the only zero, as shown in Figure 3.5. Thus no change in sign occurs, causing our proposed algorithms to miss this zero and hence overestimate the maximum generalised roundness.

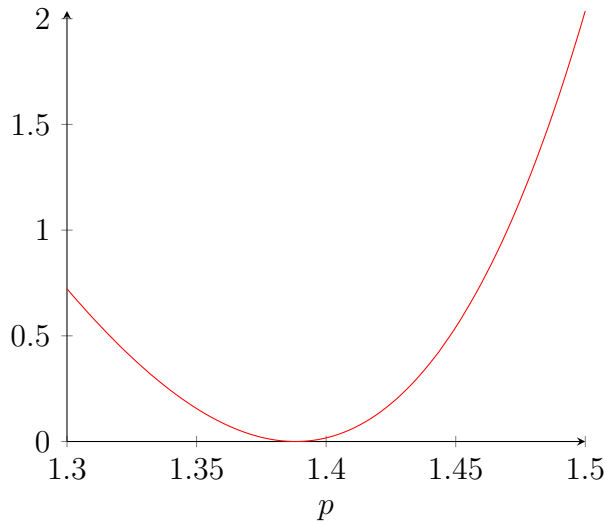


Figure 3.5: For $G = C_5$, $\det D_p$ has a zero without a change in sign.

The same behaviour occurs in C_7 and C_9 , and we suspect the following, although it is surprisingly difficult to prove due to the presence of various exponentials of different bases.

Conjecture 3.2.3. *For C_{2n+1} , the cycle graph on $2n + 1$ vertices, $\det D_p$ is of the form*

$$2(1 + 2^p + 3^p + \dots + n^p)(\dots)^2.$$

This is not the only example of this phenomenon; see for example the graph depicted in Figure 3.6, for which $\det D_p = -(2^{2p} + 4)(2^{2p} - 3 \times 2^p + 1)^2$.

This type of error is particularly troubling as the resulting difference between our calculation of the maximum generalised roundness and the true value is not bounded by ϵ , and may be as large as 2. Thus, we must make efforts to identify and correct such errors.

Recalling that $\det D_p$ is a smooth function, we can attempt to identify cases with two zeros close together or coincidental by analysing the stationary points. Since

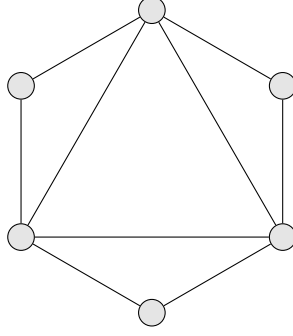


Figure 3.6: A graph on six vertices.

exact zeros of the derivative cannot be found, we calculate the finite differences $g_k = \det D_{p_k} - \det D_{p_{k-1}}$. Suppose no change in sign of $\det D_p$ is detected between p_{k-2} and p_k , that is, $\det D_p$ has the same sign for $p = p_{k-2}, p_{k-1}, p_k$. Assume without loss of generality that these determinants are all positive. We also assume that ϵ is small enough to ensure that at most one turning point exists in (p_{k-2}, p_k) . Now, a change in sign between g_{k-1} and g_k indicates the presence of such a turning point, and in particular if $g_{k-1} < 0$ and $g_k > 0$, there is a minimum turning point p^* . The determinant will be decreasing on (p_{k-2}, p^*) and increasing on (p^*, p_k) , so we can use ternary search to find p^* . If $\det D_{p^*} < 0$, as in Figure 3.7, we have a change in sign on (p_{k-2}, p^*) , and we can again find the smallest zero by bracketing. If $\det D_{p^*} > 0$, no zero has been found, so we resume the search from p_k onwards. Finally, if $\det D_{p^*} = 0$, we have a zero at the minimum turning point, as required.

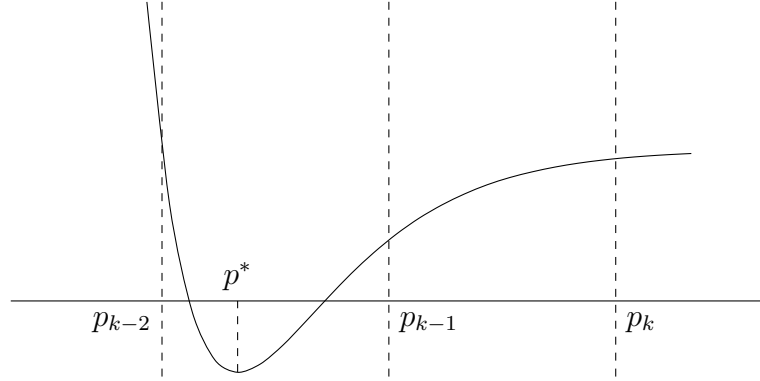


Figure 3.7: If one stationary point occurs in the interval, we may detect a zero without a change in sign.

Note that if two or more turning points appear in (p_{k-2}, p_k) , this procedure may fail to find the smallest zero. In this case, any combination of signs of g_{k-1} and g_k permits a zero at a turning point. Also, the p^* found by ternary search may only be a local minimum, so we may obtain a zero at p^* but miss an even smaller zero,

or alternatively we may fail to find a zero altogether on (p_{k-2}, p_k) if our minimum is not the smallest on this interval, as occurs in Figure 3.8.

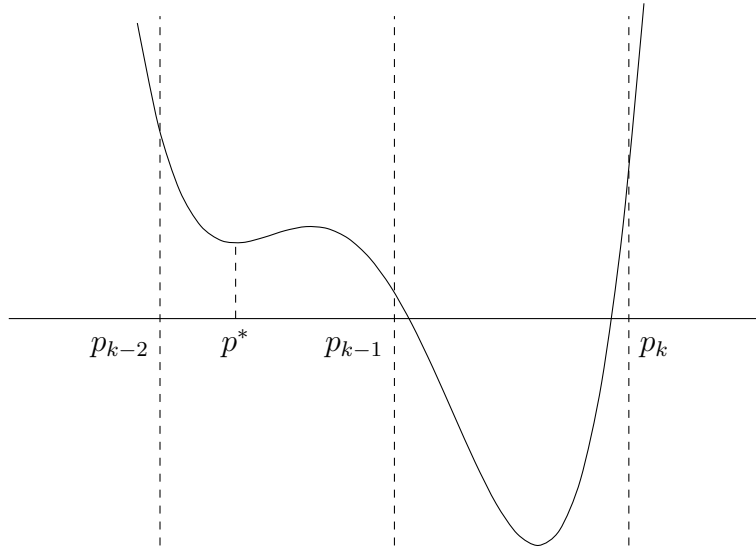


Figure 3.8: If two or more stationary points occur in the interval, a zero may be missed.

Fortunately, $\det D_p$ is not a particularly pathological function, as it is only a linear combination of exponentials. After inspection of many examples, we found occurrences of many zeros in close proximity only for large n , and only rarely did the determinant have a zero at a turning point. In a sample of graphs on 100 vertices, attempts to detect these zeros were able to improve our calculated maximum generalised roundness value in fewer than one in ten cases, and reduce it by more than ϵ much less often.

A further consideration here results from the arithmetic used in these computations. Exact arithmetic is infeasible due to the number of operations involved, so we rely on floating point representations of numbers, which introduces small errors in each calculation. It is difficult to quantify how large the final errors are in values of $\det D_p$, but values particularly close to zero are not necessarily reliable, as the true value may be of the opposite sign to the result produced by a computer. This is of particular concern if the true value is exactly zero, since we may fail to detect a change in sign in the preceding interval.

A notable example arises from the cycle graph on four vertices. Since $\wp(C_4) = 1$, a relatively small value for a graph of this size, the addition of one or more vertices and edges often creates a new graph with the same maximum generalised roundness.

As a result, 1 is a relatively common value especially among small graphs. Hence we avoid step sizes ϵ of the form $1/k$ where $k \in \mathbb{Z}^+$.

Also, in the case of zeros at turning points, the exact value will never be found so we must accept $\det D_{p^*} = 0$ up to some tolerance δ . However, if δ is too large, we may encounter false positives, where a turning point close to zero is mistaken for a zero although no zero exists on the interval in question. Examination of many examples indicates that such turning points do not occur often, and if a suitably small tolerance is chosen, then $|\det D_{p^*}| < \delta$ typically only occurs at a true zero.

Note finally that it is not immediately clear that we must analyse $\det D_p$ at all. The zeroes of $\det D_p$ correspond exactly to the discontinuities of $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$, so one may be tempted to only consider the second function and search for both zeros and discontinuities. However, this function can have removable singularities, which cannot be detected by sampling values. Again, C_5 is an example of this phenomenon, where

$$\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = \frac{5}{2(2^p + 1)}$$

with a removable singularity at $p = \log_2 \frac{3 + \sqrt{5}}{2}$, the only zero of $\det D_p$.

3.2.2 Sum of entries of the inverse of the p -distance matrix

All of the problems raised in relation to $\det D_p$ also apply to $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$, but there are some additional concerns to be addressed here. For one, the form of this quantity is likely to be significantly more complicated than the form of $\det D_p$, making it even more expensive to compute and solve symbolically. Finding the inverse of the p -distance matrix is itself an expensive operation, but this can be mitigated by instead solving the linear system $D_p \mathbf{u} = \mathbb{1}$ and then evaluating $\langle \mathbf{u}, \mathbb{1} \rangle$.

Recall that our rootfinding algorithm relies on the smoothness of the function in question. Clearly, there may be some values of p for which D_p is singular and hence $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ is undefined, but Cramer's rule guarantees that it is otherwise smooth. As we are only interested in the smaller zero of either expression, we can use the algorithm to find a zero p^* of $\det D_p$, then look for zeros of $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ on $(0, p^*)$ by the same procedure. If p^* is indeed the smallest zero of $\det D_p$, then $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ is continuous on $(0, p^*)$ as desired. However, if a smaller zero of $\det D_p$ exists, we may be searching an interval on which the function is not continuous, making the result unreliable. Again, this is rare provided that a small step size ϵ is used. Note that it is also still possible for more than one zero to appear in close proximity.

3.3 Algorithm

We now present an algorithm to numerically determine the smallest non-negative zero of either $\det D_p$ or $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$, relying on the properties of these functions as described in the previous section. The following algorithm **FINDSMALLESTZERO** iterates through values of the argument p with stepsize ϵ and at each stage looks for a change in sign or an appropriate turning point to indicate the presence of a zero. The first interval detected to contain a zero is then searched more closely by **FINDZERO**.

Algorithm 1 A procedure to approximate the smallest non-negative zero of f

Parameters: ϵ, δ

procedure FINDSMALLESTZERO(f)

$k = 0$

while $p_k = k\epsilon < 2$ **do**

$f_k = f(p_k)$

if $k \geq 1$ **then**

$g_k = f_k - f_{k-1}$

if $f_k = 0$ **then**

return p_k

else if $k \geq 1$ and $f_{k-1}f_k < 0$ **then**

▷ change in sign detected

return FINDZERO(f, p_{k-1}, p_k)

else if $k \geq 2$ and $g_{k-1}g_k < 0$ **then**

▷ turning point detected

if $g_k f_k > 0$ **then**

▷ orientation allows a zero

if $f_k > 0$ **then**

$h = f$

else

$h = -f$

$p^* = \text{FINDMIN}(h, p_{k-2}, p_k)$

if $h(p^*) < -\delta$ **then**

return FINDZERO(f, p_{k-2}, p^*)

▷ a zero occurs before the turning point

else if $h(p^*) < \delta$ **then**

return p^*

▷ a zero occurs at the turning point

$k = k + 1$

return ∞

FINDZERO numerically finds a zero of f on a given interval on which f changes sign. A simple yet robust implementation can be achieved using the bisection method, with a parameter b for the number of iterations.

Algorithm 2 A procedure to approximate a zero of f on the interval (lo, hi) , given that $f(lo)f(hi) < 0$

Parameters: b

```

procedure FINDZERO( $f, lo, hi$ )
  for  $k = 1..b$  do
     $mid = (lo + hi)/2$ 
    if  $f(lo)f(mid) < 0$  then
       $hi = mid$ 
    else
       $lo = mid$ 
  return  $(lo + hi)/2$ 

```

Alternatives such as the secant method or false position may be preferred for speed, but convergence is not guaranteed. However, most numerical computing packages include highly optimised and reliable procedures for this task, such as Maple's `fsolve` and MATLAB's `fzero`, and in practice these are the best choices. `fzero` implements the Zeroin algorithm [60] developed by Brent [10], who built upon an earlier algorithm of Dekker [22].

Similarly, FINDMIN numerically finds a local minimum of f on a given interval, for example by trisection.

Algorithm 3 A procedure to approximate a local minimum point of f on the interval (lo, hi)

Parameters: b

```

procedure FINDMIN( $f, lo, hi$ )
  for  $k = 1..b$  do
     $mid_1 = (2lo + hi)/3$ 
     $mid_2 = (lo + 2hi)/3$ 
    if  $f(mid_1) \geq f(mid_2)$  then
       $lo = mid_1$ 
    else if  $f(mid_1) \leq f(mid_2)$  then
       $hi = mid_2$ 
  return  $(lo + hi)/2$ 

```

Again, methods such as parabolic interpolation and golden-mean search are applicable here, but we can instead invoke MATLAB's `fminbnd` which uses a combination of these two methods [60]. Alternatively, an approximation is efficiently achieved by simply iterating over the interval with a finer stepsize and selecting the smallest value found.

3.4 Implementation and Validation

Writing a short program in a high level mathematical package to find roots of $\det D_p$ and $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ is very easy. The challenge is to be able to quickly and accurately find the smallest roots. We initially coded trial procedures in Maple, and this quickly threw up many of the potential problems discussed above.

In order to properly test our procedures we generated a large library of test graphs, using Maple's `RandomGraphs` package. For

$$n \in \{7, 10, 15, 20, \dots, 100\}$$

we produced 100 random graphs on n vertices with edge probability p for each value of $p \in \{0, 1, 0.2, \dots, 0.9\}$. We also produced smaller test sets for $n = 200$ and $n = 500$. Additionally, we catalogued all graphs on up to eight vertices, all trees on up to seventeen vertices, all planar graphs on up to ten vertices, various regular graphs as well as larger collections of uniform random graphs and random trees to accompany the work in later chapters. All up, the test library (which is available at <https://tinyurl.com/desilva-mgr>) contains several hundred thousand graphs.

Notation 3.4.1. For $k = 0, \dots, 99$, let $G(n, p, k)$ be the k th graph on n vertices with edge probability p in the test library.

Maple is not a very efficient tool for doing large scale numerical work, so our aim was to do the bulk of our computations using MATLAB. Both these systems were chosen because they include many inbuilt mathematical procedures and structures, and allow arbitrary precision arithmetic.

Despite Maple's computational limitations, it was very valuable to have competing programs on different platforms. To minimize possible coding errors, my supervisor and I independently coded procedures in Maple and MATLAB, and then compared our results for the test library. These programs are available in Appendices [A](#) and [B](#) respectively.

As discussed above in Figure [3.3](#), it was prohibitively slow to explicitly calculate the two functions, so our procedures needed to depend purely on numerical sampling of the functions.

3.4.1 Parameters

Any implementation of the algorithm involves a choice of a number of parameters which control the balance between accuracy and speed. Given the intended use in the chapters to follow, extreme precision was perhaps less important than speed. Our aim was to give an answer with an absolute error up to 10^{-3} , 99% of the time, for $n \leq 100$.

The parameters include:

Precision d : Both MATLAB and Maple allow the user to set the number of digits that they use when doing calculations with floating point numbers. Since we were often dealing with nearly singular matrices, insufficient precision could have caused inaccuracies in our computations.

Stepsize ϵ : We could achieve great accuracy by using very small values of ϵ , but this resulted in very long run-times. After considerable experimentation, we used $\epsilon \approx \frac{2}{5n}$ in the Maple procedure and $\epsilon = 0.0003$ in MATLAB.

$\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ upper bound p_m : It is only useful to search for a zero of $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ in the interval up to the smallest zero p^* of $\det D_p$ found. Of course $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ is not defined at p^* , so one can only sample up to some $p_m < p^*$. In some graphs, the smallest zero of $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ was very close to p^* . A simple example will be shown later in Example [3.6.1](#). If this smallest zero of $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ was greater than p_m , this would cause an overestimate of the maximum generalised roundness of up to $p^* - p_m$. In Maple, we choose $p_m = p^* - 0.0001$, ensuring that $p^* - p_m < 10^{-3}$, so any error that results from this choice is within the tolerable margin. The same outcome is achieved in MATLAB by our earlier choice of $\epsilon = 0.0003$.

Zero tolerance δ : Since we are only sampling the functions, it is hard to numerically determine whether a function has a zero at a turning point p^\dagger . This parameter was chosen so that if the function was within δ of zero at some approximation to p^\dagger then we would assume it was zero at p^\dagger . After inspection of the function values near the first zero in several test cases, we used $\delta = 10^{-n/10}$ in MATLAB. This seemed to have minimal impact, as a much higher value was used in Maple without causing conflicts. It should be noted however that for very large n , say $n \geq 170$, the value of $10^{-n/10}$ may be less than the machine epsilon and thus underflow to zero. We are therefore cautious about applying this algorithm for such large metric spaces.

Rootfinding iterations b : Once the procedures have found a change in sign, they approximate the zero using some iterative method. One obviously obtains greater accuracy with more iterations, but again at the cost of a longer run-time. If the bisection method is used, the absolute error from this process is at most $2^{-b}\epsilon$, which is certainly at most 10^{-3} for $b \geq 10$. Therefore, for the results later in the thesis we used at least ten iterations of the bisection method in Maple. On the other hand, our MATLAB program uses a smaller value of ϵ which ensures that the error is within acceptable margins even if no rootfinding is attempted, but `fzero` yields a more precise result.

Our initial choices of the parameters were informed by trial runs on the test library. In a number of cases, the two implementations gave different answers and these were examined more closely using the more hands-on method described earlier. We were thus able to correct systematic errors in both the MATLAB and Maple implementations until both programs agreed with the results of the manual process on many randomly chosen tests, and with each other on the entire dataset to an even higher degree of precision and higher frequency than we had initially desired.

3.4.2 Accuracy

There are many potential sources of error, some due to the rootfinding procedures, and some due to computational floating point error. Of these we believe that the first is by far the most important. However, it is worth commenting on the floating point errors first. In our rootfinding algorithm, it is crucial that we always correctly identify the sign of say, $\det D_p$. We will see later in this section that $|\det D_p|$ can be very small over an extended range, and so it is quite reasonable to be cautious as to whether the sign of the computed result is sufficiently reliable.

The calculation of a value of $\det D_p$ involves two main steps. The first is to calculate $d(x_i, x_j)^p$ numerically for each i and j , where we can safely assume that these values are accurate to d digits. The second step is to then calculate the determinant of a matrix, which for p close to a zero of $\det D_p$ is likely to be almost singular. Analyzing the accuracy of this second step is considerably more complicated than the first. We have throughout depended on the inbuilt and presumably state-of-the-art numerical algorithms² to calculate the determinants, although it seems rather difficult to obtain precise details of how much precision one should expect. It should be noted of course that these algorithms avoid the sort of cancellation errors that

²Maple's numerical linear algebra algorithms are based on the BLAS CLAPACK and NAG libraries [5].

might occur if one tried naively to fully expand out the determinant. Typically the determinant is calculated as a product of the diagonal elements from a suitable factorization of the matrix, so the main concern for us is really that the signs of the eigenvalues are correct.

Generally we worked with a fixed precision of $d = 15$ digits in Maple, but for testing purposes we worked with up to 200 digits in order to make sure that this was not changing what we were observing. The value of d sets the number of significant digits that each step in the computation is calculated to. For a large and complicated computation, the number of correct significant digits may be rather smaller than d .

We ran several tests to investigate the effect of changing the precision parameter. For a fixed randomly generated 100 vertex graph and a fixed value of p , we used Maple to calculate $\det D_p$ with the precision d increasing from 2 to 100. The value of $p = 0.155$ was chosen between the (apparent) first and second zeros of $\det D_p$. The results for small d are shown in Table [3.1](#).

| d | $\det D_p$ |
|-----|--------------------------|
| 2 | -1.6495×10^{-3} |
| 3 | -6.1254×10^{-6} |
| 4 | 1.2804×10^{-7} |
| 5 | 1.9240×10^{-7} |
| 6 | 1.9421×10^{-7} |
| 7 | 1.9439×10^{-7} |
| 8 | 1.9435×10^{-7} |
| 9 | 1.9436×10^{-7} |
| 10 | 1.9435×10^{-7} |
| 11 | 1.9435×10^{-7} |
| 12 | 1.9435×10^{-7} |
| 13 | 1.9435×10^{-7} |
| 14 | 1.9435×10^{-7} |
| 15 | 1.9435×10^{-7} |

Table 3.1: The first few digits of $\det D_p$ for a nearly singular 100×100 matrix.

What we generally observed was that beyond any precision d , the first $d - 3$ digits stabilized. So, for example, for all $d \geq 18$, the first 15 digits were always the

same. This gave us confidence in the values the computer was presenting to us, and also in the precision level $d = 15$ that we used for most of our experimental runs.

Similar results were observed in MATLAB, using the Symbolic Math Toolbox to calculate determinants with high precision. Again it was clear that the default of 52 bits of precision, as per IEEE Standard 754, was enough to reliably determine the sign of $\det D_p$. Further confirmation will be provided in Subsection 3.4.5, in which we compare our programs against a more stable method which is only applicable for graphs of diameter two.

The primary aim of course was not to accurately calculate determinants, but to find the zeros of $\det D_p$ and $\langle D_p^{-1} \mathbf{1}, \mathbf{1} \rangle$. For this we were depending much more on the efficacy of our rootfinding algorithms. The biggest risk here was in choosing a value of the stepsize ϵ which is too large and missing a zero (as in Figure 3.4a or Figure 3.8), or in setting the zero tolerance δ inappropriately so that either a zero at a turning point was missed, or else one was incorrectly identified.

3.4.3 Speed

The interplay between precision and speed is not always very predictable, partly due to the way the packages store floating point numbers of different precision.

In Maple, floating point numbers are stored as a pair of integers (S, E) representing $S \times 10^E$. At lower precision S can be stored as a hardware integer which leads to more efficient arithmetic operations. Higher precision obviously requires larger values of S and this can result in significantly slower computation as seen in Figure 3.9.

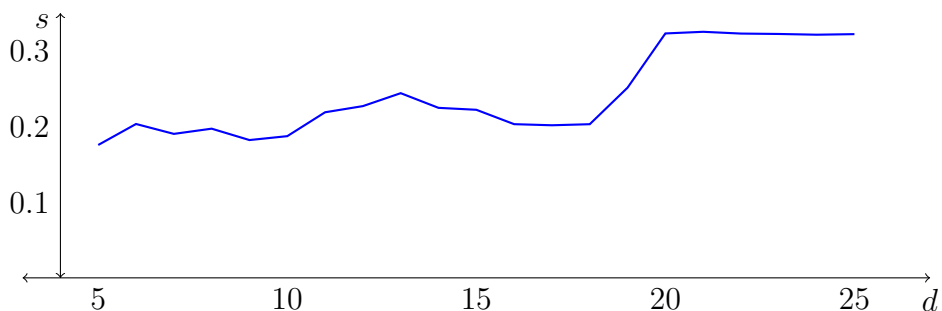


Figure 3.9: Average seconds for Maple to calculate the determinant of 70×70 matrices numerically at different values of the precision variable d (on a standard PC).

Less expected was how the size of the matrices affected the run time. As the size passed 64, there was (at least on our machines) an abrupt and significant slowdown,

which we suspect was due to some implementation detail in Maple’s linear algebra libraries.

In MATLAB, the time taken to compute $\det D_p$ was similar for all values of d up to 15, followed by a large jump and then very little change for $d \geq 16$. Again, this is likely explained by the internal details of the Symbolic Math Toolbox. The relationship between matrix size and time did not exhibit any unusual behaviour, simply growing faster than linear as one would expect.

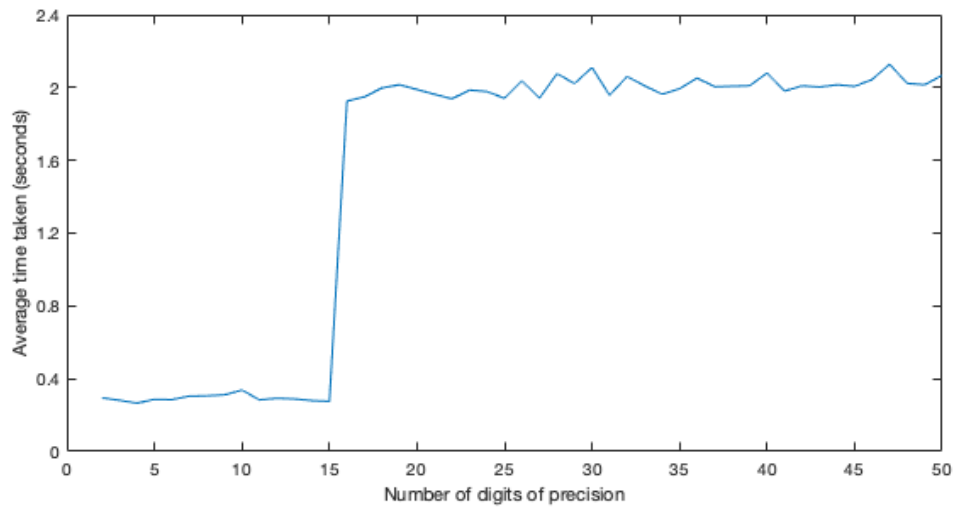


Figure 3.10: Average seconds for MATLAB to calculate the determinant of 70×70 matrices numerically at different values of the precision variable d (on a standard PC).

3.4.4 Larger graphs

For the majority of smaller graphs, the functions $\det D_p$ and $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ have easily identifiable first roots which are well separated from any larger roots, and it is easy to verify that our programs are correctly identifying the value of $p(G)$. Checking the programs for graphs of say 100 vertices is more difficult. In order to test our procedures at this scale, we numerically calculated $\det D_p$ for a large number of values of p to ascertain the general features of this function. It quickly became apparent that for larger graphs this function typically behaves quite differently to the case for smaller graphs.

This is best illustrated by looking at a particular graph, $G(100, 0.5, 0)$, from our test library, although the same issues appear to be common among graphs on 100 vertices. Plotting 250 values of $\det D_p$ (using 15 digit precision) from 0 to 0.367

produced the graph in Figure 3.11, showing a large interval on which the functions values appear to be small.

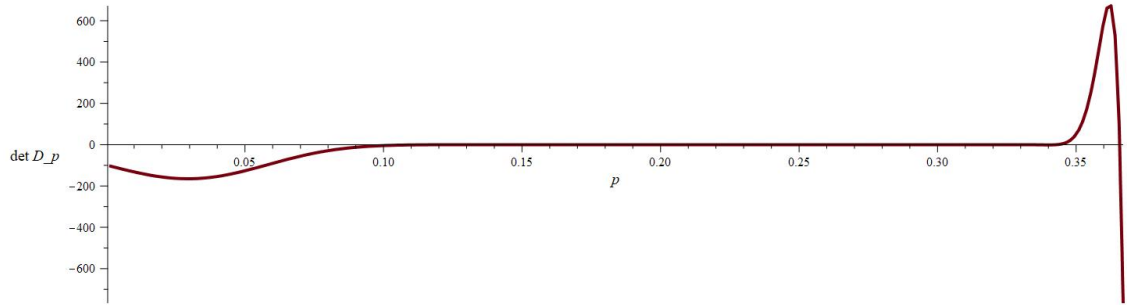


Figure 3.11: $\det D_p$ for the graph $G(100, 0.5, 0)$.

Zooming in to find the smallest zero of the function, we examined the values in the interval $[0.154, 0.17]$ (Figure 3.12). By the graph, this indicates a value of $p(G)$ of around 0.1549 which indeed matches the values found by both the Maple and MATLAB procedure (0.15488). It also shows however that the values of $\det D_p$ are extremely small in this interval, and so it is not unreasonable to worry about whether the calculation of the determinants is actually being performed accurately enough that one can place faith in the values being shown. In particular, the matrices D_p are extremely ill conditioned in this interval.

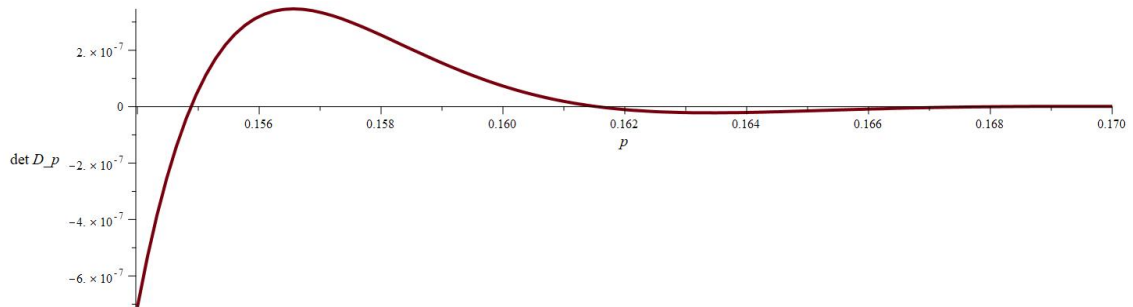


Figure 3.12: $\det D_p$ for the graph $G(100, 0.5, 0)$ near the first zero.

3.4.5 Verification

It does not appear to be easy to directly determine the accuracy of Maple or MATLAB's numerical determinant procedures. This depends rather crucially on the algorithms used, and this is not entirely transparent. Fortunately, there is another way to do these calculations which is easier to analyze.

As will be discussed in Chapter 5, a random graph on 100 vertices will have diameter 2 with high probability. This means that the entries in D_p can only take

the values 0, 1 and 2^p . If we perform the substitution $\lambda = 2^p$, then $\det D_p$ is a polynomial $q(\lambda)$, which in this case has degree at most 100. Importantly, this simpler structure does allow Maple to symbolically calculate $q(\lambda)$. For the example considered above,

$$\begin{aligned}
q(\lambda) &= \sum_{k=0}^{\infty} c_k \lambda^k \\
&= 1856810621108920971825327449735739491432122314981 \lambda^{100} \\
&\quad - 249576817146606799242693308666651592837695218064294 \lambda^{99} \\
&\quad \vdots \\
&\quad + 1243047100419479950644328663789530075664659356675536 \lambda \\
&\quad - 14190785215850538100794464003125835446457018702068.
\end{aligned}$$

We are interested in the value of $q(\lambda)$ for λ near $2^{0.156} \approx 1.114$. As q is a polynomial, it is somewhat easier to analyze the numerical accuracy.

Each of the individual terms $c_k \lambda^k$ is clearly very large. The largest of these at $\lambda = 1.114$ is $c_{54} 1.114^{54} \approx 2.67 \times 10^{80}$. On the other hand, we are hoping that $q(1.114) \approx 1.08 \times 10^{-7}$, so it is vital that each term is calculated with great accuracy.

Roughly speaking, one expects to lose up to 2 significant figures in computing each term $c_k \lambda^k$. The subtractions involved in computing $q(1.114)$ however result in a very large loss of significant figures. In order to have even a small number of accurate digits in the final value we must therefore work to a precision of nearly 100 digits.

Setting the precision d to 100 (and again for safety to 200) digits, we then evaluated $\det D_p = q(2^p)$ near the assumed first zero and also plotted the result. Maple calculates the first zero to be at $p = 0.15487944 \dots$. The graph, shown in Figure [3.13](#), is indistinguishable from the one in Figure [3.12](#). This gives us excellent evidence that our programs are not suffering from any significant problems with computational roundoff error, even with larger graphs.

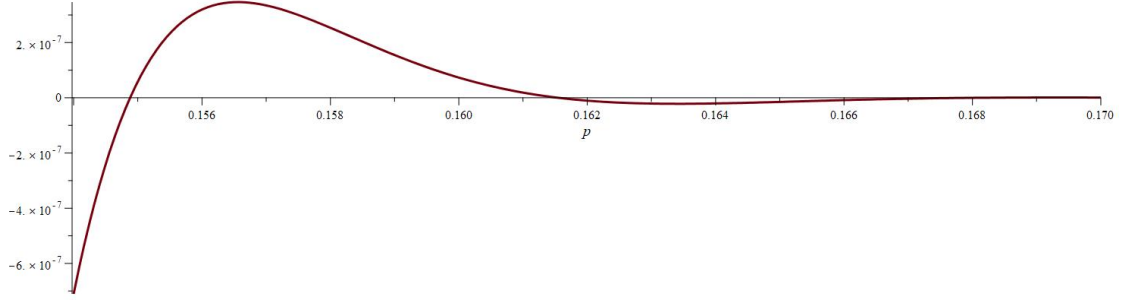


Figure 3.13: $q(2^p)$ for the graph $G(100, 0.5, 0)$ near the first zero.

We repeated this exercise with the first 200 vertex graph $G(200, 0.5, 0)$. For this graph the first turning point of $\det D_p$ has a value of about 10^{-16} . The symbolic evaluation of $q(\lambda)$ took thousands of cpu seconds, but at 300 digits of precision, identified the first zero as occurring at $\lambda = 1.0781634154\dots$ which corresponds to $p = 0.1085758610\dots$. This agrees with the value 0.10858 found by both the Maple and MATLAB procedures.

As we observe later in the thesis, for large random graphs, the maximum generalised roundness is almost always determined by the first zero of $\det D_p$, and so this is a viable way of calculating the maximum generalised roundness of most large graphs. On the other hand, as soon as the graphs have a distance greater than 2, for example in any nontrivial tree, one cannot reduce the problem to the analysis of a polynomial in one variable.

3.5 Calculation of the maximum generalised roundness of families of graph metric spaces

A wider question is to understand the distribution of the maximum generalised roundness in a family $\mathcal{F}(n)$ of n -point metric spaces. In the chapters to follow, we will examine in particular the trees, connected graphs and connected planar graphs on n labelled vertices, all with the path metric.

For small values of n , the various families of n vertex metric graphs may be small enough that one can evaluate $p(G)$ for every graph in the family. It is computationally preferable to iterate through unlabelled graphs in this way, rather than labelled graphs which may include several isomorphic copies of each graph. For small n , complete lists of all members of particular families can be generated or indeed found, using online resources such as the Combinatorial Object Server [63] or the House of Graphs [13]. The time taken to compute $p(G)$ for each $G \in \mathcal{F}(n)$

depends both on the size of the family and the size of the graphs, as well as the hardware used.

For $n = 8$, there are just over 11000 nonisomorphic³ unweighted connected graphs [64, Sequence A001349], and these could be checked in just over 20 minutes on a standard PC. The results of these calculations provided further evidence for the conjecture of Deza and Maehara [23] on c_n , the infimum over all n point metric spaces of the maximum generalised roundness. Recall that the values of c_5 , c_7 and c_8 are not currently known, although the minimising spaces are widely believed to be the complete bipartite graphs $K_{2,3}$, $K_{3,4}$ and $K_{4,4}$. Our calculations showed that there are no unweighted graphs with smaller maximum generalised roundness for these values of n .

For trees one can go somewhat further. There are almost 50000 nonisomorphic unweighted trees on seventeen vertices [64, Sequence A000055] and so it is still feasible to check each tree within two hours. We are thus able to computationally verify that for $n \leq 17$, no unweighted n -vertex tree has smaller maximum generalised roundness than the star graph on n vertices. An analogous claim was made for connected planar graphs, that the minimising graph is the complete bipartite graph $K_{2,n-2}$. This was surprisingly disproven, but only after testing all 1052805 candidates [64, Sequence A003094] for $n = 10$ over several hours.

For larger n however, the growth in size of the various families makes it infeasible to test every member of the family. Indeed for some of these families, the size grows superexponentially, so the runtime can balloon from minutes or hours to days between one value of n and the next. For example, there are over ten million nonisomorphic connected graphs to check for $n = 10$ [64, Sequence A001349].

Rather than computing $p(G)$ for each $G \in \mathcal{F}(n)$, we may attempt to take a sample of fixed size from $\mathcal{F}(n)$ according to some random process, aiming to approximate the distribution of maximum generalised roundness values in $\mathcal{F}(n)$ by instead finding the distribution among our sample. This raises its own issues, as some procedures for generating members of $\mathcal{F}(n)$ are biased, producing some with a higher probability than others. This may result in a skewed distribution, unrepresentative of the wider family. It is crucial to find a process that generates these graphs uniformly at random, in order to calculate various statistics and make valid

³For contrast, there are over 250 million connected graphs on eight labelled vertices [64, Sequence A001187]!

inferences about the underlying distribution of maximum generalised roundness values in $\mathcal{F}(n)$.

In the chapters to follow, we will examine such processes for various families of graphs, and form hypotheses from the results. Recall that the procedure used to calculate the maximum generalised roundness values may be inaccurate due to zeros coinciding or occurring close together. However, upon inspecting many examples we find that this does not appear to happen often, so the effect on averages over a large sample is typically negligible. Often, we will find that rigorous proof is available only for relatively weak statements, while the empirical evidence hints that stronger results hold.

3.6 Exact calculation of the maximum generalised roundness

Despite the difficulties that arise in calculating the exact maximum generalised roundness of graph metric spaces using Corollary 3.1.5, it is particularly useful in some circumstances. If very few points are involved, it may be feasible to calculate $\det D_p$ symbolically, but analytic solutions are still difficult to obtain. However, as discussed in Subsection 3.4.5, an exception arises if all distances are either 1 or 2 and therefore all products of distances are powers of two. It follows that $\det D_p$ is polynomial in 2^p , so we may be able to find its exact zeros after a substitution. Even if the resulting polynomial equation cannot be easily solved, numerical approaches such as the Durand-Kerner method [29, 48] can be used to find all its roots. Similarly, $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ may also be expressed entirely in terms of 2^p in such cases.

Example 3.6.1. Consider the graph on five vertices shown in Figure 3.14.

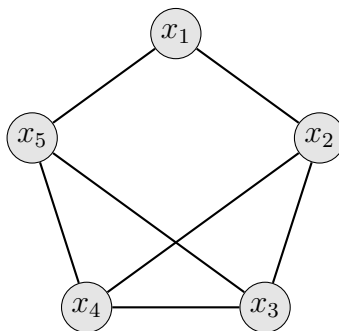


Figure 3.14: A graph on five vertices.

The p -distance matrix is

$$D_p = \begin{pmatrix} 0 & 1 & 2^p & 2^p & 1 \\ 1 & 0 & 1 & 1 & 2^p \\ 2^p & 1 & 0 & 1 & 1 \\ 2^p & 1 & 1 & 0 & 1 \\ 1 & 2^p & 1 & 1 & 0 \end{pmatrix}.$$

It follows that

$$\det D_p = -2(2^{3p} - 4 \times 2^p + 1)2^p,$$

and

$$\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = \frac{1}{2} \left(\frac{8 \times 2^{2p} - 17 \times 2^p + 4}{2^{3p} - 4 \times 2^p + 1} \right).$$

Letting $\lambda = 2^p$, it remains to solve $\lambda^3 - 4\lambda + 1 = 0$ and $8\lambda^2 - 17\lambda + 4 = 0$ on $[1, \infty)$. The only such roots are $\lambda \approx 1.8608$ and $\lambda = (17 + \sqrt{161})/16 \approx 1.8555$ respectively, so upon substituting back for p , we see that the maximum generalised roundness is

$$p = \log_2 \frac{17 + \sqrt{161}}{16} \approx 0.8918.$$

□

Unfortunately, graph metric spaces with only these distances are prevalent only in some situations. For example, when we consider path metric trees on n vertices, distances other than 1 and 2 occur in almost every tree, so both quantities of interest typically contain terms in 2^p , 3^p and so on, making the substitution $\lambda = 2^p$ useless.

To apply the formula for larger graphs, we typically require the graph to have some particular structure in order to efficiently compute $\det D_p$ and $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ and find their exact zeros. For example, in a complete bipartite graph $K_{m,n}$, the formula allows us to calculate the exact value of maximum generalised roundness p . This result follows from the work of Deza and Maehara [23], and the special cases for $m = 1, n \geq 2$ and $|m - n| \in \{0, 1\}$ were proven in [28] and Weston [78] respectively, but Sánchez's later proof in [71] is simpler, so it is replicated here.

Proposition 3.6.2 ([71, Theorem 3.1]). *The complete bipartite graph $K_{m,n}$ has maximum generalised roundness*

$$p(K_{m,n}) = \frac{2mn}{2mn - m - n}.$$

Proof. If $\det D_p = 0$, $D_p \alpha = 0$ has a nontrivial solution $\alpha \in \Pi_0$, as in the proof of Theorem [3.1.3](#). Computing the first m components of $D_p \alpha$, we have for each $j \in \{1, \dots, m\}$ that

$$2^p \sum_{\substack{i=1 \\ i \neq j}}^m \alpha_i + \sum_{i=m+1}^{m+n} \alpha_i = 0,$$

so after adding $2^p \alpha_j$ to both sides we have

$$2^p \sum_{i=1}^m \alpha_i + \sum_{i=m+1}^{m+n} \alpha_i = 2^p \alpha_j.$$

Thus $\alpha_1 = \alpha_2 = \dots = \alpha_m$. Similarly the last n components give

$$\sum_{i=1}^m \alpha_i + 2^p \sum_{i=m+1}^{m+n} \alpha_i = 2^p \alpha_j$$

for each $j \in \{m+1, \dots, m+n\}$, so $\alpha_{m+1} = \alpha_{m+2} = \dots = \alpha_{m+n}$. Recalling that $\alpha \in \Pi_0 - \{0\}$, we can now let

$$\alpha_j = \begin{cases} \frac{1}{m} & \text{if } j \in \{1, \dots, m\} \\ -\frac{1}{n} & \text{if } j \in \{m+1, \dots, m+n\} \end{cases}$$

without loss of generality. Now, we have

$$2^p(1) + (-1) = 2^p \left(\frac{1}{m} \right), \quad \text{so } p = \log_2 \frac{m}{m-1},$$

but simultaneously

$$(1) + 2^p(-1) = 2^p \left(-\frac{1}{n} \right), \quad \text{so } p = \log_2 \frac{n}{n-1}.$$

These two values are compatible if and only if $m = n$, so the determinant condition is of relevance in this case only.

On the other hand, if D_p is invertible, we can explicitly compute the inverse by exploiting its block structure. Writing

$$D_p = \begin{pmatrix} 0 & 2^p & \cdots & 2^p & 1 & 1 & \cdots & 1 \\ 2^p & 0 & \cdots & 2^p & 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 2^p & 2^p & \cdots & 0 & 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 & 0 & 2^p & \cdots & 2^p \\ 1 & 1 & \cdots & 1 & 2^p & 0 & \cdots & 2^p \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 & 2^p & 2^p & \cdots & 0 \end{pmatrix},$$

we may then use the identity

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}.$$

A lengthy process follows to calculate the required inverses using Gaussian elimination and eventually obtain

$$D_p^{-1} = \frac{1}{2^p((m-1)(n-1)2^{2p} - mn)} \begin{pmatrix} a & b & \cdots & b & c & c & \cdots & c \\ b & a & \cdots & b & c & c & \cdots & c \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b & b & \cdots & a & c & c & \cdots & c \\ c & c & \cdots & c & d & e & \cdots & e \\ c & c & \cdots & c & e & d & \cdots & e \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ c & c & \cdots & c & e & e & \cdots & d \end{pmatrix},$$

where

$$\begin{aligned} a &= -(m-2)(n-1)2^{2p} + (m-1)n, \\ b &= (n-1)2^{2p} - n, \\ c &= -2^p, \\ d &= -(m-1)(n-2)2^{2p} + m(n-1), \text{ and} \\ e &= (m-1)2^{2p} - m. \end{aligned}$$

Finally, we have

$$\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = \frac{(2mn - m - n)2^p - 2mn}{(m - 1)(n - 1)2^{2p} - mn},$$

which is zero at

$$p = \log_2 \frac{2mn}{2mn - m - n}.$$

If $m \neq n$, this is the maximum generalised roundness, as required. On the other hand, if $m = n$, this coincides with the value

$$\log_2 \frac{m}{m - 1} = \log_2 \frac{n}{n - 1}$$

obtained from the determinant condition, so in fact the determinant is zero here and the inverse does not exist. Even so, the same value

$$\log_2 \frac{2mn}{2mn - m - n}$$

is again the maximum generalised roundness, this time using the determinant condition. □

In later chapters, we aim to capitalise upon these exact maximum generalised roundness results to deduce probabilistic results about certain families of graphs. In general, we will achieve this by analysing large graphs in terms of their metric subspaces of known maximum generalised roundness, and invoking Theorem 2.1.17.

CHAPTER 4

Maximum generalised roundness of trees

Among graph metric spaces, a particularly important class is the family of metric trees, which embed naturally into ℓ_1 space. Further embedding problems arise from the fields of evolutionary biology [1, 35, 75, 77] and in computer science [4, 20], where one wishes to approximate finite metrics by a small number of tree metrics in order to improve algorithms for various optimisation problems. In this chapter, we will explore the maximum generalised roundness of path metric trees. While there is some previous understanding of the extreme values, we aim to describe the properties of typical trees, drawing on concepts from enumerative graph theory. It is relatively easy to prove the asymptotic behaviour of the maximum generalised roundness of random trees. The question of probabilistic bounds is far more difficult, and there is significant distance between the bounds we can prove rigorously and those we can determine experimentally.

4.1 p -negative type properties of trees

Recall from Chapter 2 that we defined a tree as a connected acyclic graph, and proved that there is a unique path between any pair of vertices. Given a tree, we can construct a corresponding metric space using the path metric, or the path weight metric arising from any choice of edge weights.

Definition 4.1.1. A metric tree is a weighted tree, where $|e|$ denotes the weight of edge e , endowed with the path metric, as in Proposition 2.2.15.

Definition 4.1.2. A path metric tree is a tree endowed with the path metric, as in Proposition 2.2.12.

The following theorem of Hjorth et al. established a strict lower bound for the maximum generalised roundness of a metric tree.

Theorem 4.1.3 ([47, Corollary 7.2]). *A finite metric tree (T, d) has strict 1-negative type, and hence $\wp(T, d) > 1$.*

Doust and Weston gave another proof, and obtain an explicit formula for the 1-negative type gap. This value surprisingly depends only on the edge weights, and not the internal structure of the tree.

Theorem 4.1.4 ([28, Corollary 4.13]). *The 1-negative type gap of a finite metric tree (T, d) is given by the formula*

$$\Gamma_T = \left(\sum_{e \in E(T)} |e|^{-1} \right)^{-1}.$$

Corollary 4.1.5. *The 1-negative type gap of a finite path metric tree T on n vertices is*

$$\Gamma_T = \frac{1}{n-1}.$$

Wolf [82, Corollary 4.3] provided an alternative proof of Theorem 4.1.4, using the formula for the p -negative type gap from Theorem 2.1.40.

Doust and Weston also proved a lower bound on the maximum generalised roundness of a finite tree on n vertices.

Theorem 4.1.6 ([28, Corollary 5.5]). *A finite metric tree T on $n \geq 3$ vertices has maximum generalised roundness*

$$\wp(T, d) \geq 1 + \frac{\log \left(1 + \frac{1}{(n-1)^3(n-2)} \right)}{\log(n-1)}.$$

This was later improved by Li and Weston by applying the result of Theorem 4.1.4 to Theorem 2.1.34.

Theorem 4.1.7 ([52, Corollary 3.4]). *A finite metric tree T on $n \geq 3$ vertices with diameter \mathcal{D} has maximum generalised roundness*

$$\wp(T, d) \geq 1 + \frac{\log \left(1 + \frac{1}{(n-1) \cdot \mathcal{D} \cdot \gamma(n)} \right)}{\log \mathcal{D}},$$

where

$$\gamma(n) = 1 - \frac{1}{2} \left(\frac{1}{\lfloor n/2 \rfloor} + \frac{1}{\lceil n/2 \rceil} \right).$$

Remark 4.1.8. Infinite trees also have strict 1-negative type, as the p -negative type inequality only ever applies to a finite subtree. Similarly, the formula in Theorem 4.1.4 also applies to infinite trees, where the gap is zero if the series diverges. However, an infinite tree may fail to have p -negative type for any $p > 1$. Doust and Weston [28, Theorem 5.7] constructed an example by linking an infinite sequence of star graphs to form an ‘infinite necklace’.

Note that the best possible upper bound for the maximum generalised roundness of a finite path metric tree on $n \geq 3$ vertices is simply $\wp(T) \leq 2$, achieved by the path graphs from Example 2.2.26. However, any tree with a vertex of degree at least 3 does not have maximum generalised roundness close to 2.

4.2 The trees on n vertices

The previous section details many results regarding the extremal maximum generalised roundness values among the trees on n vertices, but relatively little is known about the ‘typical’ tree. We begin by considering the family of trees on n vertices, and methods to randomly generate these trees.

For small values of n , it is possible to list all the possible trees on n vertices and evaluate the maximum generalised roundness for each one. One can then precisely identify the maximum and minimum values. However, for further questions concerning the distribution, one must make a distinction between labelled and unlabelled trees.

For a labelled tree on n vertices, one fixes vertices labelled $1, 2, \dots, n$ and then considers all the ways of attaching edges to these vertices to form a tree. Cayley’s formula [19] states that there are exactly n^{n-2} labelled trees on n vertices. Note that many of the trees formed in this way will end up being isomorphic. As a simple example, there are $\frac{4!}{2} = 12$ different labellings of the path graph on four vertices.

In contrast, unlabelled trees do not have labelled vertices, so each nonisomorphic tree is only counted once. The number t_n of unlabelled trees on n vertices does not have a known closed formula, with the asymptotic estimate

$$t_n \sim C\alpha^n n^{-\frac{5}{2}}$$

where $C \approx 0.5349$ and $\alpha = 2.9558$ due to Otter [65].

The numbers of labelled and unlabelled trees for small n are given in Table 4.1.

| n | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|-----------|---|---|----|-----|------|-------|--------|---------|-----------|
| n^{n-2} | 1 | 3 | 16 | 125 | 1296 | 16807 | 262144 | 4782969 | 100000000 |
| t_n | 1 | 1 | 2 | 3 | 6 | 11 | 23 | 47 | 106 |

Table 4.1: The number of labelled and unlabelled trees on n vertices.

If one is searching for say the n vertex tree T with the smallest value of $\wp(T)$, then it is clearly vastly more efficient to only test one tree of each isomorphism class. Enumerating all the unlabelled trees is a difficult task, but fortunately an algorithm of Wright et al. [83] provides an efficient method of generation, and for small values of n , complete lists of these trees are widely available such as at the House of Graphs [13]. Appendix C catalogues the maximum generalised roundness of each tree on $4 \leq n \leq 9$ unlabelled vertices. In the next section we shall discuss the significantly easier question of enumerating all the labelled trees, using Prüfer sequences.

If one wishes to write precise statements of the form “at least $x\%$ of the trees on n vertices have $\wp(T) < p$ ”, then evidently one must be clear whether such a statement is describing the labelled or unlabelled situation.

Past a certain point, one must obviously give up on listing all the possible trees of a given size. Instead one can obtain distributional data by finding $\wp(T)$ for a large sample of random trees. Here the situation becomes even more complicated. Random trees can be generated by a variety of models, which may generate trees with varying probability distributions. Some algorithms produce each labelled tree with equal likelihood, while others are biased. The same is possible in principle for unlabelled trees, although we are not aware of any sufficiently fast algorithm to produce large unlabelled trees uniformly at random. We therefore work only with labelled trees in the rest of this chapter. It is not surprising that our results for the distribution of $\wp(T)$ can depend critically on the random model chosen.

In the next section we shall look at the numerical data concerning the distribution of $\wp(T)$ for trees with relatively small numbers of vertices. In these cases we can enumerate all the possible trees. Following this, Section 4.4 examines the situation for larger trees, where we instead perform the same computations for a random sample of trees.

In Section 4.5 we shall show that the maximal generalised roundness of a random n point tree tends to 1 with high probability as $n \rightarrow \infty$. This will be achieved by

establishing that almost all trees on n vertices have maximum vertex degree at least some given value k , and thus these trees have maximum generalised roundness less than or equal to that of the star graph S_{k+1} , for which we have

$$\wp(S_{k+1}) = 1 + \log_2 \left(1 + \frac{1}{k-1} \right).$$

This work will depend partly on results of Moon [61] on the distribution of the maximum degree of a random tree.

More precisely, in Section 4.6 we shall prove an upper bound on the proportion ρ of trees on n vertices with the aforementioned degree property. This will finally allow us to bound the maximum generalised roundness of at least $(1-\rho)$ proportion of these trees, thus achieving meaningful descriptions of the typical tree without regard to outliers such as the path graphs.

The rigorous bounds we can prove are certainly not sharp. In Section 4.7 we shall compare these bounds with empirical data from an analysis of large samples of random graphs. In particular we shall present some results on the average maximum generalised roundness for samples of size 1000, for trees with between 10 and 100 vertices. Surprisingly, these averages agree almost exactly with

$$1 + \frac{1.2}{(n-2)^{0.46}}.$$

We also see that the standard deviations for these samples lie very close to¹

$$\frac{0.37}{(n-2)^{0.69}},$$

and so one can at least heuristically suggest expressions $L(n)$ and $U(n)$ so that, for example, 99.8% of trees on n vertices have maximum generalised roundness between $L(n)$ and $U(n)$.

4.3 The uniform model of random trees

One can attach a probability distribution to any family of graphs in order to construct a model of random graphs, and then describe the asymptotic properties of these random graphs.

¹The exponents in our approximations are in ratio 2 : 3. It is as yet unclear whether or not this relationship has deeper significance.

Definition 4.3.1. For $n = 1, 2, \dots$, let $\mathcal{G}(n)$ be a set of graphs with an associated probability measure μ_n . Then the collection

$$\mathcal{G} = (\mathcal{G}(n), \mu_n)_{n \geq 1}$$

is a model of random graphs.

Definition 4.3.2. Let P be a property of graphs. Then P holds almost surely (a.s.) in a model \mathcal{G} if

$$\mu_n(\{G \in \mathcal{G}(n) : G \text{ has } P\}) \rightarrow 1$$

as $n \rightarrow \infty$.

Using Cayley's formula for the number of labelled trees on n vertices, we can apply the uniform distribution to obtain a very simple model of random trees as follows.

Definition 4.3.3. Define the uniform model of random trees

$$\mathcal{T} = (\mathcal{T}(n), \mu_n)_{n \geq 1},$$

where $\mathcal{T}(n)$ is the set of all n^{n-2} labelled trees on n vertices each with equal probability measure $\mu_n(T) = n^{-(n-2)}$.

Definition 4.3.4. Let $T(n)$ be the discrete graph-valued random variable representing sampling from the uniform model, that is, uniformly at random from the set $\mathcal{T}(n)$.

We then apply the maximum generalised roundness function to obtain the real-valued discrete random variable $\wp(T(n))$, and aim to find its distribution features. For small n , we can simply calculate the maximum generalised roundness of each tree on n vertices. To efficiently generate the trees, we make use of a bijection [68] between the n^{n-2} trees on n vertices labelled $\{1, 2, \dots, n\}$ and the sequences $\{1, \dots, n\}^{n-2}$, where each tree is identified with its Prüfer sequence.

Algorithm to convert a tree into a sequence:

- $n - 2$ times: Remove the smallest labelled leaf and record its neighbour's vertex label as the next entry of the sequence.

Algorithm to convert a sequence into a tree:

- Calculate the required degree of each vertex, which is one more than the number of times it occurs in the sequence.
- $n - 2$ times:
 - Let i be the next term of the sequence.
 - Find the smallest labelled vertex j with required degree 1, add the edge ij to the tree and subtract one from the required degrees of i and j .
- After all $n - 2$ terms of the sequence are handled, only two vertices will remain with required degree one, with all others having required degree zero. Add an edge between these two vertices.
- The resulting graph has n vertices, whose degrees are all positive and sum to $2n - 2$. Such a graph must be a tree [61, p. 6].

As this process is invertible, it is a bijection.

By iterating through all possible Prüfer sequences, we can construct all the labelled trees on n vertices and explicitly find the distribution of $\wp(T(n))$. The example of $\wp(T(8))$ is pictured in Figure 4.1.

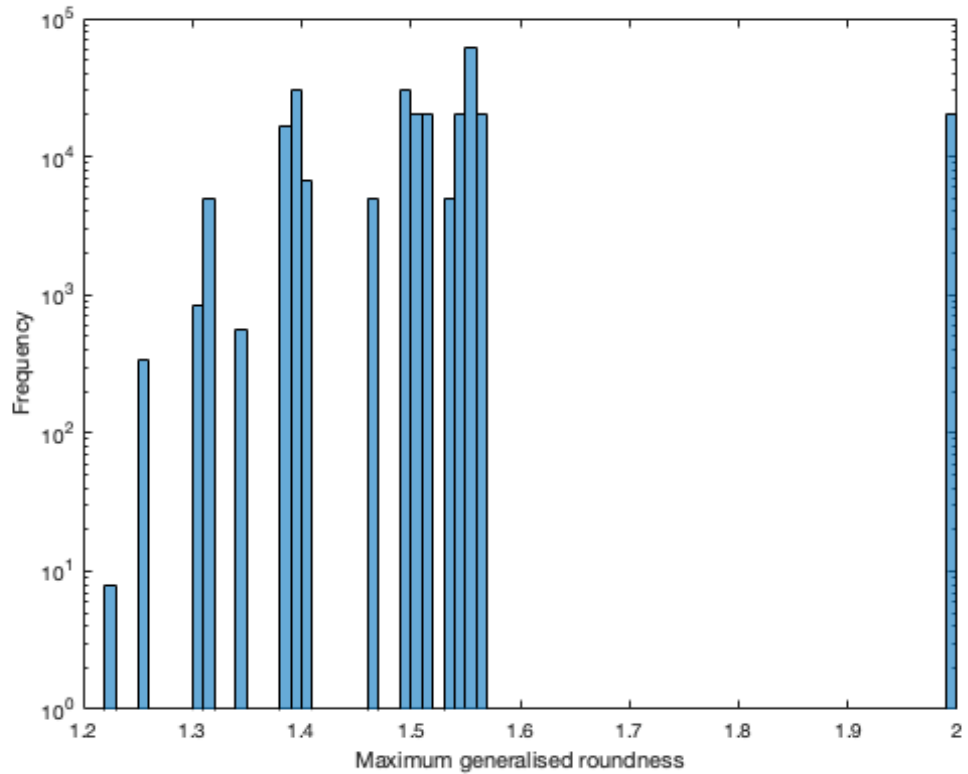


Figure 4.1: Log-scaled histogram of maximum generalised roundness values of all 262144 trees on 8 labelled vertices

4.4 Large random trees

For larger n , the number of trees prevents fast enumeration of the entire space, making it necessary to take a random sample of Prüfer sequences and hence trees. While this is a viable approach, we also present another perspective on the problem, viewing it as an attempt to generate random spanning trees of the complete graph K_n . The best known methods for producing minimum spanning trees of a weighted graph are Prim's algorithm [67] and Kruskal's algorithm [50]. Since all edges, and hence all spanning trees, are of equal weight here, we can formulate a randomised version of each algorithm.

Randomised Prim's algorithm:

- Form a tree T initially consisting of one vertex u of K_n chosen at random.
- $n - 1$ times: Randomly select two vertices $v \in V(T)$ and $w \in V(K_n) \setminus V(T)$. Add the vertex w and the edge vw to T .

Randomised Kruskal's algorithm²:

- Form a graph G initially consisting of all n vertices of K_n and no edges.
- $n - 1$ times: Randomly select an edge e of K_n , avoiding those already present in G . If e would create a cycle when added to G , discard it and pick again, as many times as required. When we find an edge which does not form a cycle in G , then add it to G .
- The resulting graph has $n - 1$ edges and is acyclic, so it is a tree.

Unfortunately, both these processes are biased, failing to generate trees uniformly at random. As an example, for $n = 4$ there are only two non-isomorphic trees to consider, namely the path graph P_4 and the star graph S_4 shown in Figure 4.2.

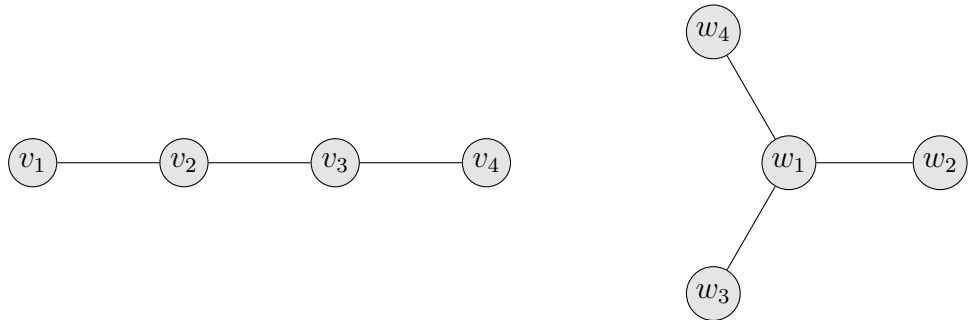


Figure 4.2: The two trees on four vertices, up to isomorphism, P_4 and S_4 .

²This is the algorithm used by the RandomTree command of Maple's GraphTheory package.

There are twelve relabellings of P_4 and four of S_4 , so an algorithm which produces trees uniformly at random should output P_4 with probability $\frac{3}{4}$ and S_4 with probability $\frac{1}{4}$. However, simple calculations reveal that the probability of obtaining P_4 from Prim's algorithm is only $\frac{2}{3}$, and from Kruskal's algorithm it is $\frac{11}{15}$. As discussed in Section 3.5, sampling from a non-uniform process may bias our results. For instance, both algorithms proposed thus far would give an underestimate of the average maximum generalised roundness among trees on four vertices.

Another algorithm assigns random weights to each edge of the complete graph K_n from the uniform distribution on $[0, 1]$, and takes the minimum spanning tree of the resulting weighted graph. Again, this fails to produce trees uniformly at random [40].

Wilson [81] instead proposed the following algorithm which uses loop-erased random walks to generate spanning trees of a given graph, in this case K_n , and proved that the resulting distribution is uniform.

Wilson's algorithm:

- Maintain a tree T , initially consisting of a single vertex chosen at random.
- $n - 1$ times:
 - Select a vertex $v \notin T$, and perform a random walk in K_n from v until a vertex of T is reached.
 - Remove any cycles traversed during this walk so that a path remains, and add all its edges and vertices to the tree.

For each n , we can now efficiently sample many such trees, using either random Prüfer sequences or Wilson's algorithm. We can then calculate the maximum generalised roundness of each tree of our sample, in order to approximate the distribution of $\wp(T(n))$. Our results for $n = 20, 30, \dots, 90$ are depicted in Figure 4.3, and the basic distribution features are summarised in Table 4.2.

| n | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
|--------------|--------|--------|--------|--------|--------|--------|--------|--------|
| \wp_{\max} | 1.4900 | 1.4046 | 1.3250 | 1.3028 | 1.2642 | 1.2842 | 1.2363 | 1.2213 |
| \wp_{\min} | 1.1806 | 1.1451 | 1.1390 | 1.1288 | 1.1188 | 1.1223 | 1.1089 | 1.1117 |
| $\bar{\wp}$ | 1.3189 | 1.2603 | 1.2265 | 1.2044 | 1.1859 | 1.1727 | 1.1621 | 1.1538 |

Table 4.2: The maximum, minimum and average maximum generalised roundness of 1000 trees on n vertices selected uniformly at random.

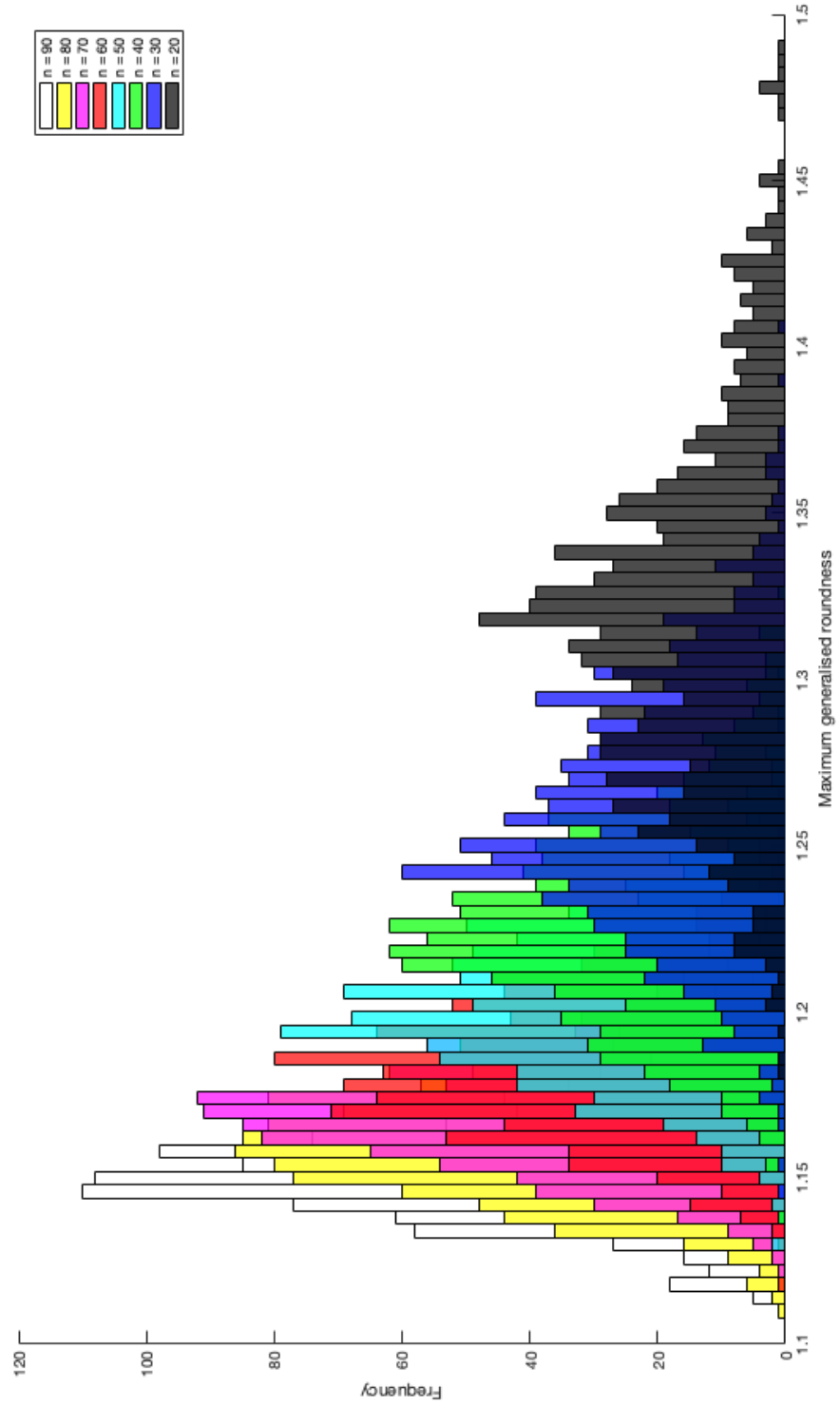


Figure 4.3: Histograms of maximum generalised roundness values of 1000 trees on n vertices selected uniformly at random, $20 \leq n \leq 90$

The data suggests that as n grows, the maximum generalised roundness is close to 1 with high probability, that is, for any $\epsilon, \delta > 0$, for n sufficiently large we have

$$\mu_n(\{T \in \mathcal{T}(n) \mid \wp(T) < 1 + \epsilon\}) > 1 - \delta.$$

4.5 Limiting behaviour

Our approach is to first apply Sánchez's formula from Corollary 3.1.5 to particular trees, then leverage it using our knowledge of metric subspaces to bound the maximum generalised roundness of large random trees.

The following lemma provides a simple characterisation of Theorem 2.1.17 in the context of trees.

Lemma 4.5.1. *Let T be a tree and T' be a subtree of T , as in Definition 2.2.20. Then $\wp(T) \leq \wp(T')$.*

Proof. Suppose v and w are vertices of T' . Proposition 2.2.18 states that there is a unique path between these vertices in T , and this path must also appear in T' as the subtree is connected. Thus $d_T(v, w) = d_{T'}(v, w)$, that is, T' is a metric subspace of T . The statement now follows by Theorem 2.1.17. \square

An immediate application of this shows that the path graph is indeed an outlier, and other trees on n vertices have maximum generalised roundness much less than 2.

Corollary 4.5.2. *Let T be a tree which is not a path graph. Then $\wp(T) \leq \log_2 3$.*

Proof. First note that the maximum degree of a tree is two if it is a path, and at least three otherwise. Thus S_4 , the star graph on four vertices shown in Figure 4.4, must appear as a subtree of T .

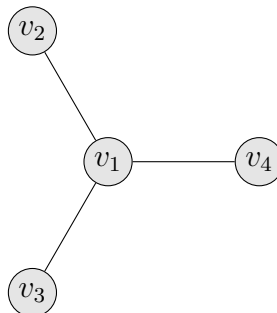


Figure 4.4: The star graph on four vertices, S_4 .

We can invoke the result of Example 2.2.27 here, but a simpler argument follows from Corollary 3.1.5

The p -distance matrix of T' is

$$D_p = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 2^p & 2^p \\ 1 & 2^p & 0 & 2^p \\ 1 & 2^p & 2^p & 0 \end{pmatrix}$$

so the determinant is

$$\det D_p = -3 \times 2^{2p}$$

and the sum of entries of the inverse is

$$\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = 2 \left(1 - \frac{2^p}{3} \right),$$

so $\wp(T') = \log_2 3$.

Finally, using Theorem 2.1.17, we can deduce that $\wp(T) \leq \log_2 3$ for any tree which is not a path graph. \square

More generally, the aim is to find a common subtree among large random trees. Let S_m be the star graph with m vertices and $m - 1$ leaves, with maximum generalised roundness

$$\wp(S_m) = 1 + \log_2 \left(1 + \frac{1}{m-2} \right) \quad (4.5.1)$$

from Example 2.2.27. For a given tree T on n vertices, S_m is a subtree if and only if the tree has a vertex of degree at least $m - 1$, that is, $\Delta(T) \geq m - 1$ where $\Delta(T)$ denotes the maximum degree over all vertices of T . Intuition suggests that as n increases, the maximum degree of typical trees on n vertices should also increase, and consequently the maximum generalised roundness will converge in probability to one.

To confirm this claim, we must have an understanding of the distribution of the maximum degree of random trees, which is governed by the following theorem of Moon.

Theorem 4.5.3 ([61], pp. 70–72]. *Let ϵ be a positive constant. Then*

$$(1 - \epsilon) \frac{\log n}{\log \log n} < \Delta(T(n)) < (1 + \epsilon) \frac{\log n}{\log \log n} \quad a.s.$$

We can then deduce the following theorem.

Theorem 4.5.4. *Let ϵ be a positive constant. Then*

$$\wp(T(n)) \leq 1 + \log_2 \left(1 + \frac{1}{(1 - \epsilon) \frac{\log n}{\log \log n} - 1} \right) \quad a.s.$$

Proof. The lower bound from Theorem [4.5.3](#) guarantees that for

$$m \geq (1 - \epsilon) \frac{\log n}{\log \log n} + 1,$$

S_m is a subtree of $T(n)$ almost surely, and therefore

$$\wp(T(n)) \leq \wp(S_m)$$

almost surely. The right hand side is given by [\(4.5.1\)](#), and substituting for m completes the proof. \square

This bound tends to 1 as $n \rightarrow \infty$, confirming our earlier hypothesis, so the maximum generalised roundness of a tree on n vertices chosen uniformly at random is arbitrarily close to 1 almost surely.

4.6 Probabilistic bounds

Theorem [4.5.4](#) gives an asymptotic bound for the maximum generalised roundness of a random tree, but it does not give any information about the rate of convergence. A more challenging problem is to answer the following question: given (small) $\delta > 0$ and $p \in (0, 1)$, how large does n need to be in order that

$$\mathbb{P}(\wp(T(n)) < 1 + \delta) \geq 1 - p. \quad (4.6.1)$$

As in the last section, we know that $\wp(T) < 1 + \delta$ if $\Delta(T) > k := \lfloor (1 - 2^{-\delta}) \rfloor^{-1}$, and so it is sufficient to determine how large n needs to be in order that $\mathbb{P}(\Delta(T(n)) > k) \geq 1 - p$, or equivalently that

$$\mathbb{P}(\Delta(T(n)) \leq k) \leq p. \quad (4.6.2)$$

We will see in Subsection [4.6.1](#) that using the bounds in Theorem [4.5.3](#) makes the computations rather intractable. However, it is clear that for any $\beta \in (0, 1)$, one has the weaker result that $\Delta(T(n)) > (\log n)^\beta$ a.s.

Our aim in this section is then, given β and p , to find a constant $M = M_{\beta,p}$ such that³ for all $n \geq M$ we have

$$\mathbb{P}\left(\Delta(T(n)) \leq (\log n)^\beta\right) \leq p \quad (4.6.3)$$

and hence

$$\mathbb{P}\left(\wp(T(n)) \leq 1 + \log_2\left(1 + \frac{1}{(\log M)^\beta - 1}\right)\right) \geq 1 - p. \quad (4.6.4)$$

Ideally, one would like a small lower bound on the values of n satisfying (4.6.3) and a small lower bound for $\wp(T(n))$ in (4.6.4), but these objectives are obviously in conflict. It is clear that if β is small then (4.6.3) is satisfied more easily, and hence one can choose a smaller value of M , but that this gives a worse upper bound for $\wp(T(n))$ in (4.6.4). In order to obtain a bound of the form of (4.6.1) one needs to choose β so that

$$\log_2\left(1 + \frac{1}{(\log M)^\beta - 1}\right) \leq \delta.$$

It should be noted that for β near 1, n needs to be very large before

$$(\log n)^\beta < \frac{\log n}{\log \log n},$$

and consequently we do not expect this method to produce even approximately sharp results.

³We ignore issues of rounding the $(\log n)^\beta$ term to an integer, as the resulting error is very small for the magnitude of n we will consider.

4.6.1 Preliminaries

According to Moon [61, p. 72],

$$\begin{aligned}
\mathbb{P}(\Delta(T(n)) \leq k) &= \frac{(n-2)!(e - \frac{1}{k!})^n}{n^{n-2}} \\
&= \frac{n}{n-1} \frac{n!}{n^n} \left(e - \frac{1}{k!}\right)^n \\
&< \frac{n}{n-1} \frac{\sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} e^{1/12n}}{n^n} \left(e - \frac{1}{k!}\right)^n \quad (4.6.5) \\
&= \sqrt{2\pi} \frac{n}{n-1} e^{1/12n} \sqrt{n} \left(1 - \frac{1}{ek!}\right)^n \\
&\leq 3\sqrt{n} \left(1 - \frac{1}{ek!}\right)^n \text{ for } n \geq 7,
\end{aligned}$$

where (4.6.5) makes use of a refinement of Stirling's formula from Robbins [69].

Taking logarithms, we have

$$\begin{aligned}
\log \mathbb{P}(\Delta(T(n)) \leq k) &\leq \log 3 + \frac{1}{2} \log n + n \log \left(1 - \frac{1}{ek!}\right) \\
&< \log 3 + \frac{1}{2} \log n - \frac{n}{ek!},
\end{aligned}$$

since $\log(1-x) > -x$ for x positive.

Now, in order for inequality 4.6.2 to hold, it suffices that

$$\log 3 + \frac{1}{2} \log n - \frac{n}{ek!} \leq \log p,$$

or upon rearranging,

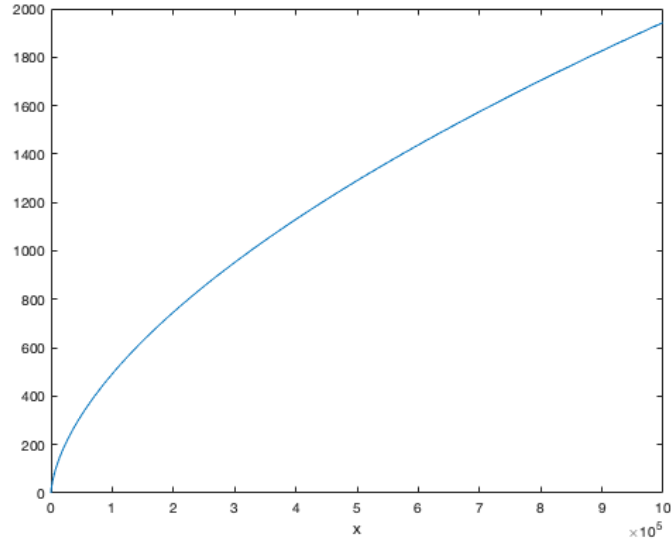
$$\frac{n}{ek!} - \frac{1}{2} \log n \geq \log \frac{3}{p}. \quad (4.6.6)$$

From this point onwards, we change our perspective from a discrete problem to a continuous one in order to make use of tools from calculus, and accordingly change our notation from n to x and from the factorial to the gamma function. Inequality (4.6.6) thus translates to

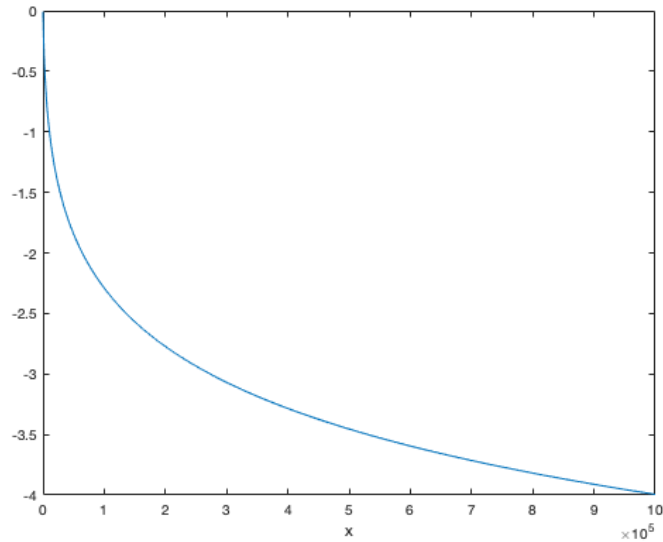
$$\frac{x}{e\Gamma(k+1)} - \frac{1}{2} \log x \geq \log \frac{3}{p}. \quad (4.6.7)$$

Theorem 4.5.3 indicates that if $k \approx \log x / \log \log x$, then for any $p > 0$, inequality (4.6.7) does indeed hold for x sufficiently large, but it is difficult to find a

suitable threshold M_p in this case. To support this conjecture, Figure 4.5a plots the left hand side of (4.6.7) with $k = \log x / \log \log x$, suggesting that the function is increasing towards infinity as required. Indeed, this expression for k appears to be somewhat close to optimal, as Figure 4.5b shows that the same does not apply if we instead use $k = \log x / \sqrt{\log \log x}$.



(a) $k = \frac{\log x}{\log \log x}$



(b) $k = \frac{\log x}{\sqrt{\log \log x}}$

Figure 4.5: Plots of the left-hand side of inequality (4.6.7) with various expressions for k in terms of n .

In order to make later inequalities tractable, we use the slightly weaker expression $k = (\log x)^\beta$, where $\beta \in (0, 1)$. The remainder of this section is therefore devoted to finding an explicit value $M_{\beta,p}$ such that

$$\mathbb{P}\left(\Delta(T(n)) \leq (\log n)^\beta\right) \leq p$$

for all $n \geq M_{\beta,p}$.

Substituting for k in inequality (4.6.7), we obtain

$$\frac{x}{e\Gamma\left((\log x)^\beta + 1\right)} - \frac{1}{2}\log x \geq \log \frac{3}{p},$$

which again appears to hold for x sufficiently large.

Our proof aims to replace the $\Gamma(k+1)$ term with a power of x , say x^α . Thus we want to find a value $M_{\beta,p}$ such that whenever $x \geq M_{\beta,p}$, we have

$$\Gamma\left((\log x)^\beta + 1\right) \leq x^\alpha \tag{4.6.8}$$

and

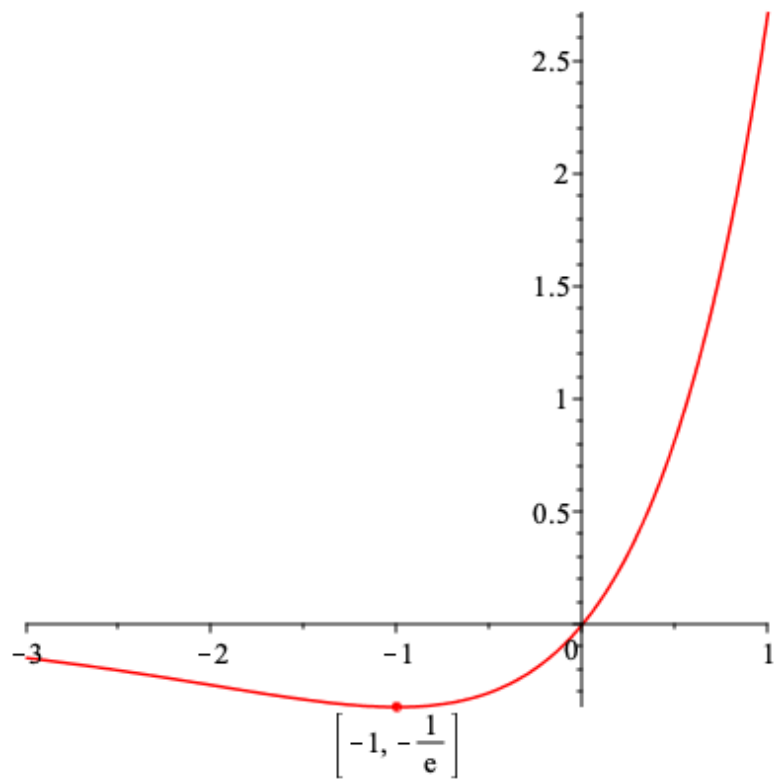
$$\frac{x^{1-\alpha}}{e} - \frac{1}{2}\log x \geq \log 3 - \log p \tag{4.6.9}$$

for some chosen $\alpha \in (0, 1)$.

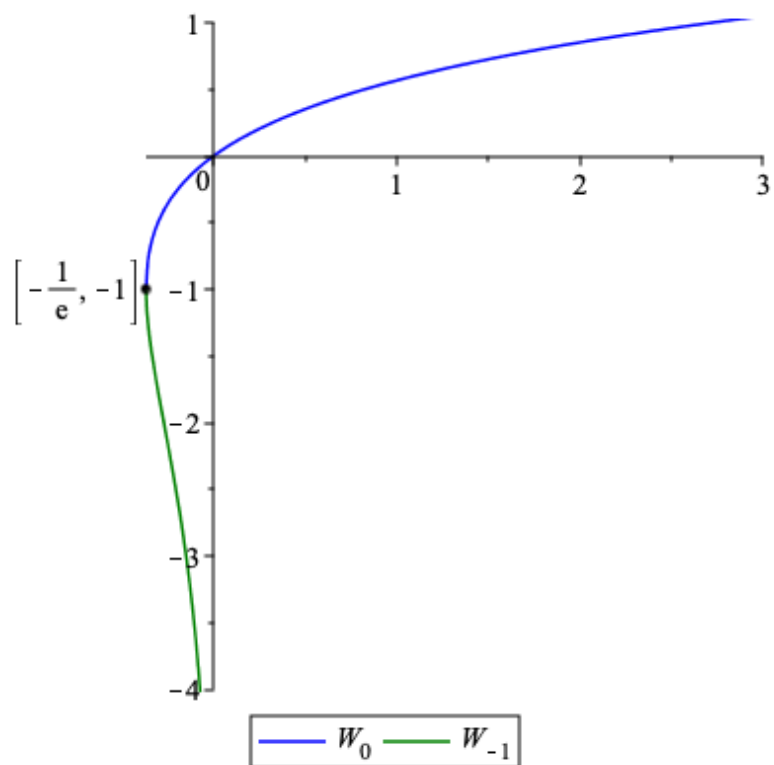
To solve these two inequalities, we will use the Lambert W function, which is the inverse relation to the function $f(t) = te^t$. The graph of f has a single stationary point at $(-1, -1/e)$, as depicted in Figure 4.6a. Thus the inverse is in fact a multi-valued function, consisting of two branches

$$\begin{aligned} W_0 &: \left[-\frac{1}{e}, \infty\right) \rightarrow [-1, \infty) \\ W_{-1} &: \left[-\frac{1}{e}, 0\right) \rightarrow (-\infty, -1] \end{aligned}$$

as shown in Figure 4.6b.



(a) A plot of $f(t) = te^t$.



(b) A plot of the inverse relation of f , the Lambert W function.

Figure 4.6: The function $f(t) = te^t$ and its inverse relation, the multi-valued Lambert W function.

4.6.2 Solution to inequality [4.6.9](#)

Inequality [\(4.6.9\)](#) is the easier of the two. Rearranging from

$$\frac{x^{1-\alpha}}{e} - \frac{1}{2} \log x \geq \log \frac{3}{p}$$

gives

$$x^{\alpha-1} \log \frac{9x}{p^2} \leq \frac{2}{e},$$

and so

$$\left(\frac{9x}{p^2}\right)^{\alpha-1} \log \frac{9x}{p^2} \leq \frac{2}{e} \left(\frac{p^2}{9}\right)^{1-\alpha}.$$

Now we let $y = \log \frac{9x}{p^2}$, so

$$\begin{aligned} ye^{(\alpha-1)y} &\leq \frac{2}{e} \left(\frac{p^2}{9}\right)^{1-\alpha} \\ (\alpha-1)ye^{(\alpha-1)y} &\geq \frac{2(\alpha-1)}{e} \left(\frac{p^2}{9}\right)^{1-\alpha} \\ (\alpha-1)y &\leq W_{-1} \left(\frac{2(\alpha-1)}{e} \left(\frac{p^2}{9}\right)^{1-\alpha} \right) \\ y &\geq \frac{1}{\alpha-1} W_{-1} \left(\frac{2(\alpha-1)}{e} \left(\frac{p^2}{9}\right)^{1-\alpha} \right) \end{aligned}$$

and hence

$$\frac{9x}{p^2} \geq \exp \left(\frac{1}{\alpha-1} W_{-1} \left(\frac{2(\alpha-1)}{e} \left(\frac{p^2}{9}\right)^{1-\alpha} \right) \right).$$

Thus the solution to inequality [\(4.6.9\)](#) is $x \geq f_1(\alpha, p)$, where

$$f_1(\alpha, p) = \frac{p^2}{9} \exp \left(\frac{1}{\alpha-1} W_{-1} \left(\frac{2(\alpha-1)}{e} \left(\frac{p^2}{9}\right)^{1-\alpha} \right) \right)$$

As inequality [\(4.6.8\)](#) does not involve p , f_1 is the only component of our final bound for $M_{\beta, p}$ which will depend on p . It is thus all the more surprising that f_1 itself has very little dependence on p . Table [4.3](#) lists some values of $\log f_1(\alpha, p)$, showing that α is far more influential than p .

For context, recall the original inequality [\(4.6.2\)](#)

$$\mathbb{P}(\Delta(T(n)) \leq k) \leq p.$$

| $\alpha \backslash p$ | 0.5 | 0.1 | 0.01 | 0.001 | 0.00001 |
|-----------------------|-----------|-----------|-----------|-----------|-----------|
| 0.1 | 2.3123 | 2.8614 | 3.3303 | 3.6507 | 4.0945 |
| 0.5 | 4.8869 | 5.6590 | 6.3695 | 6.8749 | 7.5957 |
| 0.75 | 12.2847 | 13.2135 | 14.1988 | 14.9598 | 16.1145 |
| 0.9 | 41.0544 | 41.9325 | 43.0411 | 44.0169 | 45.6819 |
| 0.95 | 98.6901 | 99.4542 | 100.4885 | 101.4614 | 103.2517 |
| 0.99 | 684.0042 | 684.5503 | 685.3255 | 686.0938 | 687.6100 |
| 0.999 | 9462.3025 | 9462.6826 | 9463.2262 | 9463.7694 | 9464.8549 |

Table 4.3: Values of $\log f_1(\alpha, p)$ for various α, p .

It is evident that greatly reducing the failure probability p often requires us to increase n by only one or two orders of magnitude. We will see that compared to the scale of our eventual bound $M_{\beta,p}$, this is almost negligible.

4.6.3 Solution to inequality [4.6.8](#)

The other inequality to be solved is [\(4.6.8\)](#), that is,

$$\Gamma\left((\log x)^\beta + 1\right) \leq x^\alpha.$$

Let

$$F_{\alpha,\beta}(x) = \frac{\Gamma\left((\log x)^\beta + 1\right)}{x^\alpha}.$$

We want $F_{\alpha,\beta}(x) \leq 1$, that is, $\log F_{\alpha,\beta}(x) \leq 0$. Simplifying, we have

$$\begin{aligned} \log F_{\alpha,\beta}(x) &= \log \Gamma\left((\log x)^\beta + 1\right) - \log(x^\alpha) \\ &= \log \Gamma(y^\beta + 1) - \alpha y, \end{aligned}$$

letting $y = \log x$. Now using the same Stirling-type inequality as in [\(4.6.5\)](#),

$$\begin{aligned} \log F_{\alpha,\beta}(x) &\leq \log \left(\sqrt{2\pi} (y^\beta)^{y^\beta + \frac{1}{2}} e^{-y^\beta} e^{\frac{1}{12y^\beta}} \right) - \alpha y \\ &= \frac{1}{2} \log 2\pi + \beta \left(y^\beta + \frac{1}{2} \right) \log y - y^\beta + \frac{1}{12y^\beta} - \alpha y \\ &\leq \beta y^\beta \log y + \beta \log y - y^\beta - \alpha y \end{aligned} \tag{4.6.10}$$

where the last inequality holds for y sufficiently large. We can then use the fact that y^β exceeds $\beta \log y$ to simplify further, although the difference between these two quantities may be substantial for large y , which will result in a suboptimal bound. Therefore

$$\begin{aligned}\log F_{\alpha,\beta}(x) &\leq \beta y^\beta \log y - \alpha y \\ &= y^\beta (\beta \log y - \alpha y^{1-\beta}),\end{aligned}$$

and finally

$$\beta \log y - \alpha y^{1-\beta} \leq 0 \tag{4.6.11}$$

for y sufficiently large, as required.

It remains to analyse just how large y (and hence x) must be in order for inequalities (4.6.10) and (4.6.11) to hold.

4.6.3.1 Solution to inequality 4.6.10

Inequality (4.6.10) holds when

$$\frac{1}{2} \log y^\beta \geq \frac{1}{2} \log 2\pi + \frac{1}{12y^\beta},$$

that is,

$$\frac{y^\beta}{2\pi} \log \frac{y^\beta}{2\pi} \geq \frac{1}{12\pi}.$$

Substituting $z = \log \frac{y^\beta}{2\pi}$, we have

$$\begin{aligned}ze^z &\geq \frac{1}{12\pi} \\ z &\geq W_0\left(\frac{1}{12\pi}\right)\end{aligned}$$

We can then substitute back for y to obtain

$$y \geq \left(2\pi \exp\left(W_0\left(\frac{1}{12\pi}\right)\right)\right)^{\frac{1}{\beta}},$$

and thus $x \geq f_2(\beta)$, where

$$f_2(\beta) = \exp\left(\left(2\pi \exp\left(W_0\left(\frac{1}{12\pi}\right)\right)\right)^{\frac{1}{\beta}}\right).$$

Upon evaluating the constants in this expression, we find that

$$f_2(\beta) \approx \exp\left(6.4477^{\frac{1}{\beta}}\right).$$

In particular, f_2 is decreasing, so in fact inequality (4.6.10) becomes less restrictive on x as β approaches 1. Table 4.4 lists several values of $f_2(\beta)$ for $\beta \in (0, 1)$.

| β | $f_2(\beta)$ |
|---------|-----------------------|
| 0.5 | 1.13×10^{18} |
| 0.75 | 16287.4690... |
| 0.9 | 2782.8156... |
| 0.95 | 1226.8893... |
| 0.99 | 713.5465... |
| 0.999 | 638.9054... |

Table 4.4: Values of $f_2(\beta)$ for various $\beta \in (0, 1)$.

The limiting value is approximately

$$\lim_{\beta \rightarrow 1^-} f_2(\beta) \approx 631.2591.$$

4.6.3.2 Solution to inequality (4.6.11)

Inequality (4.6.11) holds when

$$\beta \log y - \alpha y^{1-\beta} \leq 0,$$

so upon rearranging we have

$$y^{\beta-1} \log y \leq \frac{\alpha}{\beta}.$$

Here we let $z = \log y$, simplifying the inequality to

$$ze^{(\beta-1)z} \leq \frac{\alpha}{\beta},$$

that is,

$$(\beta - 1)ze^{(\beta-1)z} \geq \frac{\alpha(\beta - 1)}{\beta}.$$

Now, if $\alpha > e \frac{\beta}{1-\beta}$ then this inequality holds for all z , and hence for all $x > 1$. Otherwise, we require

$$(\beta - 1)z \leq W_{-1} \left(\frac{\alpha(\beta - 1)}{\beta} \right)$$

and thus

$$z \geq \frac{1}{\beta - 1} W_{-1} \left(\frac{\alpha(\beta - 1)}{\beta} \right).$$

Again we substitute back to find

$$y \geq \exp \left(\frac{1}{\beta - 1} W_{-1} \left(\frac{\alpha(\beta - 1)}{\beta} \right) \right)$$

so the solution is $x \geq f_3(\alpha, \beta)$, where

$$f_3(\alpha, \beta) = \begin{cases} 1 & \text{if } \alpha > \frac{e\beta}{1-\beta} \\ \exp \left(\exp \left(\frac{1}{\beta - 1} W_{-1} \left(\frac{\alpha(\beta - 1)}{\beta} \right) \right) \right) & \text{if } \alpha \leq \frac{e\beta}{1-\beta}. \end{cases}$$

Note however that the first case can only apply if $\frac{e\beta}{1-\beta} < 1$, that is, $\beta < \frac{1}{1+e}$, whereas we are primarily concerned with β close to 1.

4.6.4 Conclusion

Combining the above results, we have that inequalities (4.6.8) and (4.6.9) hold for all $x \geq M_{\beta,p}$, where

$$M_{\beta,p} = \max \left(f_2(\beta), \min_{\alpha \in (0,1)} [\max(f_1(\alpha, p), f_3(\alpha, \beta))] \right). \quad (4.6.12)$$

Now for fixed β and p , f_1 is increasing and approaches infinity as $\alpha \rightarrow 1$, while f_3 is decreasing in the first argument. Therefore the optimal choice for α is that which equates $f_1(\alpha, p)$ and $f_3(\alpha, \beta)$. This value can be found numerically using the bisection method, and comparing to $f_2(\beta)$ gives $M_{\beta,p}$. Furthermore, as documented earlier in Table 4.3, the value of p has almost no effect so all values in Table 4.5 were calculated with $p = 0.001$.

| β | $\log M_{\beta,0.001}$ |
|---------|-----------------------------|
| 0.7 | 29.0466... |
| 0.75 | 493.5846... |
| 0.8 | 47507.2819... |
| 0.85 | $8.8059... \times 10^7$ |
| 0.9 | $7.8520... \times 10^{14}$ |
| 0.95 | $3.2384... \times 10^{38}$ |
| 0.99 | $3.9150... \times 10^{280}$ |

Table 4.5: Values of $M_{\beta,0.001}$ for various $\beta \in (0, 1)$.

Note that in all of these cases,

$$f_2(\beta) \ll \min_{\alpha \in (0,1)} [\max(f_1(\alpha, p), f_3(\alpha, \beta))],$$

so $M_{\beta,p}$ is calculated exclusively from the latter expression.

Summarising, we have now established the following result.

Theorem 4.6.1. *Suppose that $\beta, p \in (0, 1)$ and that $M_{\beta,p}$ is defined as in (4.6.12). Then for all $n \geq M_{\beta,p}$, we have*

$$\mathbb{P} \left(\Delta(T(n)) \leq (\log n)^\beta \right) \leq p,$$

and hence

$$\mathbb{P} \left(\wp(T(n)) < 1 + \log_2 \left(1 + \frac{1}{(\log M_{\beta,p})^\beta - 1} \right) \right) \geq 1 - p$$

for $n \geq M_{\beta,p}$.

Using the values from Table 4.5, we have:

$$\begin{aligned} \beta = 0.7 : \quad & \mathbb{P}(\wp(T(n)) \leq 1.1304) \geq 0.999 \text{ for } n \geq 4.1189 \times 10^{12} \\ \beta = 0.8 : \quad & \mathbb{P}(\wp(T(n)) \leq 1.0002617) \geq 0.999 \text{ for } n \geq 10^{20632} \\ \beta = 0.9 : \quad & \mathbb{P}(\wp(T(n)) \leq 1 + 5.6714 \times 10^{-14}) \geq 0.999 \text{ for } n \geq 10^{3.4101 \times 10^{14}} \\ \beta = 0.99 : \quad & \mathbb{P}(\wp(T(n)) \leq 1 + 2.3571 \times 10^{-278}) \geq 0.999 \text{ for } n \geq 10^{1.7003 \times 10^{280}}. \end{aligned}$$

In general, as $\beta \rightarrow 1$, we have for $n \geq M_{\beta,p}$ that

$$\mathbb{P} \left(\wp(T(n)) \lesssim 1 + \frac{1}{(\log M_{\beta,p})^\beta \log 2} \right) \geq 1 - p,$$

since $\log_2(1+t) \approx \frac{t}{\log 2}$ for small t .

4.6.5 An aside: further investigation of inequality [4.6.8](#)

In this subsection, we present our more general findings on inequality [\(4.6.8\)](#), that is,

$$\Gamma \left((\log x)^\beta + 1 \right) \leq x^\alpha.$$

The reader is advised that the results obtained in this subsection are not essential to the overall direction of the chapter, but it is an interesting problem in its own right. Also, the contrast between this inequality and an analogue with $k = \log n / \log \log n$ as in Theorem [4.5.3](#) is indicative of the wider difficulties encountered with the latter form.

We first ask: for what values of α and β does the inequality hold for all $x \geq 1$? The following example is illustrative of our approach.

Proposition 4.6.2. *For $\alpha = \beta = \frac{1}{2}$, inequality [\(4.6.8\)](#) holds for all $x \geq 1$, that is,*

$$\Gamma \left(\sqrt{\log x} + 1 \right) \leq \sqrt{x}.$$

Proof. Let

$$F(x) = \frac{\Gamma \left(\sqrt{\log x} + 1 \right)}{\sqrt{x}}.$$

Then $F(1) = 1$ and

$$F'(x) = \frac{\Gamma \left(\sqrt{\log x} + 1 \right) \left(\psi \left(\sqrt{\log x} + 1 \right) - \sqrt{\log x} \right)}{2x^{3/2} \sqrt{\log x}},$$

where ψ is the digamma function

$$\psi(z) = \frac{d}{dz} \log \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}.$$

Binet's first integral for the logarithm of the gamma function [\[76\]](#) gives

$$\log \Gamma(z) = \left(z - \frac{1}{2} \right) \log z - z + \frac{1}{2} \log(2\pi) + \int_0^\infty \left(\frac{1}{2} - \frac{1}{t} + \frac{1}{e^t - 1} \right) \frac{e^{-tz}}{t} dt,$$

and differentiating gives

$$\psi(z) = \log z - \frac{1}{2z} - \int_0^\infty \left(\frac{1}{2} - \frac{1}{t} + \frac{1}{e^t - 1} \right) e^{-tz} dt.$$

Bernstein's theorem [6] guarantees that this integral is totally monotone and hence non-negative, so we have

$$\psi(z) \leq \log z - \frac{1}{2z} \leq z - 1 - \frac{1}{2z}$$

for $z \geq 1$. It follows that

$$\psi(\sqrt{\log x} + 1) - \sqrt{\log x} \leq -\frac{1}{2(\sqrt{\log x} + 1)} < 0,$$

so $F'(x)$ is negative for all $x \geq 1$ and hence $F(x) \leq 1$, that is,

$$\Gamma(\sqrt{\log x} + 1) \leq \sqrt{x},$$

for all $x \geq 1$ as required. □

Next, we attempt to generalise to $\alpha, \beta \in (0, 1)$. Let

$$F_{\alpha, \beta}(x) = \frac{\Gamma((\log x)^\beta + 1)}{x^\alpha}$$

as in Subsection 4.6.3. Note that $F_{\alpha, \beta}(1) = 1$ for all $\alpha, \beta \in (0, 1)$, so imitating the proof of Proposition 4.6.2 above, we hope to prove that

$$F'_{\alpha, \beta}(x) = \frac{\Gamma((\log x)^\beta + 1) [\psi((\log x)^\beta + 1) \beta (\log x)^{\beta-1} - \alpha]}{x^{\alpha+1}}$$

is negative for all $x \geq 1$. Letting

$$G(x, \beta) = \beta \frac{\psi((\log x)^\beta + 1)}{(\log x)^{1-\beta}},$$

we require

$$\alpha \geq G(x, \beta)$$

for all $x \geq 1$, that is $\alpha \geq g(\beta)$ where

$$g(\beta) = \max_{x \geq 1} G(x, \beta).$$

As an example, using Maple we can evaluate

$$\begin{aligned} g(0.5) &= \max_{x \geq 1} G(x, 0.5) \\ &= \max_{y \geq 0} G(e^y, 0.5) \\ &\approx 0.2347, \end{aligned}$$

achieved when $x \approx 12.6332$. This confirms the result of Proposition [4.6.2](#), but does not extend its conclusion to all $\alpha \in (0, 1)$.

Generalising to other values of β , we note that if x is held constant then G is an increasing function of β . It follows that g is also increasing in β , so for larger β , our choice of α is more restricted. In fact, Maple finds that $g(0.8) \approx 1.1800$, achieved when $y \approx 139.8136$, that is, $x \approx 5.2514 \times 10^{60}$. Thus, for some values of β close to 1, there is no $\alpha \in (0, 1)$ for which we cannot prove the desired inequality

$$\Gamma\left((\log x)^\beta + 1\right) \leq x^\alpha$$

for all $x \geq 1$ by this method of analysing its derivative.

For some applications, we may not require inequality [\(4.6.8\)](#) to hold for all $x \geq 1$. Relaxing this condition slightly permits a proof not only for α judiciously chosen as a function of β , but indeed all α .

Proposition 4.6.3. *For any $\alpha, \beta \in (0, 1)$, we have*

$$\Gamma\left((\log x)^\beta + 1\right) = o(x^\alpha),$$

and hence

$$\Gamma\left((\log x)^\beta + 1\right) \leq x^\alpha$$

for all x sufficiently large.

We make use of the following lemma.

Lemma 4.6.4. *If $\log f(x) = o(\log g(x))$ and $g(x) \rightarrow \infty$ as $x \rightarrow \infty$, then we have $f(x) = o(g(x)^c)$ for any $c > 0$.*

Proof. For any $c > 0$, there exists $M > 0$ such that for all $x > M$,

$$\log f(x) < c \log g(x).$$

As $\log g(x) \rightarrow \infty$ as $x \rightarrow \infty$, this property remains true if any real constant is added to the right hand side⁴ that is, $\forall c > 0, \exists M > 0$ such that $\forall x > M$,

$$\log f(x) < c \log g(x) + k.$$

Exponentiating, we have

$$f(x) < e^k g(x)^c.$$

Now e^k can take any positive value, so

$$f(x) = o(g(x)^c)$$

as required. □

Proof of Proposition 4.6.3. Recalling Stirling's formula, we have

$$\begin{aligned} \Gamma(z) &= \frac{1}{z} \Gamma(z+1) \\ &= \sqrt{\frac{2\pi}{z}} \left(\frac{z}{e}\right)^z \left(1 + O\left(\frac{1}{z}\right)\right), \end{aligned}$$

so taking logarithms we have

$$\begin{aligned} \log(\Gamma(z)) &= \frac{1}{2} \log(2\pi) - \frac{1}{2} \log z + z \log \frac{z}{e} + \log \left(1 + O\left(\frac{1}{z}\right)\right) \\ &= O(z \log z). \end{aligned}$$

Thus we have

$$\begin{aligned} \log \left(\Gamma \left((\log x)^\beta + 1 \right) \right) &= \log \left(\Gamma \left((\log x)^\beta \right) \right) + \log \left((\log x)^\beta \right) \\ &= O \left((\log x)^\beta \log (\log x)^\beta \right) + \log \left((\log x)^\beta \right) \\ &= O \left(\beta (\log x)^\beta \log \log x \right) \\ &= o(\log x), \end{aligned}$$

as $\beta < 1$ and $\log \log x = o \left((\log x)^{1-\beta} \right)$. It then follows by Lemma 4.6.4 that

$$\Gamma \left((\log x)^\beta + 1 \right) = o(x^\alpha)$$

⁴Note that this may not hold if $g(x)$ is bounded; consider for example $f(x) = \exp(1/x)$ and $g(x) = \exp(1/x^2)$.

for any $0 < \alpha < 1$, and hence

$$\Gamma\left((\log x)^\beta + 1\right) \leq x^\alpha$$

for sufficiently large x . □

Finally, recall that inequality (4.6.8) originally arose from our choice to substitute $k = (\log x)^\beta$ in inequality (4.6.7), but we suspect that sharper bounds for $\Delta(T(n))$ exist for $k = \log n / \log \log n$. In this case, the desired inequality is

$$\Gamma\left(\frac{\log x}{\log \log x} + 1\right) \leq x^\alpha. \quad (4.6.13)$$

Naturally, we would like to obtain a result analogous to Proposition 4.6.3, but these efforts fail as

$$\begin{aligned} \log \Gamma\left(\frac{\log x}{\log \log x} + 1\right) &= \log \Gamma\left(\frac{\log x}{\log \log x}\right) + \log\left(\frac{\log x}{\log \log x}\right) \\ &= O\left(\frac{\log x}{\log \log x} \log \frac{\log x}{\log \log x}\right) + \log \log x - \log \log \log x \\ &= O\left(\frac{\log x}{\log \log x} (\log \log x - \log \log \log x)\right) + \log \log x - \log \log \log x \\ &= O\left(\log x - \frac{\log x \log \log \log x}{\log \log x}\right) + \log \log x - \log \log \log x, \end{aligned}$$

which is not necessarily $o(\log x)$. This also hints at the difficulties encountered in our efforts to solve inequality (4.6.13). Following the method of Subsection 4.6.3, we let

$$F_\alpha(x) = \frac{\Gamma\left(\frac{\log x}{\log \log x}\right)}{x^\alpha},$$

and aim to show that $\log F_\alpha(x) \leq 0$ for x sufficiently large. However, simplifying using the lower bound from Robbins's version of Stirling's formula from [69], we have

$$\begin{aligned} \log F_\alpha(x) &\geq \frac{1}{2} \log 2\pi + \left(\frac{\log x}{\log \log x} + \frac{1}{2}\right) (\log \log x - \log \log \log x) \\ &\quad - \frac{\log x}{\log \log x} + \frac{1}{12 \frac{\log x}{\log \log x} + 1}, \end{aligned}$$

in which the dominant positive term $\log x$ barely outweighs the negative terms.

4.7 Empirical p -negative type properties of random trees

The rigorous bounds given in the last section have some shortcomings. In general, in order to ensure that

$$\mathbb{P}(\wp(T(n)) \leq 1 + \delta) \geq 1 - p$$

for a small value of δ , the bounds obtained on n are extremely large. Furthermore it is not immediately clear how these bounds depend on the size of δ . In this section we shall examine empirical data which indicates that the rigorously proven bounds are indeed far short of what appears to be true. As an example, in Section [4.6](#) we proved that 99.9% of random graphs have maximum generalised roundness less than 1.1304 as long as the number of vertices is at least 4.12×10^{12} . However, taking a sample of one thousand random trees on $n = 300$ vertices, each of the trees has maximum generalised roundness smaller than this. Indeed, the empirical data suggests that the the mean maximum generalised roundness of n vertex trees has a quite simple relationship to n .

The same empirical approach indicates additional information about the distribution of $\wp(T(n))$. Recall from Example [2.2.26](#) and Corollary [4.5.2](#) that this distribution is supported on a subset of $(1, \log_2 3] \cup \{2\}$. To refine this further, we study the lower bound

$$d_n^* = \min\{\wp(T) : T \in \mathcal{T}(n)\}.$$

Considering the star graph on n vertices, we clearly have

$$d_n^* \leq \wp(S_n) = 1 + \log_2 \left(1 + \frac{1}{n-2}\right).$$

Our experimentation suggests that this bound cannot be improved.

Conjecture 4.7.1. *The star graph has the smallest maximum generalised roundness among all path metric trees on n vertices, that is,*

$$d_n^* = 1 + \log_2 \left(1 + \frac{1}{n-2}\right).$$

This conjecture has been verified by calculating the maximum generalised roundness of each tree up to isomorphism for $n = 3, 4, \dots, 17$ by our MATLAB procedure,

and for $n = 3, 4, \dots, 12$ by our Maple procedure. This result is ‘computationally dependent’, and for larger n , there may be graphs with maximum generalised roundness close enough to that of S_n to prevent us from definitively identifying the smallest value.

To study other distribution features, we generated many samples of $T(n)$ for various n , and analysed the following statistics.

Definition 4.7.2. Let $m_{n,N}$, $M_{n,N}$, $\bar{\varphi}_{n,N}$ and $s_{n,N}$ be random variables describing the minimum, maximum, mean and standard deviation of N samples of $T(n)$.

Clearly $m_{n,N} \geq d_n^*$ and $M_{n,N} \leq 2$. Our earlier work in Section 4.6 shows that for fixed N , we have $M_{n,N} \rightarrow 1$ in distribution and the same is true of $m_{n,N}$. We also expect $\bar{\varphi}_{n,N} \rightarrow 1$ and $s_{n,N} \rightarrow 0$.

We calculated the maximum generalised roundness of 1000 trees sampled from $T(n)$ for each $n \in \{10, 11, \dots, 99, 100\}$. A plot of the minimum, maximum and average in the sample can be seen in Figure 4.7.

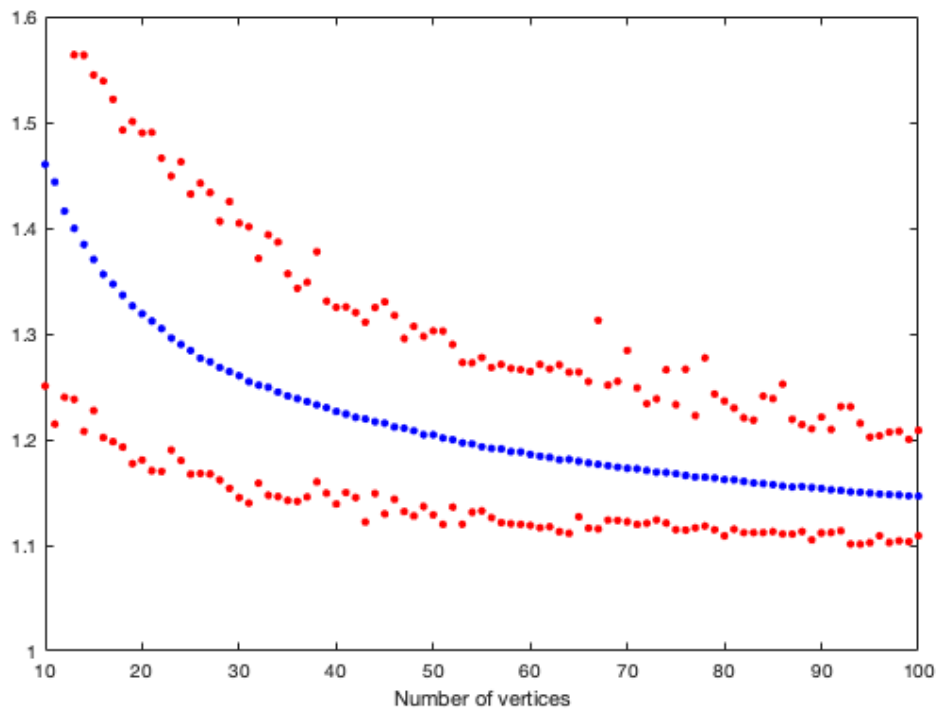


Figure 4.7: A plot of the sample mean, minimum and maximum of the maximum generalised roundness values of $N = 1000$ trees on n vertices generated uniformly at random.

An obvious direction of inquiry relates to how these statistics vary with n . Specifically, what is the rate of decay, and can we fit the data points with a simple curve? After trying various curves, we found that

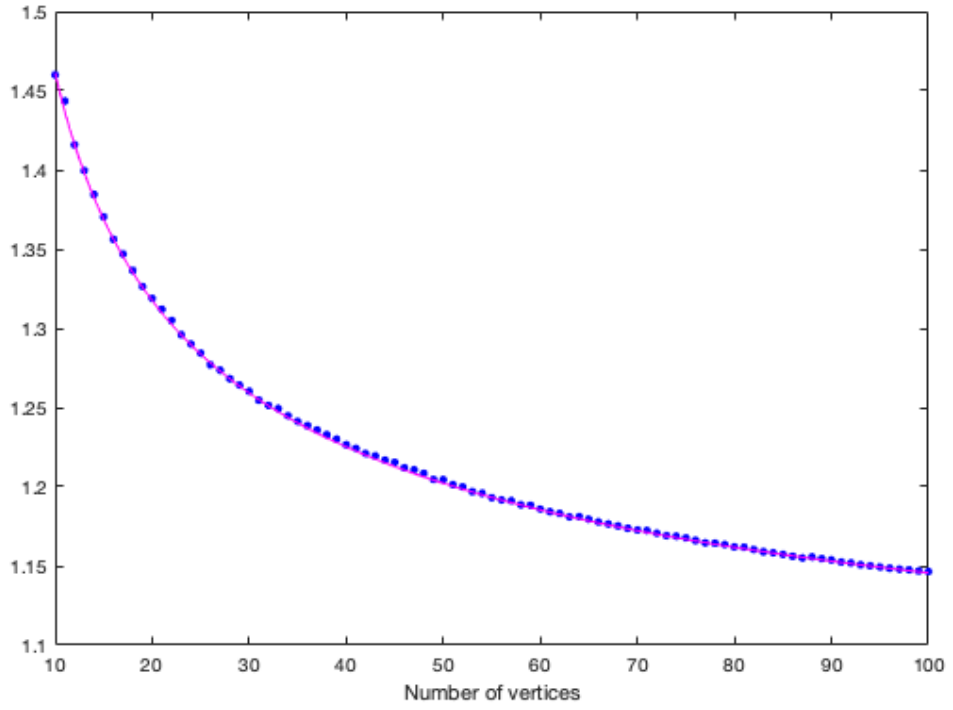
$$\bar{\phi}_{n,1000} \approx 1 + \frac{1.2}{(n-2)^{0.46}}, \quad (4.7.1)$$

as shown in Figure 4.8a. It is surprisingly difficult to distinguish different curves of the form $a(n-2)^{-b}$ as the data is inherently noisy, with especially large relative and absolute error for small n . Indeed, including one or more lower-order terms may also improve the approximation. Even so, it is a pleasant surprise that we can so closely estimate the sample mean.

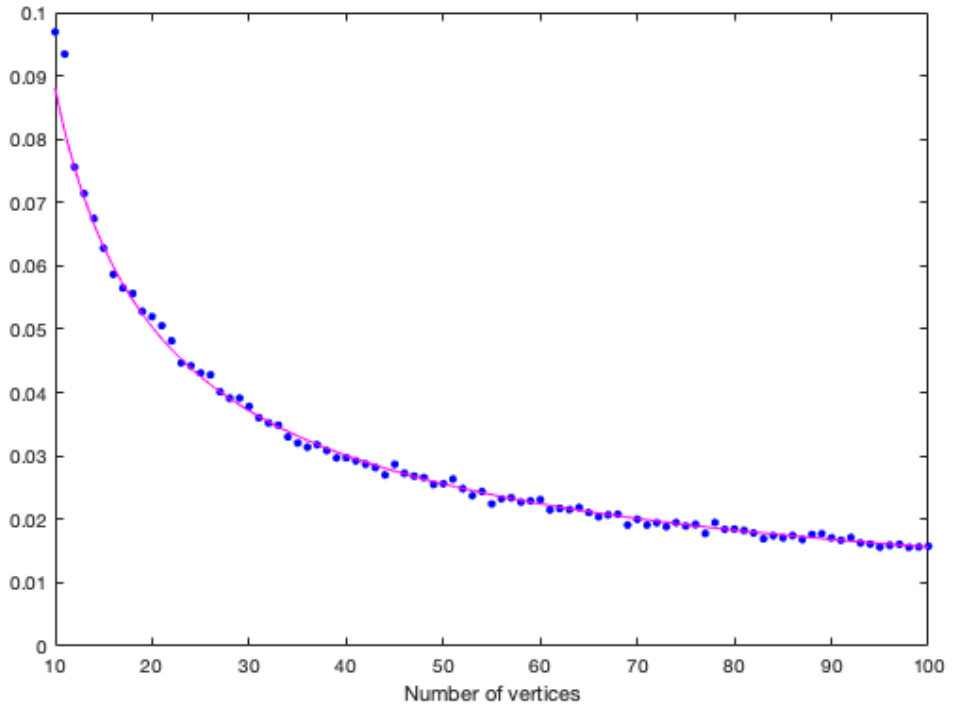
Similarly, the standard deviation is quite well approximated by

$$s_{n,1000} \approx \frac{0.37}{(n-2)^{0.69}}, \quad (4.7.2)$$

as shown in Figure 4.8b.



(a) A plot of the sample mean $\bar{\varphi}_{n,1000}$, fitted with the curve (4.7.1).



(b) A plot of the sample standard deviation $s_{n,1000}$, fitted with the curve (4.7.2).

Figure 4.8: Curve fitting for the sample mean and standard deviation of one thousand trees on n vertices, $10 \leq n \leq 100$.

While the mean value varies little with the sample chosen, it is to be expected that the extreme values will not be so repeatable, and hence it is difficult to meaningfully fit a curve through these points. Indeed, among ten samples for $n = 50$, we observe values of $\bar{\varphi}_{50,1000}$ between 1.2031 and 1.2052, whereas

$$1.1207 \leq m_{50,1000} \leq 1.1444$$

and

$$1.2888 \leq M_{50,1000} \leq 1.3546,$$

where all values are rounded to four decimal places.

However, it is notable that $[m_{n,1000}, M_{n,1000}]$ is a $(1000 - 1)/(1000 + 1) \approx 99.8\%$ prediction interval, that is, an additional sample will be either less than $m_{n,1000}$ or greater than $M_{n,1000}$ each with probability about 0.001. Under the assumption that the distribution is approximately normal, one would expect the following heuristic bounds:

$$\begin{aligned} \mathbb{P}(\varphi(T(n)) < \bar{\varphi}_{n,1000} - 3s_{n,1000}) &< 0.001 \\ \text{and} \quad \mathbb{P}(\varphi(T(n)) < \bar{\varphi}_{n,1000} + 3s_{n,1000}) &> 0.999, \end{aligned}$$

and hence

$$m_{n,1000} \approx \bar{\varphi}_{n,1000} - 3s_{n,1000} \text{ and } M_{n,1000} \approx \bar{\varphi}_{n,1000} + 3s_{n,1000}.$$

However, Figure [4.9](#) indicates that the sample minimum is often slightly outside the predicted range and the sample maximum even more so, suggesting that the distribution of $\varphi(T(n))$ has long tails⁵, particularly on the right side. We correct for this by extending to four standard deviations on the left side and six on the right, to obtain the bounds

$$L(n) = 1 + \frac{1.2}{(n-2)^{0.46}} - 4 \left(\frac{0.37}{(n-2)^{0.69}} \right)$$

and

$$U(n) = 1 + \frac{1.2}{(n-2)^{0.46}} + 4 = 6 \left(\frac{0.37}{(n-2)^{0.69}} \right),$$

⁵This indicates that the underlying distribution is not normal, and further investigation of this distribution is a direction for future research.

so that ‘typically’

$$\mathbb{P}(\wp(T(n)) < L(n)) < 0.001 \text{ and } \mathbb{P}(\wp(T(n)) > U(n)) < 0.001.$$

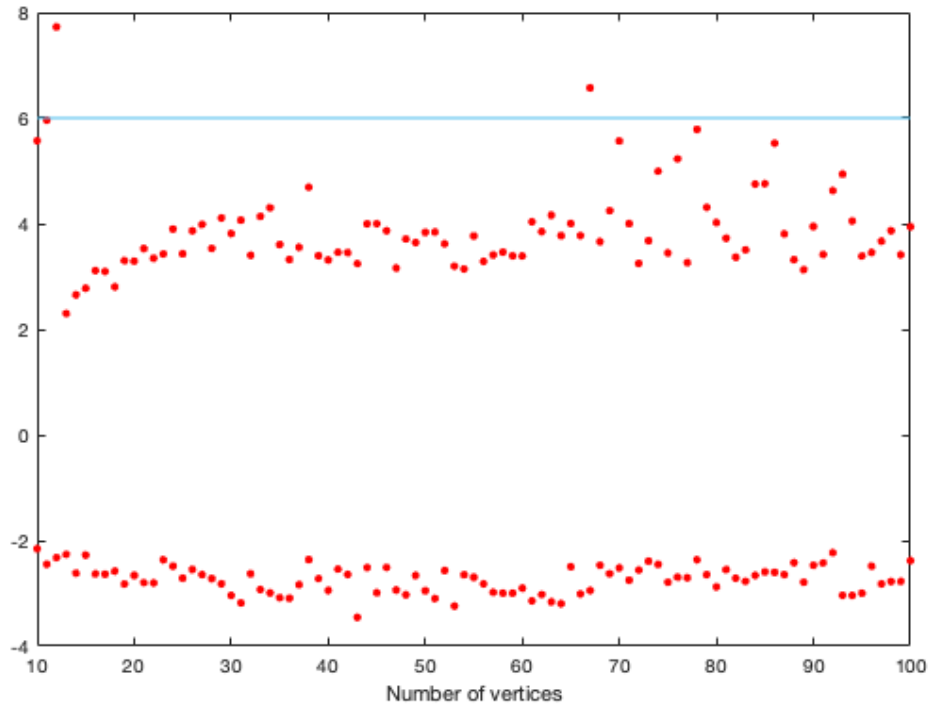


Figure 4.9: A plot of the z -scores of our observed values of $m_{n,1000}$ and $M_{n,1000}$, relative to $N(\bar{\wp}_{n,1000}, s_{n,1000}^2)$.

The most notable outlier occurs for $n = 67$, where the calculated value of $M_{67,1000}$ is more than six standard deviations from the mean. Repeating the process forty times, each with a new sample of 1000 trees, the maximum Z -score exceeded four only twice, and never exceeded five. Similar behaviour was also observed for other values of n , with samples occasionally containing a tree of particularly large maximum generalised roundness. The evidence thus suggests that despite the large sample size, $M_{n,1000}$ is usually quite far from the actual maximum even among non-degenerate trees. We suspect that this maximum is achieved by some tree very close in structure to the path graph, such as the graph is pictured in Figure 4.10 which is the best known candidate on 100 vertices. Its maximum generalised roundness is 1.5608, well over twenty standard deviations above the mean. This type of tree will be discussed in more detail in Section 6.4.

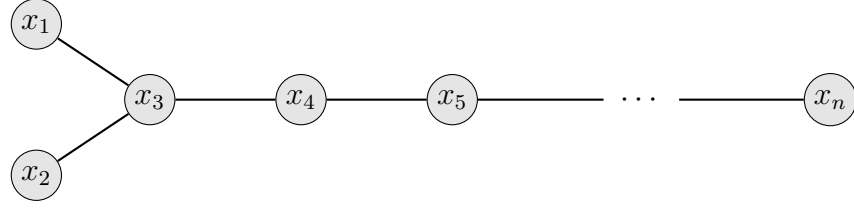


Figure 4.10: A graph on 100 vertices with unusually large maximum generalised roundness.

As remarked in Section [4.2](#), generating trees by some non-uniform random process may change the distribution. Recall from Section [4.4](#) that alternative models of random trees could be produced from randomised versions of Prim’s algorithm and Kruskal’s algorithm, which we will denote by $\mathcal{T}^{(P)}$ and $\mathcal{T}^{(K)}$ respectively. Our data, available at <https://tinyurl.com/desilva-mgr>, shows that $\mathcal{T}^{(K)}$ yields a smaller sample mean, maximum and minimum, and the results from $\mathcal{T}^{(P)}$ are much smaller still. Recalling that the maximum generalised roundness is small for ‘star-like’ trees and large for ‘path-like’ trees, these results are consistent with our earlier calculations for $n = 4$ and suggest that the distributions $\wp(T^{(P)}(n))$ and $\wp(T^{(K)}(n))$ are produced by similar distortions of $\wp(T(n))$ for larger n .

CHAPTER 5

Maximum generalised roundness of connected graphs

Recall from Chapter 2 that any metric space can be represented by a complete graph with appropriately chosen edge weights. The reverse construction is more interesting, as we seek to construct metric spaces from graphs and relate properties of the metric space to its underlying graph theoretic properties. Of particular interest are path metric graphs, where our metric is the distance between the vertices of an unweighted graph, and we can calculate the maximum generalised roundness accordingly.

Previous work on this subject has focused on the extremal maximum generalised roundness values among families of graphs. As a simple example, we have already established that path metric graphs have maximum generalised roundness in $(0, 2]$, with the exception of the complete graph. We will see that the lower bound can be improved when we restrict ourselves to graphs on n vertices. However, little is known about the ‘typical’ graph on n vertices, in any sense of the word.

In this chapter, we will examine the maximum generalised roundness properties of the path metric graphs on n vertices, and apply ideas from random graph theory to study the properties of typical graphs, as well as those of sparse and dense graphs. As in the previous chapter, we are able to prove significant new theorems on the limiting behaviour, and we have empirical evidence for stronger heuristic results.

5.1 p -negative type properties of graphs

The existing literature contains various results on the maximum generalised roundness of graph metric spaces in terms of the diameter of the graph. Recall that the diameter of (G, d) is

$$\text{diam}(G, d) = \max_{v, w \in V(G)} d(v, w).$$

For path metric graphs in particular, the diameter is always a positive integer. If $\text{diam}(G, d) = 1$, the graph is complete, so $\wp(G) = \infty$ from Example [2.2.21](#).

Deza and Maehara proved a sharp lower bound for the maximum generalised roundness of a path metric graph on n vertices with diameter two.

Theorem 5.1.1 ([\[23\]](#), Theorem 6]. *Let*

$$c_{n,2}^* = \min\{\wp(G, d) : |G| = n \text{ and } \text{diam}(G, d) = 2 \text{ where } d \text{ is the path metric}\}.$$

Then

$$c_{n,2}^* = \log_2 \left(\frac{1}{1 - \gamma(n)} \right)$$

where

$$\gamma(n) = \frac{1}{2} \left(\frac{1}{\lfloor n/2 \rfloor} + \frac{1}{\lceil n/2 \rceil} \right).$$

Naturally, we would like to study the analogous quantity over all metric graphs on n vertices.

Definition 5.1.2. Let c_n be the infimum of the maximum generalised roundness among all n -vertex metric graphs, that is,

$$c_n = \inf\{\wp(G, d) : |G| = n\}.$$

The first notable lower bound for c_n was expressed purely in terms of n by Weston.

Theorem 5.1.3 ([\[79\]](#), Theorem 4.3]. *Define the sequence ζ recursively by $\zeta(1) = \zeta(2) = 1$ and $\zeta(k) = \zeta(k-1) + \zeta(k-2) + 1$ for $k \geq 3$. Then*

$$c_n \geq \log_2 \left(1 + \frac{1}{(2n)^{2\zeta(n)}} \right).$$

Weston recognised that this bound was likely to be far from sharp, and later generalised Theorem [5.1.1](#) to general metric graphs. Note that the earlier theorem makes reference to the diameter of a path metric graph. In the general case, we instead use the scaled diameter

$$\mathcal{D}(G, d) = \frac{\text{diam}(G, d)}{\min_{\substack{v, w \in V(G) \\ v \neq w}} d(v, w)},$$

since generalised roundness properties are invariant under scaling by Theorem 2.1.14.

Theorem 5.1.4 ([78, Theorem 4.1]). *Let*

$$c_{n,t} = \inf\{\wp(G, d) : |G| = n \text{ and } \mathcal{D}(G, d) = t\}.$$

Then

$$c_{n,t} \geq \log_t \left(\frac{1}{1 - \gamma(n)} \right)$$

where $\gamma(n)$ is as defined in Theorem 5.1.1, with equality for $t \in (1, 2]$.

It is clear that $c_n \leq c_{n,2}^*$, and Deza and Maehara conjectured that this is in fact the correct value for c_n .

Conjecture 5.1.5 ([23]). *The infimum of the maximum generalised roundness of n -vertex metric graphs is*

$$c_n = c_{n,2}^* = \log_2 \left(\frac{1}{1 - \gamma(n)} \right)$$

for all n .

Deza and Maehara [23] also provided the explicit construction for an n -vertex graph with this maximum generalised roundness. Recall from Proposition 3.6.2 that the maximum generalised roundness of the complete bipartite graph $K_{m,n}$ with the path metric is

$$\wp(K_{m,n}) = \log_2 \left(\frac{2nm}{2nm - n - m} \right).$$

Choosing parts of size $\lceil n/2 \rceil$ and $\lfloor n/2 \rfloor$ achieves the desired result.

On the other hand, recall from Proposition 2.2.22 that the maximum generalised roundness of a path metric graph is at most 2, with the exception of complete graphs, where $\wp(K_n) = \infty$ as proved in Example 2.2.21. While a maximum generalised roundness of two is achieved by the path graphs in Example 2.2.26, it is in some sense almost as much of an outlier as the value ∞ . Path metric graphs cannot have $\wp(G) = 2 - \epsilon$ for small ϵ , as Sánchez established the following gap in maximum generalised roundness values.

Theorem 5.1.6 ([71, Theorem 3.2]). *If G is a connected path metric graph, then $\wp(G) \notin (\log_2(2 + \sqrt{3}), 2)$.*

Proof. First, we note that for $n < 4$, G is either a complete graph or a path graph, with maximum generalised roundness ∞ or 2 respectively, so the theorem is satisfied.

Now, when $n \geq 4$, we can again discount complete graphs and path graphs. Sánchez proved that a graph which is neither of these must contain a metrically embedded cycle, in which case $\wp(G) \leq \log_2 3$, or a metrically embedded copy of one of the three graphs depicted in Figure 5.1.

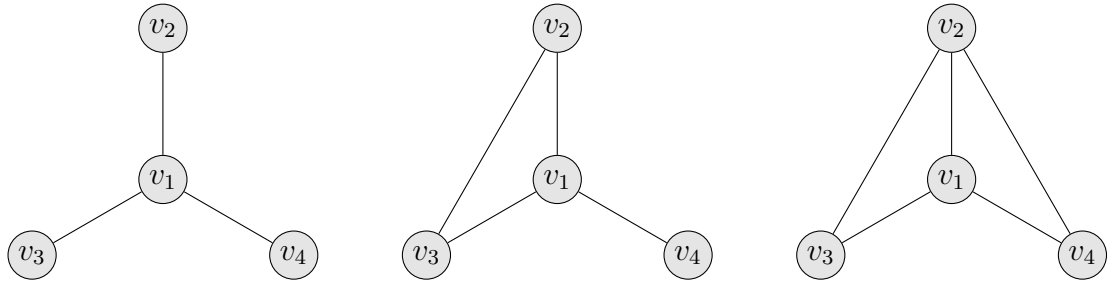


Figure 5.1: Three graphs on four vertices.

These graphs have maximum generalised roundness $\log_2 3$, $\log_2(2 + \sqrt{3})$ and $\log_2 3$ respectively by 3.1.5, so $\wp(G) \leq \log_2(2 + \sqrt{3})$ by Theorem 2.1.17. \square

5.2 The graphs on n vertices

Aside from these extreme values, little is known about the ‘typical’ graph on n vertices. We seek to understand the maximum generalised roundness properties of the family of path metric graphs on n vertices for large n . Recall from Definition 2.2.11 that a graph must be connected in order to define the path metric.

Just as in Section 4.2, we must distinguish between labelled and unlabelled connected graphs. Let ℓ_n be the number of connected graphs on n vertices labelled $1, 2, \dots, n$. In the labelled case, there is no known simple formula for ℓ_n , but instead each term of the sequence can be calculated from the recurrence

$$\sum_{k \geq 1} k \ell_k 2^{\binom{n-k}{2}} = n 2^{\binom{n}{2}}.$$

A proof using exponential generating functions is provided by Wilf [80, p. 87]. We will see later that of the $2^{\binom{n}{2}}$ graphs on n vertices, relatively few are not connected. Note however that many of the graphs counted by ℓ_n are isomorphic to each other.

On the other hand, we denote u_n for the number of connected graphs on n unlabelled vertices, counting only one representative of each isomorphism class.

Again there is no known closed form for u_n , although $u_n \approx \ell_n/n!$ provides a rough approximation.

The numbers of labelled and unlabelled connected graphs for small n are given in Table 5.1.

| n | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|---|---|----|-----|-------|---------|-----------|-------------|
| ℓ_n | 1 | 4 | 38 | 728 | 26704 | 1866256 | 251548592 | 66296291072 |
| u_n | 1 | 2 | 6 | 21 | 112 | 85 | 11117 | 261080 |

Table 5.1: The number of labelled and unlabelled connected graphs on n vertices.

Appendix C catalogues the maximum generalised roundness of each tree on four or five unlabelled vertices. Echoing our discussion of labelled and unlabelled trees in Section 4.2, it will eventually be infeasible to list all the connected graphs on n vertices, so we must at some point take a random sample of these graphs. We are not able to efficiently generate unlabelled connected graphs uniformly at random, nor with another useful probability distribution. For the labelled case however, there are well-established methods to generate random graphs, which can be easily adapted to produce only connected graphs.

In the next section, we discuss the Erdős-Rényi model of random graphs, in which edges appear independently with some fixed probability q . In Section 5.4 we study the special case where $q = \frac{1}{2}$, and prove that the maximum generalised roundness of a random graph from this “uniform model” approaches zero with high probability as the number of vertices $n \rightarrow \infty$. Section 5.5 then discusses the probabilistic bounds implied by this proof, as the experimental data suggests that our known bounds are far from sharp. Indeed, sampling 1000 graphs on n vertices for $10 \leq n \leq 100$, the average lies quite close to

$$\frac{1.9}{(n-2)^{0.55}}$$

with standard deviation approximately

$$\frac{0.91}{(n-2)^{1.18}}.$$

Further analysis of the extremal values allows us to estimate 99.8% intervals $[L(n), U(n)]$ for the maximum generalised roundness of a uniform random graph.

In Section 5.6, we identify that the “uniform model” almost exclusively represents graphs with about half the possible edges, and we must consider other values of q in order to understand the properties of graphs with either fewer or more edges. We are able to extend the results of Section 5.4 for general q , and then in Section 5.7 we examine empirical data for additional insights on the relationship between maximum generalised roundness and edge probability.

5.3 Random graphs

Random graphs can be produced by various models, most notably the model proposed by Gilbert [38] and developed further by Erdős and Rényi [33], in which we place each possible edge independently with probability q .

Definition 5.3.1. Define the Erdős-Rényi model of random graphs

$$\mathcal{G}(q) = (\mathcal{G}(n), \mu_{n,q})_{n \geq 1},$$

where $\mathcal{G}(n)$ is the set of all labelled graphs on n vertices and $\mu_{n,q}(G) = q^m(1-q)^{\binom{n}{2}-m}$ where $m = |E(G)|$.

Notation 5.3.2. In particular, if $q = \frac{1}{2}$, all graphs $G \in \mathcal{G}(n)$ have equal measure, so we refer to the special case $\mathcal{G}(\frac{1}{2})$ as the uniform model, and abbreviate it to \mathcal{G} .

Definition 5.3.3. Let $G(n, q)$ be the discrete graph-valued random variable representing sampling from the Erdős-Rényi model, that is, from the set $\mathcal{G}(n)$ according to the probability measure $\mu_{n,q}$. We will abbreviate $G(n, \frac{1}{2})$ to $G(n)$.

For a full exposition of this model of random graphs, we direct the reader to Blum, Hopcroft and Kannan [7] or Bollobás [9]. For our purposes, it is necessary to modify the Erdős-Rényi model slightly by excluding disconnected graphs, to ensure that the path length is defined for all pairs of vertices. Here, we consider only constant values of q , and under this assumption, we can establish a bound for the proportion of graphs to be discarded.

Proposition 5.3.4. *Let $G = G(n, q)$ be a random graph from the Erdős-Rényi model. Then G is of diameter two, and hence connected, almost surely.*

Proof. Let $X_n(G)$ be the number of pairs of vertices (v, w) in G where $d(v, w) > 2$. Such pairs are not connected by an edge, and do not have a common neighbour. Then by the Markov inequality we have

$$\mathbb{P}(X_n(G) \geq 1) \leq \mathbb{E}(X_n(G)).$$

Using the linearity of the expectation, we can evaluate this expression, as in the proof of [7, Theorem 8.5]. We have

$$\begin{aligned}
\mathbb{E}(X_n(G)) &= \mathbb{E}\left(\sum_{v,w \in V} \mathbb{P}(d(v,w) > 2)\right) \\
&= \binom{n}{2} \times \mathbb{P}(d(v,w) > 2) \\
&= \binom{n}{2} \times \mathbb{P}((v,w) \notin E) \times \left(\prod_{u \in V \setminus \{v,w\}} \mathbb{P}((u,v), (u,w) \notin E)\right) \\
&= \binom{n}{2} (1-q)(1-q^2)^{n-2},
\end{aligned}$$

so

$$\begin{aligned}
\mathbb{P}(X_n(G) = 0) &= \mu_{n,q}(\{G \in \mathcal{G}(n,q) : \text{diam}(G) \leq 2\}) \\
&\geq 1 - \binom{n}{2} (1-q)(1-q^2)^{n-2}.
\end{aligned}$$

Letting $n \rightarrow \infty$, the result follows immediately. \square

In fact, these properties may remain true if q is allowed to vary with n up to certain thresholds, as demonstrated by Blum, Hopcroft and Kannan. These results can be stated concisely using the following asymptotic notation.

Notation 5.3.5. Let $f(n) = \omega(g(n))$ denote that f grows asymptotically faster than g , that is,

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \infty.$$

Proposition 5.3.6 ([7, p. 253]). *The following thresholds govern the properties of the random graph $G = G(n, q(n))$.*

- If $q(n) = \omega\left(\frac{\log n}{n}\right)$, then G is connected almost surely.
- If $q(n) = \omega\left(\frac{\log n}{n}\right)$, then $\text{diam}(G) = O(\log n)$ almost surely.
- If $q(n) = \omega\left(\sqrt{\frac{2 \log n}{n}}\right)$, then $\text{diam}(G) = 2$ almost surely.

Although we will not consider the construction $\mathcal{G}(q(n))$ further, these facts assist in understanding the space of random graphs.

Definition 5.3.7. Define the connected Erdős-Rényi model

$$\mathcal{G}^*(q) = (\mathcal{G}^*(n), \mu_{n,q}^*)_{n \geq 1},$$

where $\mathcal{G}^*(n)$ is the set of all connected labelled graphs on n vertices with probability measure

$$\mu_{n,q}^*(G) = \begin{cases} \frac{\mu_{n,q}(G)}{\sum_{G' \in \mathcal{G}^*(n)} \mu_{n,q}(G')} & \text{if } G \text{ is connected} \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 5.3.8. *Let c be a positive constant. Any graph property which holds in $\mathcal{G}(q)$ with probability $1 - o(n^{-c})$ must also hold with probability $1 - o(n^{-c})$ (and hence almost surely) in $\mathcal{G}^*(q)$.*

Proof. This is a consequence of the fact that

$$\mu_{n,q}(\mathcal{G}^*(n)) \geq 1 - \binom{n}{2}(1-q)(1-q^2)^{n-2},$$

from the proof of Proposition [5.3.4](#). □

Remark 5.3.9. Note that $\mathcal{G}^*(\frac{1}{2})$ generates the connected graphs on n vertices uniformly at random. We will refer to this model as simply \mathcal{G}^* , the connected uniform model.

Notation 5.3.10. Let $G^*(n, q)$ be the discrete graph-valued random variable representing sampling from the connected Erdős-Rényi model. We will abbreviate $G^*(n, \frac{1}{2})$ to $G^*(n)$.

In the next two sections we will study the maximum generalised roundness under the connected uniform model, before returning to the more general connected Erdős-Rényi model in Section [5.6](#).

5.4 p -negative type properties of uniform random connected graphs

We first generated a sample of graphs $G^*(n)$ for various values of n , and calculated the maximum generalised roundness of each graph. This allowed us to approximate the distribution of $\wp(G^*(n))$, which is supported on $[c_{n,2}^*, 2] \cup \{\infty\}$ by Theorem [5.1.1](#). The resulting data suggested that as n grows, the maximum generalised roundness is usually small with occasional larger values, as seen in Figure [5.2](#) and Table [5.2](#).

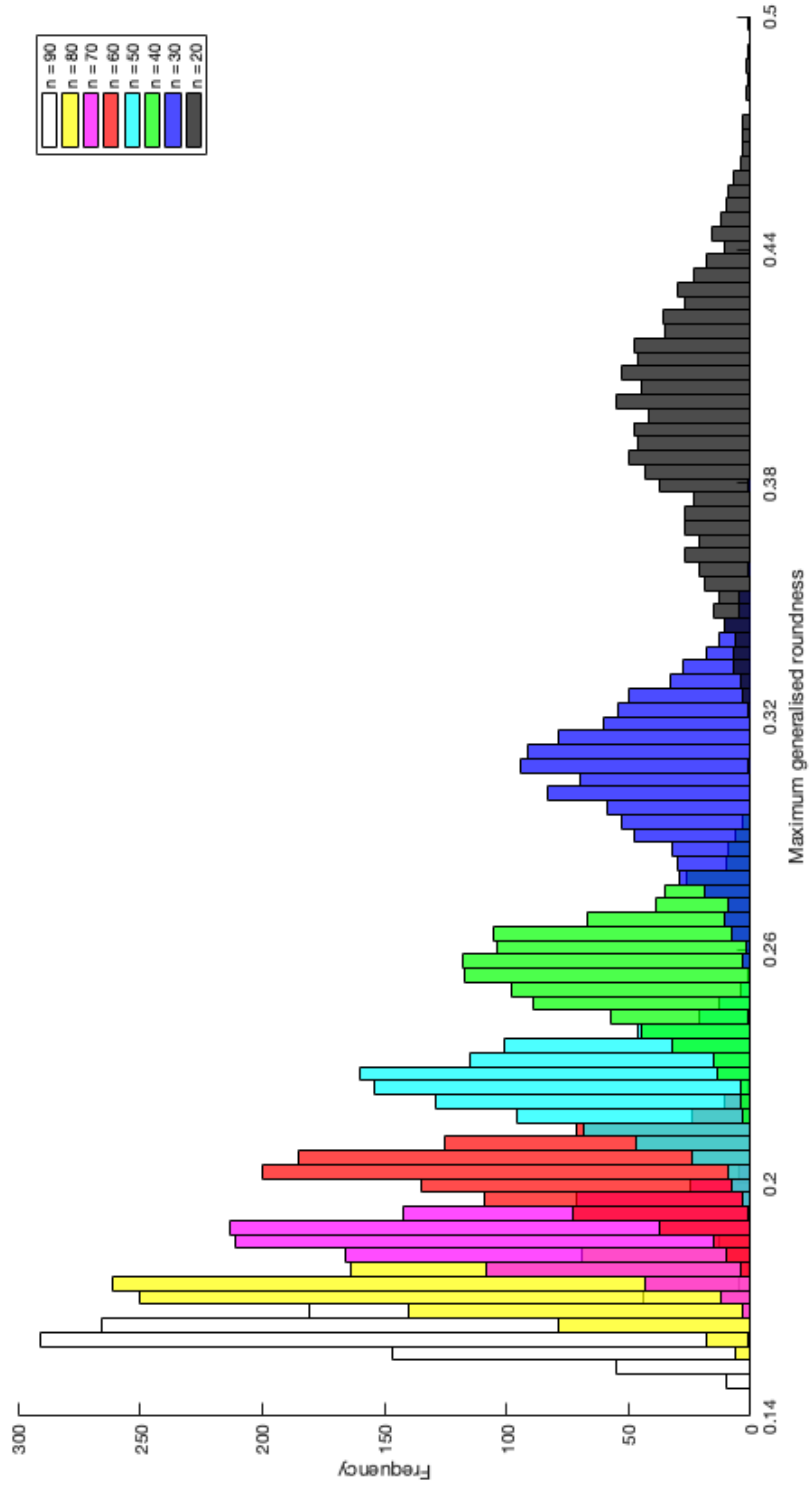


Figure 5.2: Histograms of maximum generalised roundness values of 1000 graphs on n vertices from the connected uniform model, $20 \leq n \leq 90$

| n | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
|--------------|--------|--------|--------|--------|--------|--------|--------|--------|
| \wp_{\max} | 0.4976 | 0.3808 | 0.2944 | 0.2529 | 0.2236 | 0.2026 | 0.1865 | 0.1775 |
| \wp_{\min} | 0.3092 | 0.2442 | 0.2164 | 0.1923 | 0.1771 | 0.1606 | 0.1552 | 0.1473 |
| $\bar{\wp}$ | 0.3984 | 0.3065 | 0.2558 | 0.2250 | 0.2028 | 0.1859 | 0.1723 | 0.1616 |

Table 5.2: The maximum, minimum and average maximum generalised roundness of 1000 graphs on n vertices from the connected uniform model.

To prove these results, we will identify a graph with small maximum generalised roundness which appears with high probability as a metric subspace of $G^*(n)$. This will allow us to conclude that with the same probability, $G^*(n)$ has maximum generalised roundness no larger than that of our chosen graph.

Note however that this does not correspond exactly to the notion of subgraphs from Definition 2.2.6. Unless all valid edges are inherited by the subgraph, some distances of 1 in the original graph will differ from the corresponding distances in the subgraph.

Example 5.4.1. It is easy to confirm that

$$\wp(C_4) = 1 < \infty = \wp(K_5).$$

Although the 4-cycle C_4 is clearly a subgraph of the complete graph K_5 , this does not contradict Theorem 2.1.17 on metric subspaces. We note that C_4 is not a metric subspace of K_5 since the distances do not agree. Some pairs of vertices in C_4 are separated by a distance of 2, but there are no such pairs in K_5 . \square

To overcome this problem, we introduce the notion of induced subgraphs.

Definition 5.4.2. Let $G = (V, E)$ be a graph and let $V' \subseteq V$. The induced subgraph $G[V']$ is the subgraph of G with vertex set V' and edge set $\{(v, w) \in E : v, w \in V'\}$, that is, all the edges of G which join vertices of V' .

Theorem 5.4.3. *If H is an induced subgraph of G with diameter 2, then the path metric on H is the restriction of the path metric on G to the vertices of H .*

Proof. Let $u, v \in H$. We must prove that $d_G(u, v) = d_H(u, v)$.

It is trivial to see that the shortest path in G cannot be improved upon by restricting to H , so $d_G(u, v) \leq d_H(u, v)$.

Suppose for a contradiction that $d_G(u, v) < d_H(u, v)$. Since $d_H(u, v) \leq 2$, we must have $d_G(u, v) = 1$ and $d_H(u, v) = 2$, which is impossible. Therefore the distance between u and v in G and H coincide, so the metric in H is the restriction of the metric in G . \square

Our aim is to find a particular small graph of diameter two which is likely to occur as an induced subgraph of a large random graph, and hence state a probabilistic upper bound on the maximum generalised roundness of the large graph.

A simple application of this idea is sufficient to prove a weak result suggested by Figure 5.2

Theorem 5.4.4. *Under the uniform model, $\wp(G(n)) < 1$ almost surely.*

Proof. Consider the complete bipartite graph $K_{2,3}$, whose maximum generalised roundness is $\wp = \log_2(12/7) < 1$ by Theorem 3.6.2. Suppose a graph G has an induced subgraph H which is isomorphic to $K_{2,3}$. Since $\text{diam } H = 2$, $d_H = d_G|_H$ and hence $\wp(G) \leq \wp(H) < 1$.

We then ask how likely it is that $G = G(n)$ contains an embedded copy of $K_{2,3}$. A crude underestimate can be found by partitioning the vertices of G into the following $\lfloor n/5 \rfloor$ sets: $\{1, 2, 3, 4, 5\}$, $\{6, 7, 8, 9, 10\}$, and so on. There are $\binom{5}{2} = 10$ potential edges among the first five vertices, giving 2^{10} possibilities for the induced subgraph $G[\{1, 2, 3, 4, 5\}]$. Of these, ten are isomorphic to $K_{2,3}$, so

$$\mathbb{P}(G[\{1, 2, 3, 4, 5\}] \simeq K_{2,3}) = \frac{5}{512}.$$

The same probability applies to each of the other sets of five vertices, so the probability that no copy of $K_{2,3}$ appears is at most

$$1 - \left(1 - \frac{5}{512}\right)^{\lfloor n/5 \rfloor}.$$

Thus $\wp(G(n)) < 1$ almost surely. \square

Corollary 5.4.5. *Under the connected uniform model, $\wp(G^*(n)) < 1$ almost surely also.*

Proof. This follows immediately from Theorem 5.4.4 by applying Proposition 5.3.8. \square

Remark 5.4.6. This result confirms that trees have unusual maximum generalised roundness, compared to connected graphs in general.

The proof of Theorem 5.4.4 suggests that this theorem is far from the strongest available result. We considered only very few of the subgraphs of G , and compared them only against the particular graph $K_{2,3}$. Figure 5.2 suggests that a far stronger result is true, namely that $\wp(G) < \epsilon$ almost surely for any $\epsilon > 0$. To prove this hypothesis, we look for metrically embedded subgraphs with even smaller maximum generalised roundness. We return to the complete bipartite graphs studied in Theorem 3.6.2, and for simplicity use the special case where both parts are of equal size.

Theorem 5.4.7. *The maximum generalised roundness of the complete bipartite graph $K_{m,m}$ with the path metric is*

$$\wp(K_{m,m}) = \log_2 \left(1 + \frac{1}{m-1} \right).$$

Applying Theorem 5.4.3 gives the following immediate corollary.

Corollary 5.4.8. *Let G be a graph containing an induced copy of $K_{m,m}$. Then*

$$\wp(G) \leq \log_2 \left(1 + \frac{1}{m-1} \right).$$

We seek a function $m(n)$ which increases slowly to infinity, such that $G^*(n)$ contains an induced copy of $K_{m(n),m(n)}$ almost surely. Here, we make use of the following theorem of Palka, which was proven using the second moment method. While the theorem applies to random graphs from the uniform model \mathcal{G} , the same result will hold when we later move to the connected uniform model \mathcal{G}^* , as discussed in Proposition 5.3.8.

Theorem 5.4.9 ([66, Theorem 2]). *Given two non-decreasing sequences of positive integers $\mathbf{r} = (r_1, r_2, \dots)$ and $\mathbf{w} = (w_1, w_2, \dots)$, such that the sequence $\mathbf{r} + \mathbf{w}$ is increasing, define*

$$B_n(\mathbf{r}, \mathbf{w}) = \max \{ r_k + w_k : \text{there is a copy of } K_{r_k, w_k} \text{ in } G(n) \}.$$

Suppose that $\frac{w_k}{r_k} \rightarrow c$ as $k \rightarrow \infty$, where $0 < c \leq 1$ is a constant. Then for every $\epsilon > 0$,

$$\mathbb{P}(B_n(\mathbf{r}, \mathbf{w}) < (2 - \epsilon) \log_2 n) = o(n^{-1-\delta})$$

and

$$\mathbb{P}(B_n(\mathbf{r}, \mathbf{w}) > (2 + \epsilon) \log_2 n) = o(n^{-k}),$$

where $0 < \delta < 1$ is a constant and k is any integer.

In particular, we choose $\mathbf{r} = \mathbf{w} = (1, 2, 3, \dots)$, and define

$$m(n) = \frac{1}{2} B_n(\mathbf{r}, \mathbf{w}) = \max \{m : \text{there is a copy of } K_{m,m} \text{ in } G(n)\}.$$

Then for every $\epsilon > 0$, we have

$$(1 - \epsilon) \log_2 n < m(n) < (1 + \epsilon) \log_2 n \quad a.s. \quad (5.4.1)$$

Note that only the lower bound here is useful to us.

We now know that $G(n)$ contains an induced copy of $K_{m(n), m(n)}$ almost surely, and that this subgraph has diameter 2, so we have

$$\begin{aligned} \wp(G(n)) &\leq \wp(K_{m(n), m(n)}) \\ &= \log_2 \left(1 + \frac{1}{m(n) - 1} \right) \\ &< \log_2 \left(1 + \frac{1}{(1 - \epsilon) \log_2 n - 1} \right) \quad a.s. \end{aligned}$$

and therefore by Proposition [5.3.8](#), we have the following theorem.

Theorem 5.4.10. *In the connected uniform model,*

$$\wp(G^*(n)) < \log_2 \left(1 + \frac{1}{(1 - \epsilon) \log_2 n - 1} \right) \quad a.s.$$

for every $\epsilon > 0$.

Note that this bound tends to zero as $n \rightarrow \infty$, so the maximum generalised roundness of a path metric graph chosen uniformly at random from $\mathcal{G}^*(n)$ can be made arbitrarily small almost surely, by choosing n sufficiently large.

5.5 Empirical p -negative type properties of uniform random connected graphs

Although we now know that $\wp(G^*(n))$ eventually approaches zero, Theorem [5.4.10](#) says very little about the rate of this convergence. An obvious question is: how

large must n be so that $G^*(n)$ contains $K_{m,m}$ ‘with high probability’?

To be more precise, for $\epsilon > 0$ and $0 < p < 1$, we seek $M = M_{\epsilon,p}$ such that if $n \geq M$ then

$$\mathbb{P}(\wp(G^*(n)) < \epsilon) > 1 - p.$$

Again we simplify by interchanging $G^*(n)$ for the more accessible (and almost identical) $G(n)$. Following Palka’s proof of Theorem 5.4.9 (see [66, Theorem 2]), we have

$$\mathbb{P}(G(n) \text{ has an induced copy of } K_{m,m}) > 1 - \frac{m^6}{\frac{1}{2}(n - 4m + 2)(n - 4m + 1)}. \quad (5.5.1)$$

To ensure that this probability is greater than $1 - p$, it is sufficient to have

$$(n - 4m + 2)(n - 4m + 1) > \frac{2m^6}{p}.$$

Solving this inequality, we obtain

$$M = \frac{8m - 3 + \sqrt{1 + \frac{8m^6}{p}}}{2}.$$

This bound is still quite large. As an example, setting $p = 0.001$ gives $M = 1218$ for $m = 3$, proving that

$$\mathbb{P}(\wp(G^*(n)) \leq \wp(K_{3,3}) < 0.585) > 0.999$$

for $n \geq 1218$. However, the data suggests that we should be able to do better. For n as small as 100, calculating the maximum generalised roundness of ten thousand such random graphs, we do not obtain any values exceeding 0.167. Indeed, in a sample of 10,000 random graphs on thirty vertices, all had an induced subgraph isomorphic to $K_{3,3}$, and none had maximum generalised roundness greater than 0.368.

Likewise, for $m = 4$, the bound above gives $M = 2877$, when in fact 10,000 random graphs on sixty vertices all had an copy of $K_{4,4}$, and had maximum generalised roundness not exceeding 0.217, much less than $\wp(K_{4,4}) \approx 0.415$. Testing for the presence of larger complete bipartite subgraphs quickly becomes computationally infeasible, but it is clear that at least for small m , this bound leaves room for improvement.

Thus it appears that our theorems do not fully describe the observed maximum generalised roundness values. While we have established that $\wp(G^*(n))$ converges almost surely to zero, the data suggests that we are some way short of tight bounds, perhaps because inequality (5.5.1) is not very tight or because the prevalence of complete bipartite graphs $K_{m,m}$ are not the strongest cause of the convergence.

In a similar way to Definition 4.7.2, for a sample of N graphs on n vertices, we denote by $m_{n,N}$, $M_{n,N}$, $\bar{\wp}_{n,N}$ and $s_{n,N}$ the sample minimum, maximum, mean and standard deviation of the maximum generalised roundness values. From the previous section, we expect all four statistics to approach zero in distribution as $n \rightarrow \infty$ for fixed N .

We calculated these statistics for a sample of $N = 1000$ graphs from the connected uniform model¹ for each $n \in \{10, 11, \dots, 99, 100\}$. A plot of the minimum, maximum and average in the sample can be seen in Figure 5.3.

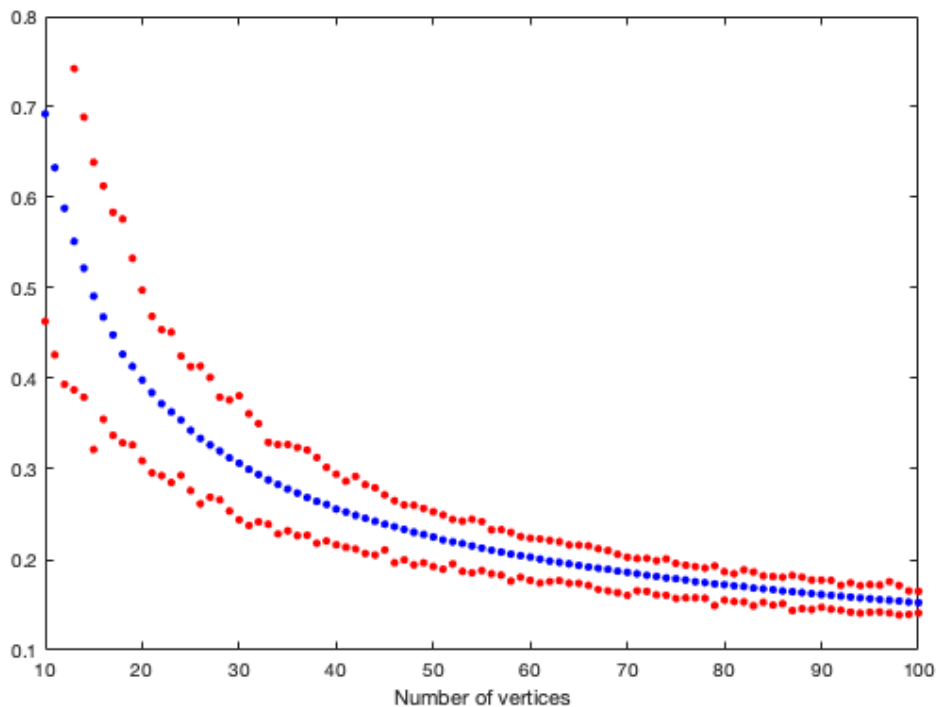


Figure 5.3: A plot of the sample mean, maximum and minimum of the maximum generalised roundness values of $N = 1000$ graphs on n vertices generated by the connected uniform model.

¹In order to guarantee that the mean was well-defined, we had to also exclude the complete graph. Fortunately, there is a vanishingly small chance that $G^*(n) = K_n$, so this has a negligible effect on our results.

Again we would like to study how these quantities decay with n and fit curves to these data points. For the mean, we used

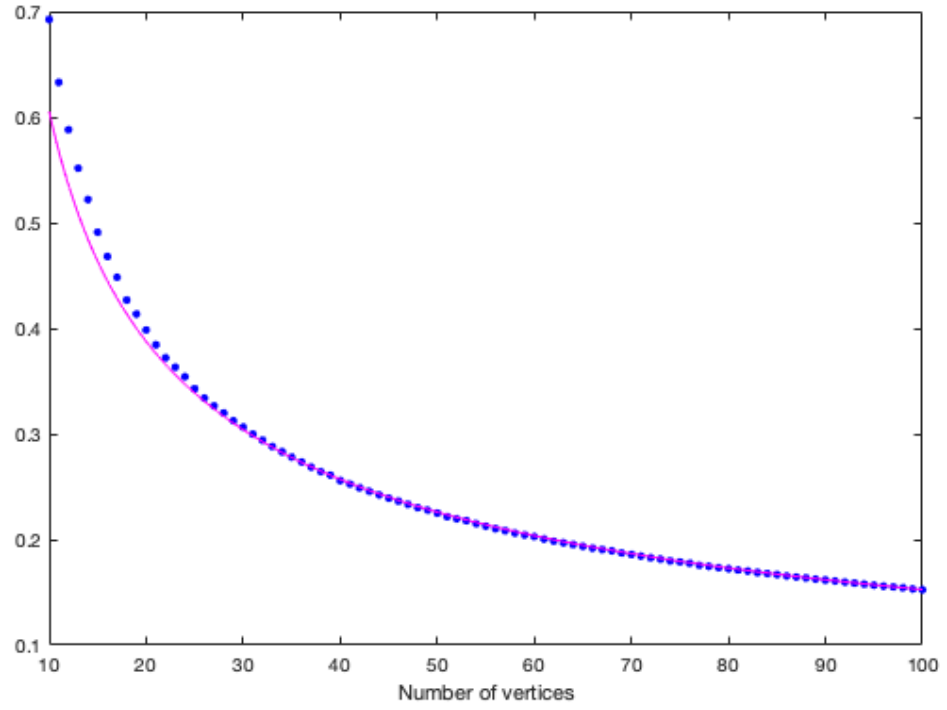
$$\bar{\varrho}_{n,1000} \approx \frac{1.9}{(n-2)^{0.55}}, \quad (5.5.2)$$

as shown in Figure [5.4a](#).

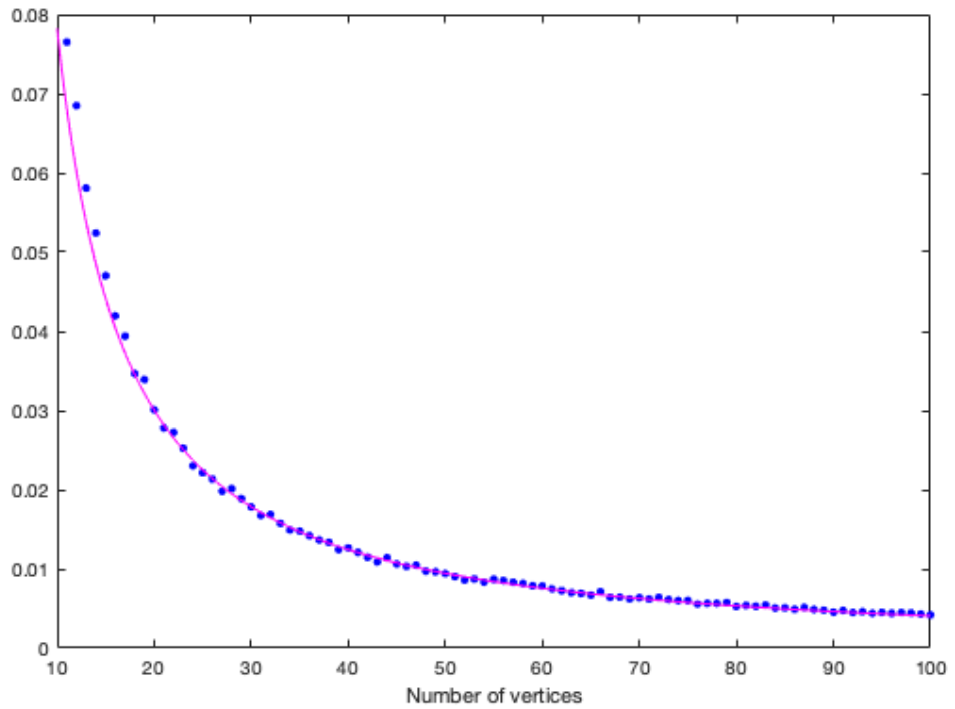
The standard deviation is quite well approximated by

$$s_{n,1000} \approx \frac{0.91}{(n-2)^{1.18}}, \quad (5.5.3)$$

as shown in Figure [5.4b](#).



(a) A plot of the sample mean $\bar{\varphi}_{n,1000}$, fitted with the curve (5.5.2).



(b) A plot of the sample standard deviation $s_{n,1000}$, fitted with the curve (5.5.3).

Figure 5.4: Curve fitting for the sample mean and standard deviation of one thousand connected graphs on n vertices, $10 \leq n \leq 100$.

The sample minimum and maximum are again slightly further dispersed from the mean than we would expect under the normal distribution. Figure 5.5 shows that $m_{n,1000}$ and $M_{n,1000}$ are often up to four standard deviations from $\bar{\varphi}_{n,1000}$. We therefore have the approximate bounds

$$L(n) = \frac{1.9}{(n-2)^{0.55}} - 4 \left(\frac{0.91}{(n-2)^{1.18}} \right)$$

and

$$U(n) = \frac{1.9}{(n-2)^{0.55}} + 4 \left(\frac{0.91}{(n-2)^{1.18}} \right),$$

so that ‘typically’

$$\mathbb{P}(\varphi(T(n)) < L(n)) < 0.001 \text{ and } \mathbb{P}(\varphi(T(n)) > U(n)) < 0.001.$$

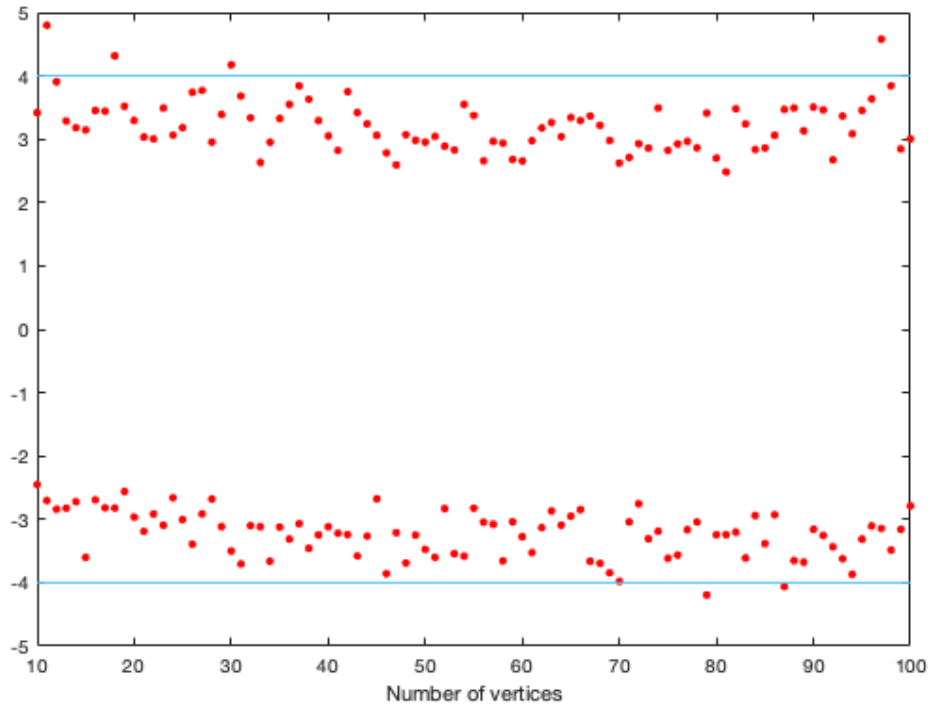


Figure 5.5: A plot of the z -scores of our observed values of $m_{n,1000}$ and $M_{n,1000}$, relative to $N(\bar{\varphi}_{n,1000}, s_{n,1000}^2)$.

5.6 p -negative type properties of atypical connected graphs

Our analysis thus far has focused on $G^*(n)$, which samples connected graphs on n vertices uniformly at random. However, it is important to identify that the structure of such a graph is not as random as it might initially appear. We have already found that $G^*(n)$ is typically very different from any tree on n vertices. On average, we expect $G^*(n)$ to have about half the possible edges, while a tree has only $n - 1$ edges, so a natural direction is to analyse the relationship between the number of edges of a graph and its maximum generalised roundness.

We begin by describing the number of edges of $G(n)$ more precisely, as an approximation to $G^*(n)$. Recalling from Definition [5.3.1](#) that each edge appears independently with probability $\frac{1}{2}$, we see that the number of edges $E_n = |E(G(n))|$ is binomially distributed, with

$$E_n \sim \text{Bin} \left(\binom{n}{2}, \frac{1}{2} \right),$$

so we have

$$\mathbb{E}(E_n) = \frac{1}{2} \binom{n}{2} = \frac{n(n-1)}{4}$$

and

$$\text{Var}(E_n) = \binom{n}{2} \frac{1}{2} \left(1 - \frac{1}{2} \right) = \frac{n(n-1)}{8},$$

so $\sigma(E_n) < n/3$. Thus for large n , the normal approximation to the binomial distribution suggests that well over 99% of our sample will be comprised of graphs where

$$\frac{n^2 - 5n}{4} < E_n < \frac{n^2 + 3n}{4}.$$

It should be noted however that not all such graphs have the property of Theorem [5.4.10](#), that is, some have relatively large maximum generalised roundness. One way to construct such a graph is to recall from Example [2.2.21](#) that complete graphs have infinite maximum generalised roundness, and accordingly create large complete subgraphs where possible.

Example 5.6.1. For n even, consider the graph L_n formed by connecting two copies of $K_{n/2}$ with a single edge. The example of L_{10} is depicted in [5.6](#).

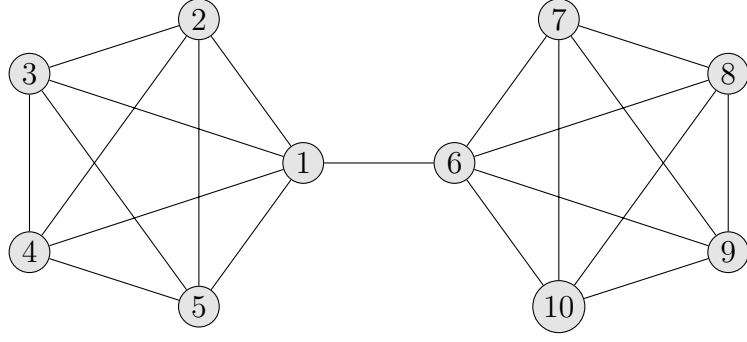


Figure 5.6: The graph L_{10} .

More precisely, let $n = 2m$, and let $L_n = (V, E)$ where $V = \{1, \dots, 2m\}$ and

$$E = \{\{i, j\} : 1 \leq i < j \leq m\} \cup \{\{i, j\} : m+1 \leq i < j \leq 2m\} \cup \{\{1, m+1\}\},$$

so the number of edges is

$$|E| = 2 \left(\frac{m(m-1)}{2} \right) + 1 = \frac{n^2 - 2n + 4}{4}.$$

The p -distance matrix is

$$D_p = \begin{pmatrix} 0 & 1 & \cdots & 1 & 1 & 2 & \cdots & 2 \\ 1 & 0 & \cdots & 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 & 2 & 3 & \cdots & 3 \\ 1 & 2 & \cdots & 2 & 0 & 1 & \cdots & 1 \\ 2 & 3 & \cdots & 3 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 2 & 3 & \cdots & 3 & 1 & 1 & \cdots & 0 \end{pmatrix}.$$

Calculating L_{2m} for several values of m , a clear pattern emerges where

$$\begin{aligned} \det D_p &= ((m-1) \times 4^p - (m-1) \times 3^p - 2(m-1) \times 2^p + (2m-3)) \\ &\quad \times ((m-1) \times 4^p - (m-1) \times 3^p + 2(m-1) \times 2^p + 1) \end{aligned}$$

and

$$\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = \frac{2(2(m-1) \times 2^p - (m-1) \times 3^p + 1)}{(m-1) \times 4^p - (m-1) \times 3^p + 2(m-1) \times 2^p + 1}.$$

In each case, the maximum generalised roundness is governed by $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ rather than $\det D_p$. Solving for the numerator to equal zero gives

$$2 \times 2^p - 3^p + \frac{1}{m-1} = 0,$$

and letting $m \rightarrow \infty$ we obtain

$$\lim_{m \rightarrow \infty} \wp(L_{2m}) = \log_{\frac{3}{2}} 2 \approx 1.709511.$$

We confirmed this result by calculating the values in Table 5.3 using our MATLAB program.

| m | $\wp(L_{2m})$ |
|-----|---------------|
| 10 | 1.7499 |
| 20 | 1.7290 |
| 30 | 1.7224 |
| 40 | 1.7191 |
| 50 | 1.7172 |
| 60 | 1.7159 |
| 70 | 1.7149 |
| 80 | 1.7143 |
| 90 | 1.7137 |
| 100 | 1.7133 |

Table 5.3: Values of $\wp(L_{2m})$ for various m , to four decimal places.

Thus L_n has a similar number of edges to the graphs of $G^*(n)$, but has much greater maximum generalised roundness. \square

Outliers notwithstanding, Theorem 5.4.10 describes the behaviour of graphs with about $n(n-1)/4$ edges. In order to describe graphs with significantly more or fewer edges, it is useful to define the density of a graph.

Definition 5.6.2. Let $G = (V, E)$ be a graph. Then the density of G is

$$\rho(G) = \frac{|E|}{\binom{|V|}{2}},$$

the ratio of the number of edges to the number of possible edges.

Remark 5.6.3. A graph with density close to 1 is said to be dense, whereas a graph with density close to 0 is said to be sparse.

The Erdős-Rényi random graph $G(n, q)$, and thus also the connected Erdős-Rényi random graph $G^*(n, q)$, typically has

$$E_{n,q} \approx \frac{n(n-1)}{2}q$$

edges, so we can simply change the value of q in order to produce graphs of density approximately q . Indeed, $\rho(\mathcal{G}^*(n, q)) \rightarrow q$ as a consequence of the law of large numbers.

We can then apply the full version of Theorem [5.4.9](#) to dense or sparse graphs to obtain a very similar theorem and proof to the $q = \frac{1}{2}$ case explored in Theorem [5.4.10](#).

Theorem 5.6.4 ([\[66\]](#) Theorem 2). *Given two non-decreasing sequences of positive integers $\mathbf{r} = (r_1, r_2, \dots)$ and $\mathbf{w} = (w_1, w_2, \dots)$, such that the sequence $\mathbf{r} + \mathbf{w}$ is increasing, define*

$$B_n(\mathbf{r}, \mathbf{w}) = \max \{r_k + w_k : \text{there is a copy of } K_{r_k, w_k} \text{ in } G(n, q)\}.$$

Suppose that $\frac{w_k}{r_k} \rightarrow c$ as $k \rightarrow \infty$, where $0 < c \leq 1$ is a constant, and denote

$$f = \max \left(\frac{1}{q}, \frac{1}{1-q} \right) \text{ and } h = \min \left(\frac{1}{q}, \frac{1}{1-q} \right).$$

Then for every $\epsilon > 0$,

$$\mathbb{P} \left(B_n(\mathbf{r}, \mathbf{w}) < (2 - \epsilon) \frac{\log n}{\log f} \right) = o(n^{-1-\delta})$$

and

$$\mathbb{P} \left(B_n(\mathbf{r}, \mathbf{w}) > (2 + \epsilon) \frac{\log n}{\log h} \right) = o(n^{-k}),$$

where $0 < \delta < 1$ is a constant and k is any integer.

We can now prove the main theorem of this section.

Theorem 5.6.5. *In the connected Erdős-Rényi model,*

$$\wp(G^*(n, q)) < \log_2 \left(1 + \frac{1}{(1 - \epsilon) \frac{\log n}{\log f} - 1} \right) \quad a.s.,$$

where $f = \max\left(\frac{1}{q}, \frac{1}{1-q}\right)$.

Proof. Following the structure of the proof of Theorem 5.4.10, we apply Theorem 5.6.4. We choose $\mathbf{r} = \mathbf{w} = (1, 2, 3, \dots)$ and define

$$m(n) = \frac{1}{2}B_n(\mathbf{r}, \mathbf{w}) = \max\{m : \text{there is a copy of } K_{m,m} \text{ in } G(n, q)\},$$

so that for every $\epsilon > 0$ we have

$$(1 - \epsilon)\frac{\log n}{\log f} < m(n) < (1 + \epsilon)\frac{\log n}{\log h} \quad a.s. \quad (5.6.1)$$

The lower bound now guarantees that $K_{m(n), m(n)}$ almost surely appears as an induced subgraph of $G(n, q)$, with diameter 2, so

$$\begin{aligned} \wp(G(n, q)) &\leq \wp(K_{m(n), m(n)}) \\ &= \log_2 \left(1 + \frac{1}{m(n) - 1}\right) \\ &< \log_2 \left(1 + \frac{1}{(1 - \epsilon)\frac{\log n}{\log f} - 1}\right) \quad a.s., \end{aligned}$$

and invoking Proposition 5.3.8 again relates $G(n, q)$ and $G^*(n, q)$ to complete the proof. \square

5.7 Empirical p -negative type properties of atypical connected graphs

Although $\wp(G^*(n, q)) \rightarrow 0$ in distribution for any fixed q , we still suspect that for given n , the distribution of $\wp(G^*(n, q))$ still depends significantly on the edge probability q . The sparsest connected graphs on n vertices are the trees, with maximum generalised roundness greater than one by Theorem 4.1.3. At the other extreme, the densest graph is K_n , with infinite maximum generalised roundness from Example 2.2.21. Even omitting one edge of K_n still leaves a graph with maximum generalised roundness

$$1 + \log_2 \left(1 + \frac{1}{n - 2}\right).$$

Halfway between these extremes, we recall from Theorem 5.4.4 that $\wp(G(n)) < 1$ almost surely. This suggests that both sparse and dense graphs have unusually large maximum generalised roundness as compared to those in between.

This conjecture can be experimentally verified and indeed expanded upon for particular values of n . For example, we computed $\wp(G^*(100, q))$ for q chosen uniformly at random between 0.05 and 0.95, and plotted the results against the number of edges of these graphs to obtain the results in Figure 5.7.

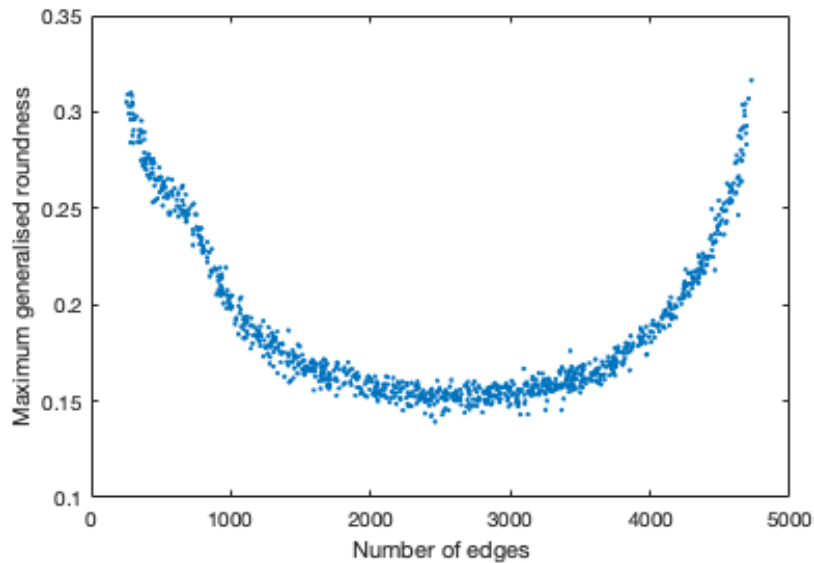


Figure 5.7: A plot of $\wp(G)$ against $|E(G)|$ for $G = G^*(100, q)$, where $q \in [0.05, 0.95]$ uniformly at random.

For very sparse ($\rho < 0.05$) and very dense ($\rho > 0.95$) graphs, we typically obtain maximum generalised roundness values well outside the range of this chart. The corresponding values of q have been omitted for practicality, as our graph generating algorithm has a low probability of yielding a connected graph when q is very small.

This visualisation shows that the maximum generalised roundness of a graph of $G^*(n, q)$ is strongly linked to its density, with the typical maximum generalised roundness smallest for graphs of density approximately one-half, and increasing as the graph becomes either more dense or more sparse. We can then infer the same result for q , with $\wp(G^*(n, q))$ generally smallest for $q \approx \frac{1}{2}$ and increasing as q moves away from $\frac{1}{2}$ in either direction. Recall however that the relationship between density and maximum generalised roundness is not without exceptions, as demonstrated by Example 5.6.1.

Note also that this relationship does not contradict the result of Theorem [5.6.5](#). While this theorem showed that $\wp(G^*(n, q))$ approaches zero with high probability for any q , the rate of convergence appears to depend not only on n but also on q .

CHAPTER 6

Future Work

This chapter outlines assorted partial results and demonstrates potential directions for further discoveries in this field.

6.1 Confluence of determinants for connected uniform random graphs

In our study of the connected uniform model of random graphs in Chapter 5, we often graphed the determinant $\det D_p$ for a sample graph $G^*(n)$. For n sufficiently large, we frequently observed the behaviour seen in Figure 6.1. Beginning at a common point $(0, (-1)^{n+1}n)$, the determinant would briefly increase in magnitude before decreasing towards its first zero, then remain very close to zero over an unexpectedly large interval.

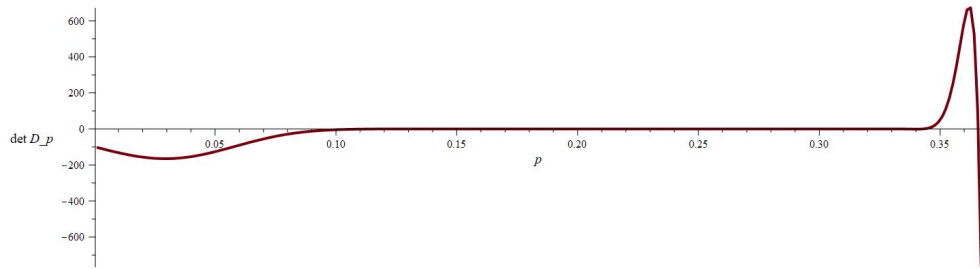


Figure 6.1: $\det D_p$ for the graph $G(100, 0.5, 0)$.

A surprising observation is that this behaviour seems to be almost universal in graphs of sufficient size. Figure 6.2 shows a plot of $\det D_p$ for one thousand graphs each on one hundred vertices, where each curve is evaluated numerically at points 0.0001 apart. We see a similar initial dip in all plotted curves, after which a ‘pinching’ phenomenon takes effect.

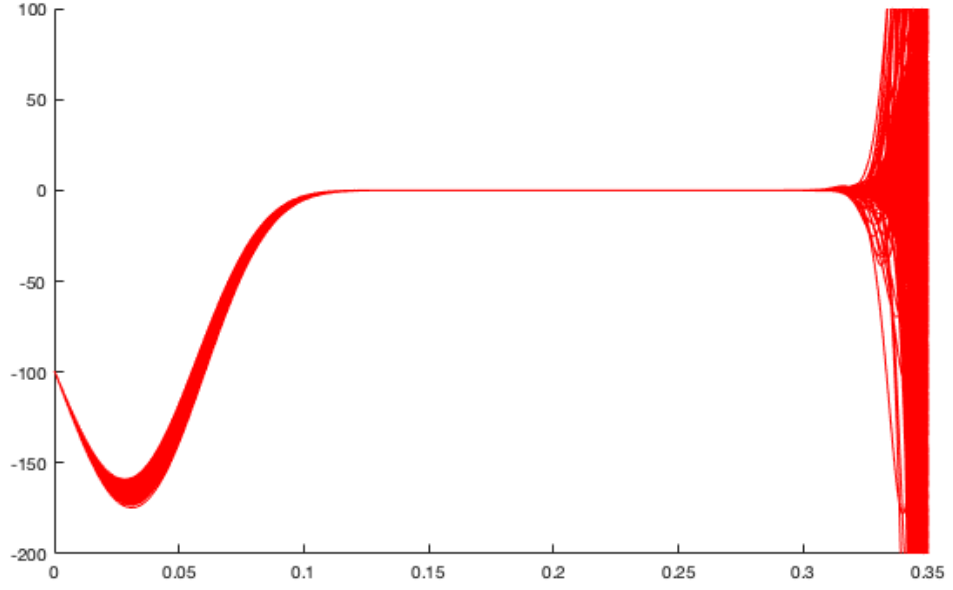


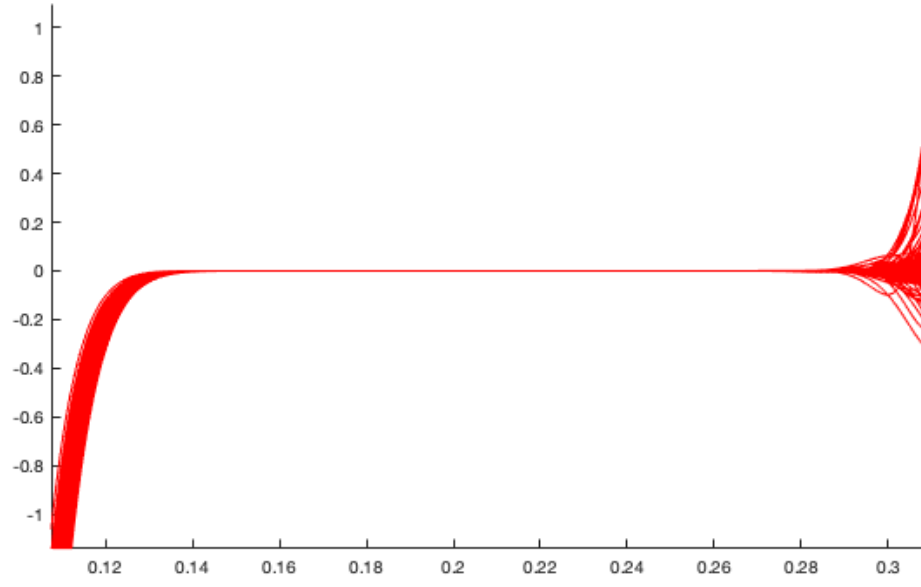
Figure 6.2: A plot of $\det D_p$ for $N = 1000$ graphs on $n = 100$ vertices generated by the connected uniform model.

A more detailed investigation in Figure 6.3 reveals that:

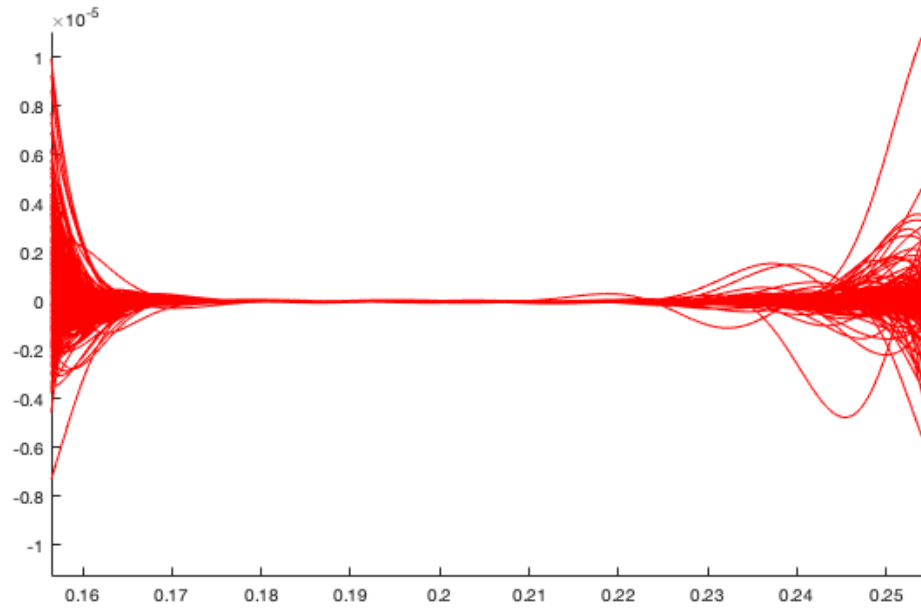
- $|\det D_p| < 1$ for $p \in [0.12, 0.3]$
- $|\det D_p| < 1 \times 10^{-5}$ for $p \in [0.16, 0.25]$
- $|\det D_p| < 1 \times 10^{-6}$ for $p \in [0.17, 0.23]$
- $|\det D_p| < 1 \times 10^{-7}$ for $p \in [0.18, 0.21]$.

This behaviour is particularly remarkable when we consider the algebraic expressions in question. Recall from Subsection 3.4.4 that the substitution $\lambda = 2^p$ transforms the function $\det D_p$ into a polynomial $q(\lambda)$ with integer coefficients on the order of 10^{80} . These coefficients are not consistent between the different graphs of our sample, exhibiting significant variety. It is surprising that any one of these polynomials remains so close to zero on a nontrivial interval, let alone all the polynomials we encountered in a random sample. Unfortunately we have not been able to provide a complete explanation for this phenomenon.

The corresponding graphs for $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ are less useful, as they have many discontinuities corresponding to the many zeroes of $\det D_p$.

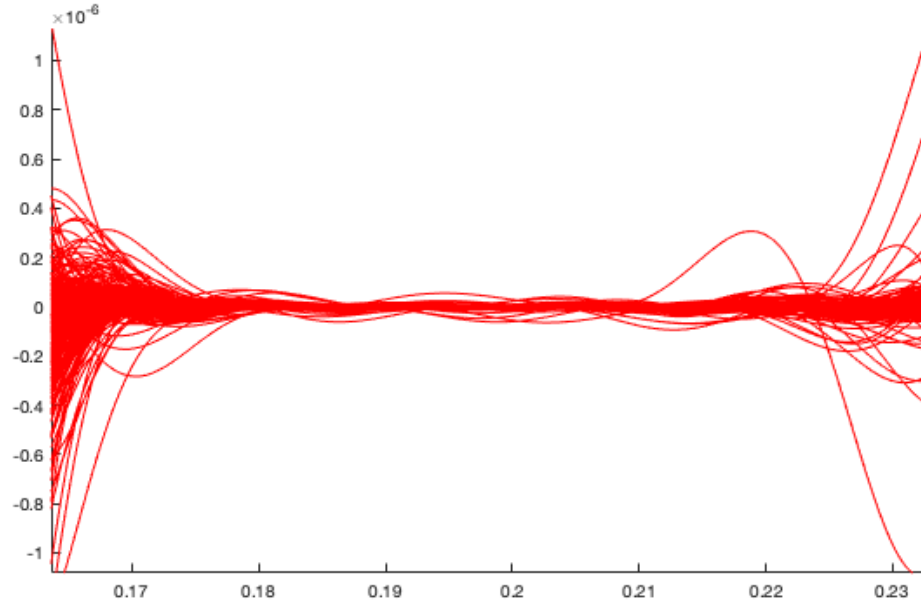


(a) $|\det D_p| < 1$

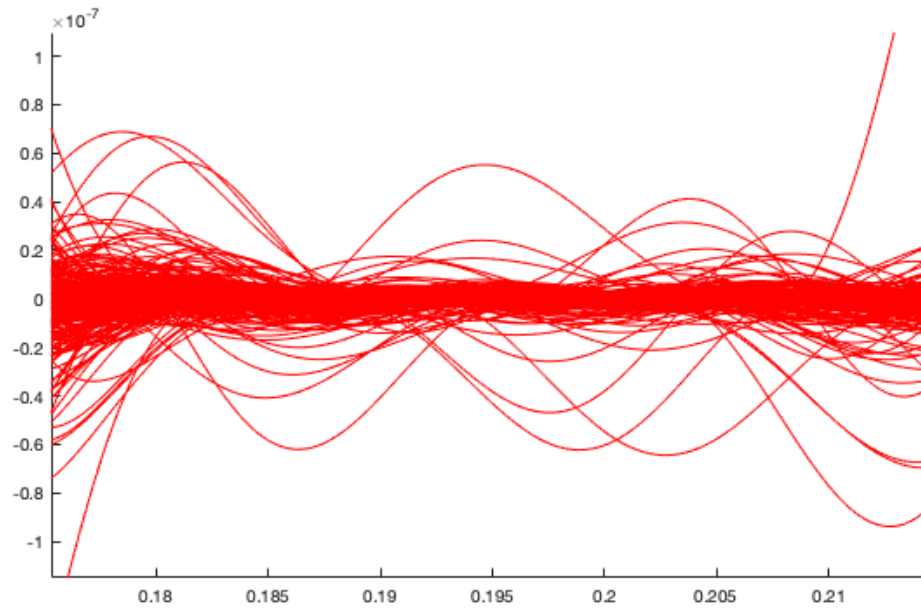


(b) $|\det D_p| < 1 \times 10^{-5}$

Figure 6.3: More detail of the ‘pinching’ effect from Figure [6.2](#).



(c) $|\det D_p| < 1 \times 10^{-6}$



(d) $|\det D_p| < 1 \times 10^{-7}$

Figure 6.3: More detail of the ‘pinching’ effect from Figure 6.2. (continued)

6.2 Weston tree conjecture

In a 2019 seminar at UNSW Sydney, Anthony Weston conjectured that trees can be distinguished (up to isomorphism) by their maximum generalised roundness, with the obvious exception of the path graphs, all of which have $\wp(T) = 2$. To test this claim, we obtained lists of trees (up to graph isomorphism) on $3 \leq n \leq 17$ vertices from the Combinatorial Object Server [63], which provides a web interface to various programs to generate combinatorial objects. In particular, trees are generated by the `nauty` program developed by McKay and Piperno [56]. Computation of the maximum generalised roundness of each tree using our MATLAB program has failed to reveal two non-trivial trees with equal maximum generalised roundness, so the problem remains open. However, one should note that this testing was very much ‘computationally dependent’. With almost fifty thousand trees to test for $n = 17$, we encountered several pairs of trees whose maximum generalised roundness agree to four or even five decimal places, so a more careful manual analysis was needed to confirm that these values were indeed distinct. This phenomenon will be intensified for larger n as a consequence of the pigeon-hole principle, perhaps requiring a more precise algorithm for the numerical evaluation of $\wp(G)$.

Notably, the conjecture does not extend to path metric graphs in general. The graphs in Figure 6.4 both have $\wp(G) = 1$, and similarly $\wp(G) = \log_2 \left(\frac{3+\sqrt{5}}{2} \right)$ for both graphs depicted in Figure 6.5.

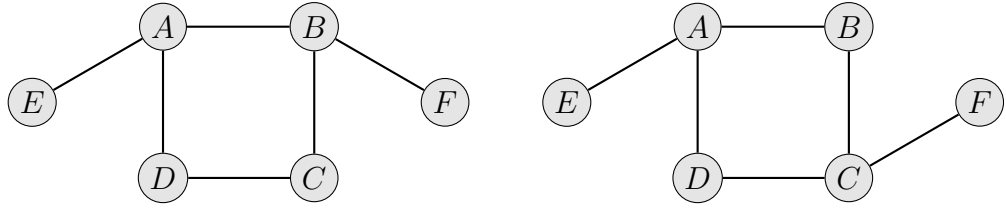


Figure 6.4: Two non-isomorphic six-vertex graphs with equal maximum generalised roundness.

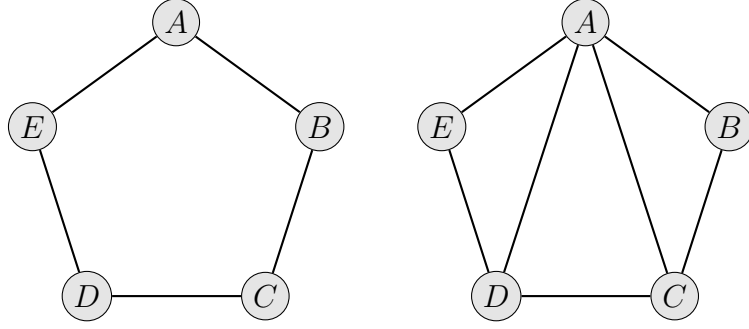


Figure 6.5: Two non-isomorphic five-vertex graphs with equal maximum generalised roundness.

6.3 Sánchez formula types

Recall Sánchez's formula for the maximum generalised roundness of a finite metric space, from Corollary [3.1.5](#):

$$\wp(X, d) = \min\{p \geq 0 : \det D_p = 0 \text{ or } \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = 0\},$$

where D_p is the p -distance matrix and $\mathbb{1}$ is a vector of ones of the appropriate dimension. We aim to classify finite metric spaces by the criterion which determines this minimum.

Definition 6.3.1. A finite metric space (X, d) is of Type I if

$$\wp(X, d) = \min\{p \geq 0 : \det D_p = 0\},$$

or Type II if

$$\wp(X, d) = \min\{p \geq 0 : \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = 0\}.$$

Remark 6.3.2. Note that these two classes are mutually exclusive unless $\wp(X, d)$ is infinite, as D_p is not invertible when $\det D_p = 0$.

We then wish to describe these two classes of metric spaces in terms of their graph theoretic properties. The only existing result in the literature is due to Murugan [\[62\]](#), concerning vertex transitive graphs.

Definition 6.3.3. Let $G = (V, E)$ be a (possibly infinite) graph. G is vertex transitive if for any $v, w \in V$ there is an edge-preserving bijection $f : V \rightarrow V$ such that $f(v) = w$.

Murugan listed several examples of vertex transitive graphs, including the cycle graph C_n , the complete graph K_n , the complete bipartite graph $K_{n,n}$, as well as the Peterson graph, Platonic solids, Hamming cubes and Cayley graphs of groups.

Theorem 6.3.4 ([62, Theorem 3.1]). *A finite vertex transitive graph must be of Type I.*

Proof. A consequence of Definition 6.3.3 is that for a finite vertex transitive graph, each row of the distance matrix is a permutation of the first row.

It follows that $\mathbb{1}$ is an eigenvector of D_p . Let the corresponding eigenvalue be λ , and note that $\lambda = \sum_{j=1}^n d(x_1, x_j)^p > 0$. Now we have $D_p \mathbb{1} = \lambda \mathbb{1}$, and multiplying on the left by D_p^{-1} , if it exists, yields $D_p^{-1} \mathbb{1} = \frac{1}{\lambda} \mathbb{1}$. Then $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = \frac{n}{\lambda} \neq 0$ for any p , and the result follows. \square

A surprising pattern emerged in our study of trees, as calculations showed that the maximum generalised roundness was always determined by the sum of entries of the inverse.

Conjecture 6.3.5. *All path metric trees are of Type II.*

An analogous method of proof to that used for vertex transitive graphs seems unlikely, as the situation is more complicated here. For path metric graphs, $\det D_p$ can have zeros, but it appears that $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ always has a smaller zero.

Example 6.3.6. Let T be the tree on six vertices in Figure 6.6.

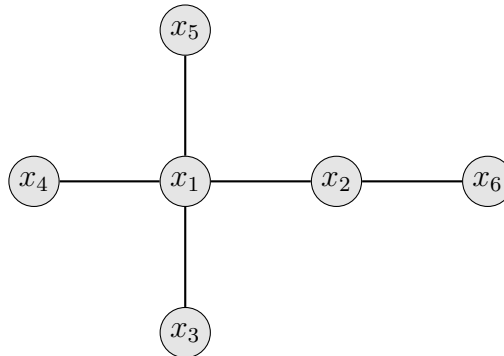


Figure 6.6: A tree on six vertices.

Maple calculates the determinant of the p -distance matrix

$$\det D_p = 4^p (3 \times 16^p - 6 \times 12^p + 3 \times 9^p - 2 \times 4^p - 6 \times 3^p + 3),$$

which can be plotted as in Figure 6.7.

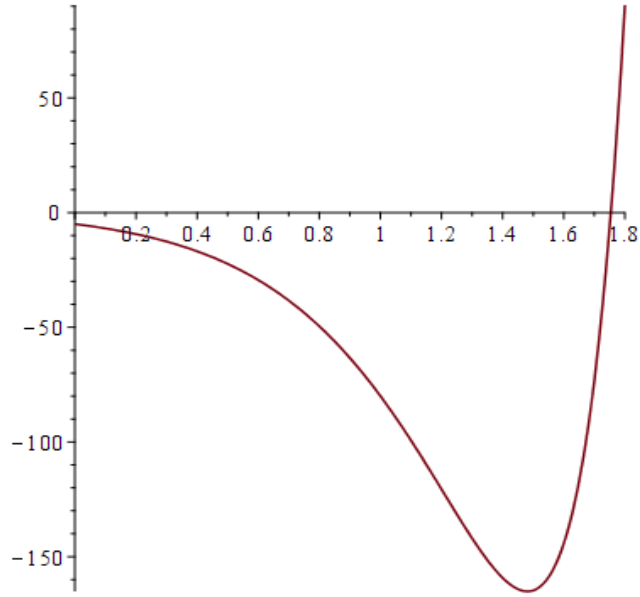


Figure 6.7: The determinant of the p -distance matrix of T from Example [6.3.6](#).

Applying `fsolve` on $(1.7, 1.8)$ finds a zero at $p \approx 1.7540$. However, $\langle D_p^{-1} \mathbf{1}, \mathbf{1} \rangle$ is zero at $p \approx 1.4080$, so T is of Type II. \square

The situation for random graphs is quite different, as seen in Figure [6.8](#). Taking one thousand connected graphs on n vertices chosen uniformly at random, we find that the proportion of Type II graphs increases over the first few values of n , reaching a peak around 92%, before decreasing steadily. However, even for $n = 100$, more than 10% are of Type II.

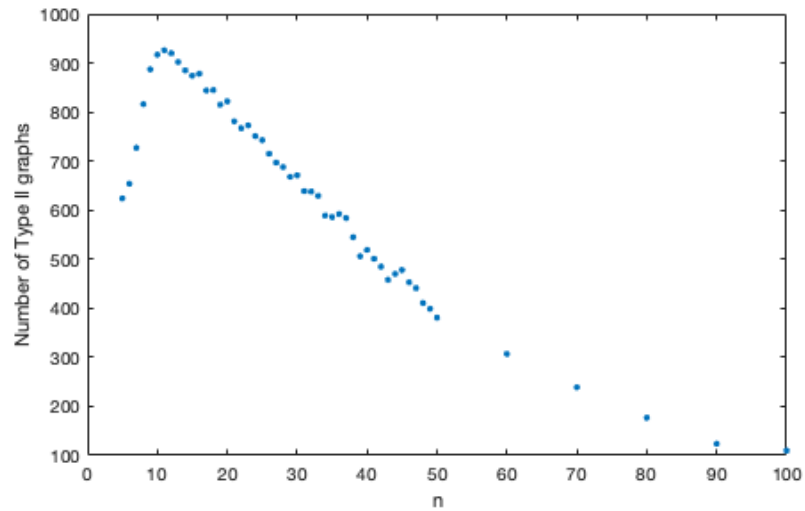
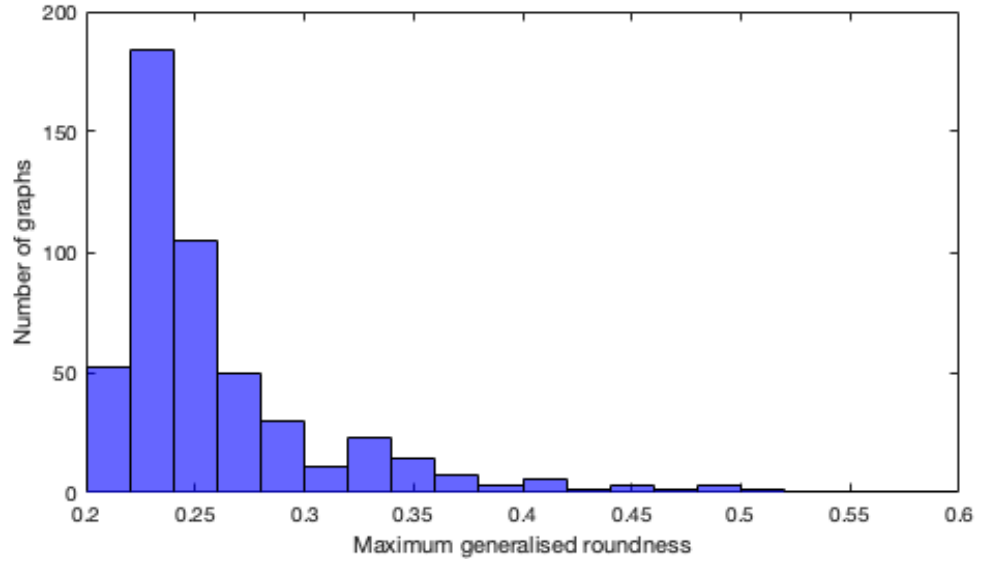
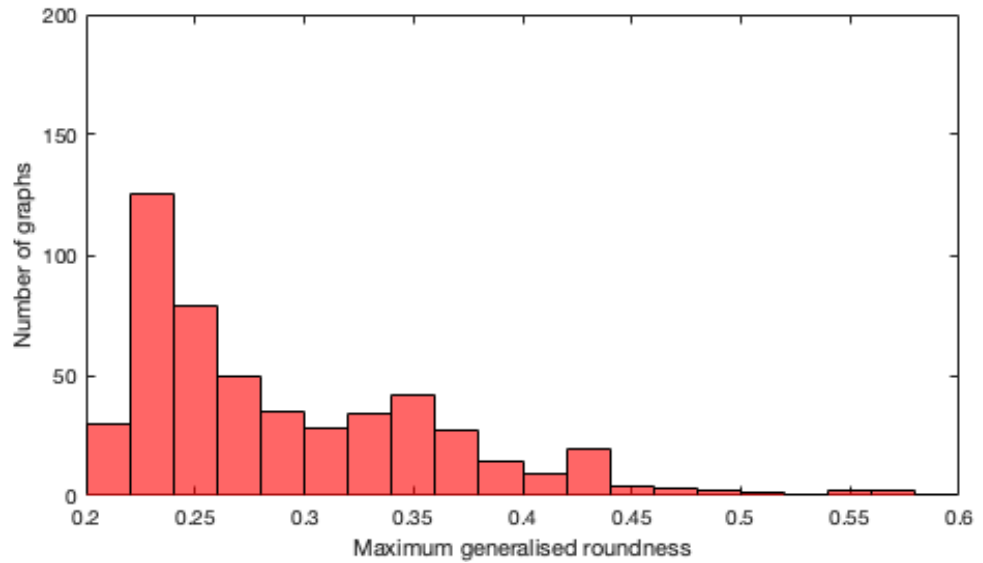


Figure 6.8: The number of Type II graphs among one thousand connected uniform random graphs on n vertices.

In fact, there appears to be some correlation between the maximum generalised roundness of a graph and its type. Our data indicates that graphs of high maximum generalised roundness are disproportionately of Type II, as seen in Figure 6.9



(a) Type I



(b) Type II

Figure 6.9: Histograms of maximum generalised roundness values among a total of one thousand graphs from the connected Erdős-Rényi model $G^*(50, q)$, where $q \in (0.05, 0.95)$ uniformly at random.

6.4 Infinite metric trees

Infinite trees have also been studied in some detail, beginning with the work of Doust and Weston [28] who showed that infinite metric trees also have strict 1-negative type, but may have maximum generalised roundness of 1. Caffarelli, Doust and Weston [16] then showcased classes of spherically symmetric trees, k -regular trees, infinitely bifurcating trees and comb graphs, all of which have maximum generalised roundness one. A well known open question in the field asks whether there are any infinite metric trees with maximum generalised roundness strictly between 1 and 2.

One direction of investigation was to extend the existing work on comb graphs, which are a particular type of metric tree.

Definition 6.4.1. Let S be a subset of the positive integers. Define the comb graph $C_S = (V, E)$, where

$$V = \{(k, 0) : k \in \mathbb{Z}^+\} \cup \{(k, 1) : k \in S\}$$

and

$$E = \{((k, 0), (k + 1, 0)) : k \in \mathbb{Z}^+\} \cup \{((k, 0), (k, 1)) : k \in S\},$$

with the path metric. The edges $((k, 0), (k, 1))$ are known as teeth.

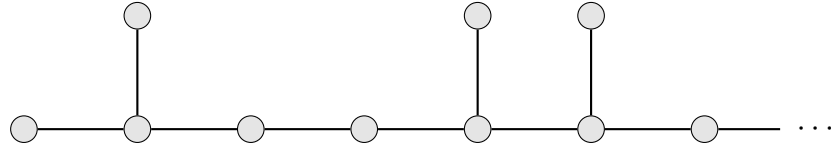


Figure 6.10: An example of a comb graph.

Theorem 6.4.2 ([16] Corollary 4.5). *Let S be a nonempty subset of \mathbb{Z}^+ . If the distance between any two consecutive teeth is uniformly bounded by some $r > 0$, then $\wp(C_S) = 1$.*

It is unknown whether the same result applies if this condition is relaxed in any way.

We conjectured that the forked path graph Y shown in Figure 6.11 has maximum generalised roundness strictly greater than one and less than two, for suitably chosen edge weights w_i . In particular, we suspected that an appropriate sequence of edge weights must be rapidly increasing to infinity, in order to provide the starkest contrast to the condition studied in Theorem 6.4.2.

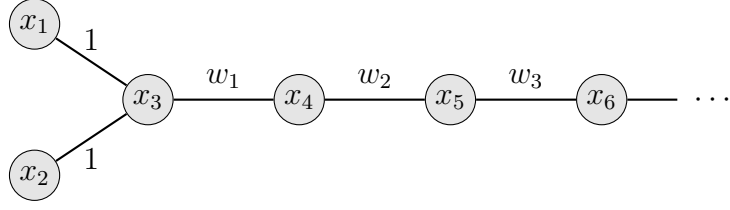


Figure 6.11: The forked path graph.

We aimed to first study the maximum generalised roundness of the truncations Y_n as in Figure 6.12.

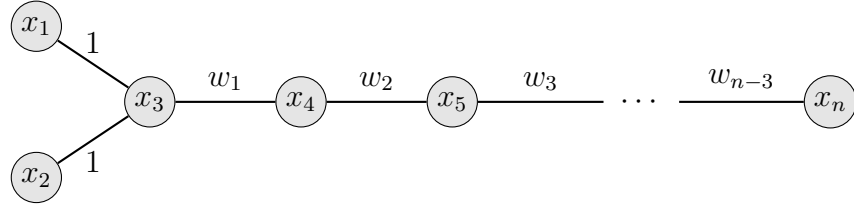


Figure 6.12: A truncation of the forked path graph.

Our method involved analysing the related graphs Z_n in which each vertex x_i ($i \geq 4$) is joined by an edge of weight w_{i-3} not only to x_{i-1} , but to each x_j for $j < i$. The example of Z_5 is depicted in Figure 6.13.

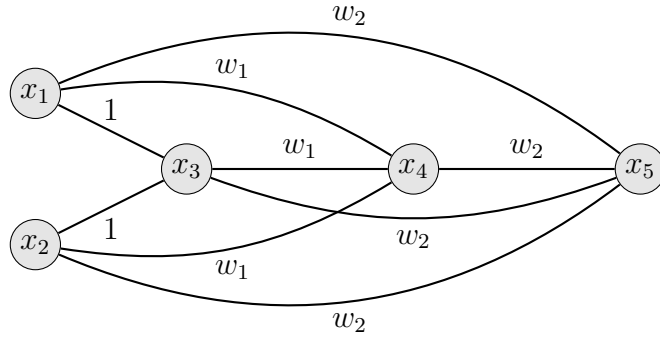


Figure 6.13: A truncation of the forked path graph with added edges.

Recall from Corollary 3.1.5 that the maximum generalised roundness of a finite metric space can be evaluated using the p -distance matrix of the space. The graph in Figure 6.13 has p -distance matrix as follows.

$$\begin{aligned}
D_p^{(n)} &= \begin{pmatrix} 0 & 2 & 1 & w_1 & w_2 & \cdots & w_{n-3} \\ 2 & 0 & 1 & w_1 & w_2 & \cdots & w_{n-3} \\ 1 & 1 & 0 & w_1 & w_2 & \cdots & w_{n-3} \\ w_1 & w_1 & w_1 & 0 & w_2 & \cdots & w_{n-3} \\ w_2 & w_2 & w_2 & w_2 & 0 & \cdots & w_{n-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ w_{n-3} & w_{n-3} & w_{n-3} & w_{n-3} & w_{n-3} & \cdots & 0 \end{pmatrix} \\
&= \begin{pmatrix} D_p^{(n-1)} & w_{n-3} \mathbb{1} \\ w_{n-3} \mathbb{1}^T & 0 \end{pmatrix},
\end{aligned}$$

We see that additional vertices contribute a row and column whose entries are all w_{n-3} , away from the main diagonal. This structure allows for easier computation of $\wp(Z_n)$, using the identity

$$\det \left(\begin{array}{ccc|c} \ddots & & & \vdots \\ & A & & B \\ & & \ddots & \vdots \\ \hline \cdots & C & \cdots & d \end{array} \right) = (d - CA^{-1}B) \det A,$$

where the right hand side is zero when A is not invertible. Thus we have

$$\begin{aligned}
\det(D_p^{(n)}) &= \left(-w_{n-3} \mathbb{1}^T (D_p^{(n-1)})^{-1} w_{n-3} \mathbb{1} \right) \det D_p^{(n-1)} \\
&= -w_{n-3}^2 \left\langle (D_p^{(n-1)})^{-1} \mathbb{1}, \mathbb{1} \right\rangle \det D_p^{(n-1)}.
\end{aligned}$$

Note that the right hand side has its first zero at $\wp(Z_{n-1})$, as discussed in a 2019 seminar by Gavin Robertson at UNSW Sydney.

Next, we remove the artificial edges to recover Y_n . This perturbs the distance matrix somewhat. For example, Y_5 has distance matrix

$$D_p^{(5)} = \begin{pmatrix} 0 & 2 & 1 & w_1 + 1 & w_2 + w_1 + 1 \\ 2 & 0 & 1 & w_1 + 1 & w_2 + w_1 + 1 \\ 1 & 1 & 0 & w_1 & w_2 + w_1 \\ w_1 + 1 & w_1 + 1 & w_1 & 0 & w_2 \\ w_2 + w_1 + 1 & w_2 + w_1 + 1 & w_2 + w_1 & w_2 & 0 \end{pmatrix}.$$

Similarly, for larger n , we typically replace w_{n-3} by $w_{n-3} + w_{n-4} + \dots$. If the sequence of edge weights w_i is rapidly increasing, for example $w_i = 10^{10^i}$, then each distance changes by a relatively small fraction. We conjectured that a small perturbation of the distance matrix would result in only a small change in the maximum generalised roundness, that is, the maximum generalised roundness is a continuous function of the distances between points. Since $\det D_p$ is a smooth function, the only point of contention relates to turning points at zeros.

Recall the example of C_5 , the cycle graph on five vertices, whose determinant is plotted in Figure 6.14.

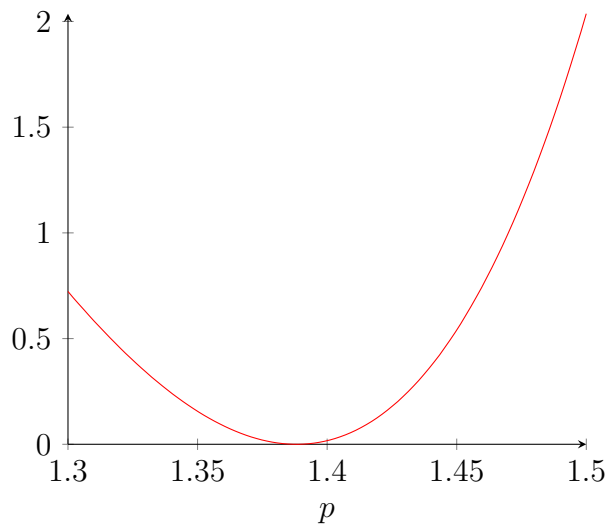


Figure 6.14: A plot of $\det D_p$ for the cycle graph on five vertices.

It is plausible that a perturbation of the edge weights in the underlying graph could shift the turning point upwards, so

$$\varphi = \min\{p \geq 0 : \det D_p = 0 \text{ or } \langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = 0\}$$

would change abruptly. This quantity would then fail to be continuous in the distances comprising D . As it happens, in the example of C_5 , no perturbation has this property. More surprisingly, we were not able to find any graph which serves as a counterexample of this form. For instance, we identified five graphs (up to isomorphism) on six vertices where the maximum generalised roundness is achieved at a maximum turning point of $\det D_p$, pictured in Figure 6.15.

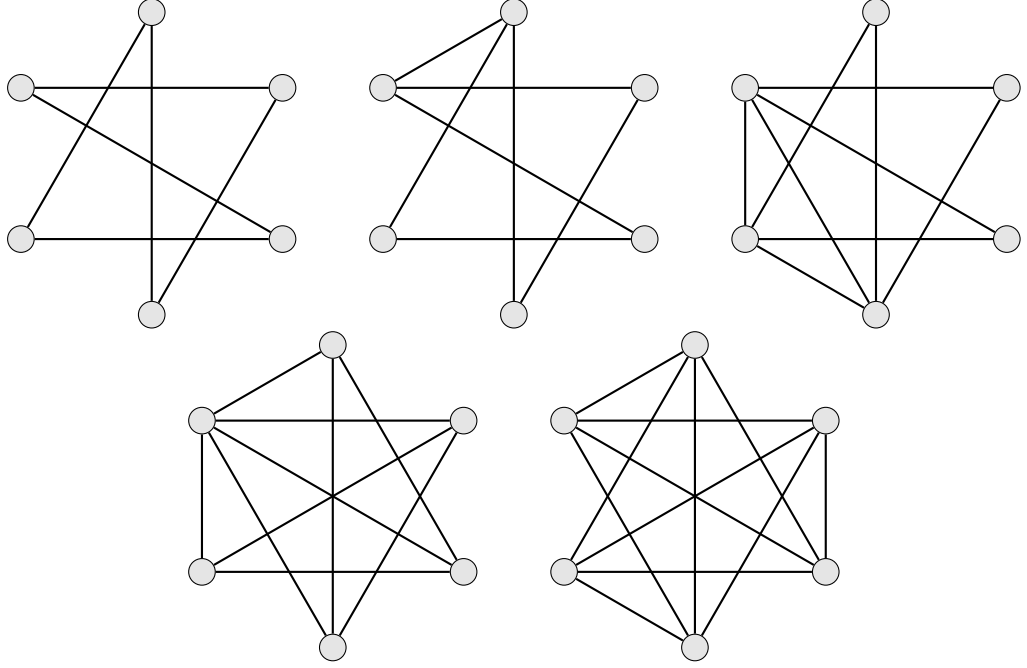


Figure 6.15: All connected graphs on six unlabelled vertices where the maximum generalised roundness occurs at a maximum turning point of $\det D_p$.

Small modifications to edge weights caused one of three effects:

- shifted the turning point above the axis;
- shifted the turning point along the axis, or
- caused $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$ to have a zero smaller than those of $\det D_p$, close to the original maximum generalised roundness.

None of these cases nor any larger examples provided a counterexample to the continuity of φ , so the question of continuity remains open, as does the viability of modelling the forked path Y by the modified truncations Z_n .

6.5 Planar graphs

We examined the maximum generalised roundness of planar graphs using techniques similar to those employed in Chapters [4](#) and [5](#).

Definition 6.5.1. A graph $G = (V, E)$ is planar if it can be embedded (not necessarily isometrically) into \mathbb{R}^2 . In such an embedding, each vertex v is represented by a point in the plane and each edge (v, w) by a polyline between the corresponding points containing no other vertex and no point of any other edge.

This definition omits some technical details. For a full exposition of the subject, the reader may consult Diestel [25, Chapter 4]. Note that we will again consider only the path metric on these planar graphs.

Remark 6.5.2. It is important not to confuse the graph theoretic property of planarity with the metric space property of embedding in \mathbb{R}^2 . In fact, it is often the case that a planar graph endowed with the path metric cannot be embedded isometrically into the plane.

The following fundamental result in the study of planar graphs follows from Euler's formula.

Theorem 6.5.3. *A planar graph on n vertices has at most $3n - 6$ edges.*

A proof is provided by Diestel [25, Corollary 4.2.10].

There are two particularly famous small graphs which are not planar.

Example 6.5.4. Consider the complete graph K_5 . This graph has five vertices and ten edges, violating Theorem 6.5.3, so it is not planar. \square

Example 6.5.5. The complete bipartite graph $K_{3,3}$ has six vertices and nine edges, so it does not immediately contradict Theorem 6.5.3. However, for planar graphs without a cycle of length three, we can show that the number of edges cannot exceed $2n - 4$ using the same method of proof as Diestel above. Thus $K_{3,3}$ is also not planar. \square

As a result, any graph where K_5 or $K_{3,3}$ appears as a subgraph¹ cannot be planar. We therefore do not find large complete bipartite subgraphs as in Chapter 5. Furthermore, we can deduce from Theorem 6.5.3 that the density of a planar graph, as defined in Definition 5.6.2, is at most

$$\frac{6(n-2)}{n(n-1)} < \frac{6}{n},$$

so large planar graphs are very sparse. As a consequence of our discussion in Subsection 5.7, we expect large planar graphs to generally have large maximum generalised roundness compared to random graphs on the same number of vertices, but it is initially unclear just how large these values must be.

At first, we considered whether planar graphs have the same property as trees, that is, $\wp(G) \geq 1$. However, counterexamples are easily found. Proposition 3.6.2

¹Indeed, Kuratowski's theorem famously states that a graph is planar if and only if it has no subgraph that is a subdivision of K_5 or $K_{3,3}$.

gives $\wp(K_{2,3}) = \log_2(12/7) < 1$. More generally, larger complete bipartite graphs $K_{2,n-2}$ are always planar, with a straightforward (non-isometric) embedding in the plane shown in Figure 6.16

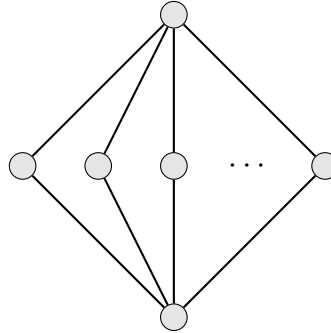


Figure 6.16: The complete bipartite graph $K_{2,n-2}$

We hypothesised that this graph has the smallest maximum generalised roundness among all planar graphs on n vertices with the path metric.

Conjecture 6.5.6. *Let G be a planar graph on n vertices, endowed with the path metric. Then*

$$\wp(G) \geq \wp(K_{2,n-2}) = \log_2 \frac{4n-8}{3n-8}.$$

We tested this conjecture by obtaining a list of planar graphs (up to graph isomorphism) from the House of Graphs [13], and calculating the maximum generalised roundness for each graph using the MATLAB program developed in Chapter 3. We found that the conjecture holds for $n \leq 9$, but fails for $n = 10$. Note that there are over ten million graphs to test for $n = 10$, so it was vital to use a very efficient algorithm and implementation. Indeed, it would have been very difficult to disprove this conjecture without such a program!

There are two ten-vertex graphs (up to isomorphism) for which

$$\wp(G) < \wp(K_{2,8}) = \log_2(32/22) \approx 0.5406.$$

One is formed by deleting an edge of $K_{2,8}$, giving $\wp(G) = 0.5398$. The other is depicted in Figure 6.17, with $\wp(G) \approx 0.5385$. The second counterexample is more instructive, as it suggests that we can construct planar graphs of small maximum generalised roundness by attaching several smaller complete bipartite graphs $K_{2,m}$.

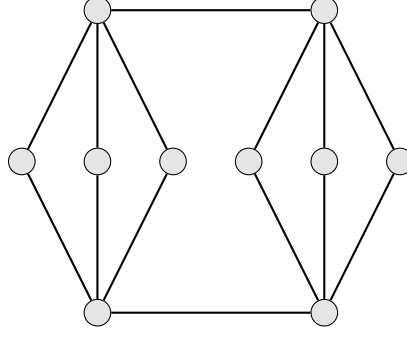


Figure 6.17: A planar graph on ten vertices

The following result would be a consequence of Corollary 6.5.6, but it too is false.

Conjecture 6.5.7. *Let G be a finite planar graph, endowed with the path metric. Then*

$$\wp(G) > \log_2 \frac{4}{3} \approx 0.4150.$$

We present a counterexample similar to that in Figure 6.17. By constructing two copies of $K_{2,7}$ with two links, we obtain a graph with $\wp(G) \approx 0.4103$. Indeed, letting G_m be the graph formed by joining two copies of $K_{2,m}$ in this way, Maple calculates for $m = 3, 4, \dots, 8$ that²

$$\begin{aligned} \det D_p &= (2^{(2m-2)p}) (2^p - 1 + 3^p)(2^p + 1 - 3^p) \\ &\quad \times ((m+1) \times 4^p + (2m-1) \times 6^p - m \times 9^p - (3m+1) \times 2^p - m \times 3^p + 2m) \\ &\quad \times ((m+1) \times 4^p - (2m-1) \times 6^p - m \times 9^p + (3m+1) \times 2^p - m \times 3^p + 2m) \end{aligned}$$

and

$$\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle = \frac{2((2m+4) \times 2^p - 3m \times 3^p + 3m)}{(m+1) \times 4^p - (2m-1) \times 6^p - m \times 9^p + (3m+1) \times 2^p - m \times 3^p + 2m}.$$

In each of these cases, G_m is of Type II, that is, $\wp(G_m)$ is the smallest zero of $\langle D_p^{-1} \mathbb{1}, \mathbb{1} \rangle$. The symbolic calculation of these quantities is prohibitively slow for $m \geq 9$, but we are confident that the pattern continues. After dividing through by m , the numerator is

$$\left(1 + \frac{2}{m}\right) 2^p - 3 \times 3^p + 3,$$

²Our attempts to prove this analytically for general n have thus far been fruitless. The primary difficulty is that unlike a complete bipartite graph, the p -distance matrix of G_m contains terms in both 2^p and 3^p , greatly complicating the required calculations.

and letting $m \rightarrow \infty$ we must solve

$$2^p - 3 \times 3^p + 3 = 0.$$

Solving numerically, we have

$$\lim_{m \rightarrow \infty} \wp(G_m) \approx 0.3159.$$

This is consistent with results obtained using our MATLAB program for larger values of m , as in Table [6.1](#).

| m | $\wp(G_m)$ |
|-----|------------|
| 10 | 0.3818 |
| 20 | 0.3488 |
| 30 | 0.3378 |
| 40 | 0.3323 |
| 50 | 0.3290 |
| 60 | 0.3269 |
| 70 | 0.3253 |
| 80 | 0.3241 |
| 90 | 0.3232 |
| 100 | 0.3225 |

Table 6.1: Values of $\wp(G_m)$ for various m , to four decimal places.

It remains to be seen whether we can use more sophisticated constructions to produce planar graphs with even smaller maximum generalised roundness.

Finally, a probabilistic result similar to those in Chapters [4](#) and [5](#) is another direction for further investigation. Again, the most straightforward approach is to consider probabilities with respect to the uniform measure on the connected planar graphs on n vertices. There are finitely many such graphs, but there is relatively little known about random connected planar graphs in the literature.

Indeed, the enumeration of these graphs is itself quite difficult. Letting g_n be the number of connected planar graphs on n labelled vertices, no explicit formula for g_n is known. The asymptotic behaviour is better understood, with progress made

using enumerative methods [55] and generating functions [39]. The best available result, due to Giménez and Noy [39], states that

$$g_n \sim gn^{-7/2}\gamma^n n!$$

where $g \approx 0.41043 \times 10^{-5}$ and $\gamma \approx 27.22688$. They then prove that for a graph chosen uniformly at random from this collection, the number of edges is asymptotically normal, with mean $\mu_n \sim \kappa n$ and variance $\sigma_n^2 \sim \lambda n$, where $\kappa \approx 2.21326$ and $\lambda \approx 0.43034$, and that the number of ‘appearances’ of a given subgraph is also normally distributed. These results may prove useful for a proof similar to those used in Chapters 4 and 5.

To make the necessary hypotheses for such a proof, we would like to first generate connected planar graphs uniformly at random, in order to compute their maximum generalised roundness using the algorithm developed in Chapter 3. Again, methods for generating these graphs are limited. The **plantri** algorithm [15] uses triangulations to produce planar graphs up to isomorphism on up to 64 vertices at great speed, often more than one million per second. The related algorithms **fullgen** [15] and **buckygen** [14] produce fullerenes, a more restricted class of planar graphs, on up to several hundred vertices. To generate larger connected planar graphs uniformly at random, one may consider the algorithm of Bodirsky, Gröpl and Kang [12], which runs in $O(n^3)$ time after $\tilde{O}(n^6) = O(n^6 p(\log n))$ preprocessing, where p is a polynomial. A recent algorithm by Griffith [45] generates connected planar graphs on up to 4489 vertices, providing another direction for further investigation.

APPENDIX A

Maple programs

The following Maple code implements the algorithm developed in Chapter [3](#).

```
with(GraphTheory) :
with(LinearAlgebra) :
with(RandomGraphs) :

#####
# mgr
#####

mgr:=proc(Dm)
# Calculate the maximal generalized roundness of a metric graph space
# with distance matrix Dm via Sanchez's formula

    local nvert, DM, p1, p2, p3, DMinv, F1, F2, t, dt, delp, debug :
    local minp, maxp, npts, eps1, eps2 :

    # Set parameters

    debug:=false :
    # search in the interval [minp, maxp]
    minp:=0.05123 : # Needs to be smaller for very large graphs
    maxp:=2.1 : # it is only bigger than this for complete graphs
    eps1:=0.2 : # how small is considered zero?
    eps2:=0.02 : # how small is considered zero?
    delp:=0.0001 : # amount to move endpoint for second check
    nvert:=RowDimension(Dm):
    npts:=5*nvert : # Number of points to check
    t:=time():

    # Find p1, the first p for which det(D_p) = 0

    DM:=p-> map( x -> evalf(x^p) , Dm) :
    F1:=p->Determinant(DM(p)):
```

```

p1:=findroot(F1, minp, maxp, npts, eps1) :

# Deal with the case that no root is found

if not(type(p1, float)) then
  p1:=infinity :
end if:
dt:=time()-t :
if debug then print("p1 is", p1, dt): end if:
t:=time():

# Now try to see if the sum of entries of the inverse are zero
# in [minp, p1-delp)

try
  p2:=findrootSE(Dm, minp, min(p1, 2.1)-delp, npts, eps2) :

# Hope to never hit this. It means we missed a root of det Dp

catch "singular" :
  print ("Singular: ", DM(1)) :
  p2:=2 :
end try:

# Deal with case of no solution

if not(type(p2, float)) then
  p2:=infinity :
end if:
dt:=time()-t :
if debug then print("p2 is", p2, dt) : end if:
if p1 <= p2 then
  return([evalf(p1, 5) , 0]) :
else
  return([evalf(p2, 5) , 1]) :
end if:
end proc:

#####
# findroot
#####

findroot:=proc(f, lb, ub, npts, eps)
# Find the first root of f between lb and ub
# npts is the number of points to test for a change of sign
# eps is a tolerance: if |f(x)| < eps then we have a root
# dirl and dirr are the signs of slopes over two consecutive intervals

```

```

# if these change sign then we have a turning point

local xl, xr, xm, yl, yr, ym, i, k :
local stepsize, dirl, dirr, t, dt, debug, bits :
debug:=false :
Digits:=15 :
bits:=15 : # number of iterations of bisection method
stepsize:=evalf((ub-lb)/npts):
xl:=evalf(lb) :
yl:=evalf(f(lb)) :
if yl = 0 then
    return xl :
end if:
dirl:=0 :
t:=time():

# Loop through the interval by stepsize

for i from 1 to npts do
    xr:=evalf(xl+stepsize) :
    yr:=evalf(f(xr)) :
    if debug then print(xl, yl, xr, yr) : end if:
    dirr:=signum(yr-yl) :

    # Check for the function changing sign

    if yl*yr < 0 then
        t:=time():

        # Sign change found. Now use bisection to find root

        for k from 1 to bits do
            xm:=evalf((xl+xr)/2.0):
            ym:=evalf(f(xm)) :
            if debug then print(xl, yl, xm, ym, xr, yr) : end if:
            if ym = 0 then
                return(xm) :
            end if:
            if yl*ym < 0 then
                xr:=xm :
                yr:=ym :
            else
                xl:=xm :
                yl:=ym :
            end if:
        end do:
        return(xm) :
    end if:
end for:

```

```

# Next check to see whether we had a turning point

elif dirl*dirr < 0 then
  # the function had a turning point !
  # better look to see whether it gets close to zero
  # first check that it is not the wrong way up
  # Only worth checking if the middle point is closer to zero
  # than the end ones
  if debug then print("turning point!", xl, yl, xr, yr) : end if:
  if abs(yl) < abs(yr) then

    # Send off to find where the turning point is

    xm:=findminB(f, xl-stepsizes, xr, 10) :
    ym:=evalf(f(xm)) :

    # Decide that we have found a root if the function is close
    # enough to zero at the (approximate) value of the turning point

    if abs(ym) < eps then
      if debug then print("Used findmin, returning", xm) : end if:
      return(xm) :
    end if:
  end if:
end if:

# Shift values to iterate to next interval.

xl:=xr :
yl:=yr :
dirl:=dirr :
t:=time() :
end do:

# If we get here, we failed to find a root

return(infinity) :
end proc:

#####
# findrootSE
#####

findrootSE:=proc(Dm, lb, ub, npts, eps)
# This procedure finds the first value of p at which the sum of the
# entries of  $Dp^{(-1)}$  is zero.

```

```

# npts is the number of points to test for a change of sign
# eps is a tolerance: if |f(x)| < eps then we have a root

local xl, xr, xm, yl, yr, ym, i, k, stepsize, dirl, dirr, t, dt :
local nvert, Dp, Dpinv, SE, C, substepsize, ii, debug, bits :
debug:=false :
Digits:=15 :
bits:=10 : # number of iterations of the bisection method
nvert:=RowDimension(Dm) :
Dp:=p->map( x -> evalf(xp) , Dm) :
Dpinv:=p->MatrixInverse(subs(q = evalf(p) , Dp(q))) :
SE:=p->add(add(Dpinv(p))[i, j] , i = 1..nvert) , j = 1..nvert) :
stepsize:=evalf((ub-lb)/npts):
xl:=evalf(lb) :
C:=Dpinv(xl) :
yl:=add(add(C[i, j] , i = 1..nvert) , j = 1..nvert) :
if yl = 0 then
    return xl :
end if:
dirl:=0 :
t:=time() :

# Loop through the interval by stepsize

for i from 1 to npts do
    xr:=evalf(xl+stepsize) :
    C:=Dpinv(xr) :
    yr:=add(add(C[i, j] , i = 1..nvert) , j = 1..nvert) :
    if debug then print("step", i, xl, yl, xr, yr) : end if:
    dirr:=signum(yr-yl) :

    # First check for a change in sign

    if yl*yr < 0 then

        # Sign change found. Do bisection to approximate root

        t:=time() :
        for k from 1 to bits do
            xm:= (xl+xr)/2.0:
            C:=Dpinv(xm) :
            ym:=add(add( C[i, j] , i = 1..nvert) , j = 1..nvert) :
            if debug then print("bisection", xl, yl, xm, ym, xr, yr) :
            end if:
            if ym= 0 then
                return(xm) :
            end if:

```



```

    if yl*ym < 0 then
        xr:=xm :
        yr:=ym :
    else
        xl:=xm :
        yl:=ym :
    end if:
end do:
return(xm) :

# Now check for a turning point

elif dirl*dirr < 0 then
    # the function appears to have a turning point !
    # better look to see whether it gets close to zero
    # first check that it is not the wrong way up
    if debug then print("turning point!", xl, yl, xr, yr) : end if:
    t:=time() :
    # First rule out turning points for which |f| does not have
    # a local min
    if abs(yl) < abs(yr) then

        # do a more careful check that it is not catching a singularity!
        # Check 20 points near here

        substepsize:= stepsize/20.0:
        xl:=xl-stepsize :
        C:=Dpinv(xl) :
        yl:=add(add(C[i, j] , i = 1..nvert) , j = 1..nvert) :

        # use the dirl from before

        for ii from -9 to 10 do
            xr:=xl+substepsize :
            C:=Dpinv(xr) :
            yr:=add(add(C[i, j] , i = 1..nvert) , j = 1..nvert) :
            dirr:=signum(yr-yl) :

            # if the y value is small enough, this is close enough!
            if abs(yr) < eps then
                if debug then print("close enough", xr, yr) : end if:
                return(xr) :

            # Also check whether there was actually a zero before
            # the turning point

            elif yl*yr < 0 then

```

```

        return(findrootSE(Dm, xr, xl, 4, eps) :

# Now check for a turning point on this finer scale
# if so, check more to find it

elif dirl*dirr < 0 then
    xm:=findminB(SE, xl-substepsize, xr, 10) :
    ym:=evalf(SE(xm)) :
    if abs(ym) < eps then
        # declare that we have found a root if it is close enough!
        return(xm) :
    end if:
end if:
end do:
end if:
end if:
xl:=xr :
yl:=yr :
dirl:=dirr :
t:=time() :
end do:

# If we get here, we didn't find a root

return(infinity) :
end proc:

#####
# FindminB
#####

findminB:=proc(f, lb, ub, steps)
# Find the turning point of a function between
# lb and ub.
# It is assumed that |f| has a local min in (lb,ub)
# But check that there isn't a root in here too
# nits is the number of iterations to do.

local stepsize, xmin, ymin, imin, i, t, debug, xl, xr, yl, yr, fgrid :
debug:=false :
if debug then print("Entering findmin", lb, ub, steps) : end if:
t:=time() :
Digits:=20 :
stepsize:=evalf((ub-lb)/steps):
xl:=lb :
yl:=evalf(f(xl)) :
ymin:=abs(yl) : # ymin stores the smallest value so far

```

```

imin:=0 :

# Loop through first on a somewhat coarse scale

for i from 1 to steps do
  xr:=xl+stepsize :
  yr:=evalf(f(xr)) :

  # Has the function changed sign near the turning point?
  # If so, find that root

  if yl*yr < 0 then
    return(findroot(f, xl, xr, 10, 0.01)) :
  else
    if debug then print("still seaching in minB", i, xr, yr, ymin) :
    end if:

    # Is this the closest y value yet?

    if abs(yr) < ymin then
      imin:=i :
      ymin:=abs(yr) :
    end if:
  end if:
  xl:=xr :
  yl:=yr :
end do:

# Now we know about where the min is, decrease the interval and
# make the grid finer

xmin:=lb+imin*stepsize :
xl:=lb + (imin-1)*stepsize :
yl:=f(xl) :
fgrid:=20 :
stepsize:= stepsize/fgrid:
for i from -(fgrid-1) to (fgrid-1) do
  xr:=xl+stepsize :
  yr:=evalf(f(xr)) :

  # Again...make sure there isn't a root there!
  if yl*yr < 0 then
    return(findroot(f, xl, xr, 10, 0.01)) :
  else
    if debug then
      print("still seaching in minB, level 2", i, xr, yr, ymin) :
    end if:
  end if:
end do:

```

```

        if abs(yr) < ymin then
            imin:=i :
            xmin:=xr :
            ymin:=abs(yr) :
        end if:
    end if:
    xl:=xr :
    yl:=yr :
end do:
if debug then print("Leaving minB", imin, xmin, ymin) : end if:
return(xmin) :
end proc:

```

APPENDIX B

MATLAB programs

The following MATLAB code implements the algorithm developed in Chapter [3](#).

```
function [p, t] = mgr(D,eps,tree)
% MGR Maximum generalised roundness.
% MGR(D,eps,tree) finds (numerically) the maximum
% generalised roundness and type of the metric space
% with distance matrix D, using steps of eps, with the
% optional flag tree for efficiency.
% MGR(D,eps) uses the default value tree = 0.
% MGR(D) uses the default value eps = 0.0003.

% for trees, begin searching at p = 1
if nargin < 3
    tree = 0;
end

if nargin == 1
    eps = 0.0003;
end

n = size(D,1);

% zero tolerance
tol = 10^(-n/10);

detDp = @(p) det(D.^p);
p1 = myfzero(detDp,tree,2.5,eps,tol);

suminv = @(p) sum(sum(inv(D.^p)));
p2 = myfzero(suminv,tree,p1,eps,tol);

if p2 < p1
    p = p2;
    t = 2;
```

```

elseif p1 < 2.5
    p = p1;
    t = 1;
else
    p = 2.5;
    t = 0;
end

```

```

function p = myfzero(f,pmin,pmax,eps,tol)
% MYFZERO Finds the smallest zero of f in [pmin,pmax].
% MYFZERO(f,pmin,pmax,eps,tol) finds (numerically) the
% smallest zero of the function f on the interval [pmin,
% pmax], using steps of eps and zero tolerance tol. The
% function returns pmax if no zero is found.
% MYFZERO(f,pmin,pmax,eps) uses the default value tol =
% 1e-3.
% MYFZERO(f,pmin,pmax) uses the default value eps =
% 0.03.

if nargin < 5
    tol = 1e-3;
end

if nargin < 4
    eps = 0.0003;
end

p = -1;
v1 = f(pmin);
df1 = 0;
for s = pmin:eps:pmax
    v2 = f(s+eps);
    df2 = v2-v1;
    if v1*v2 < 0
        p = fzero(f,[s s+eps]);
        return;
    elseif df1*df2 < 0
        x = s-eps;
        if v1 > 0 && df1 < 0
            x = fminbnd(f,s-eps,s+eps);
            y = f(x);
            if y < -tol
                p = fzero(f,[s-eps x]);
                break
            elseif y < tol
                p = x;
                break
            end
        end
    end
end

```

```

        end
    elseif v1 < 0 && df1 > 0
        x = fminbnd(@(x)-f(x),s-eps,s+eps);
        y = f(x);
        if y > tol
            p = fzero(f,[s-eps x]);
            break
        elseif y > -tol
            p = x;
            break
        end
    end
end
v1 = v2;
df1 = df2;
end

if p < 0
    p = pmax;
end

```

APPENDIX C

Catalogue of trees

The following tables list the maximum generalised roundness values of all unweighted trees on $4 \leq n \leq 9$ vertices, as confirmed by both the MATLAB and Maple programs.

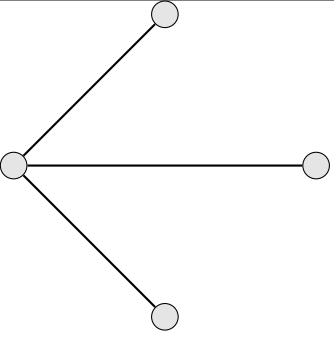
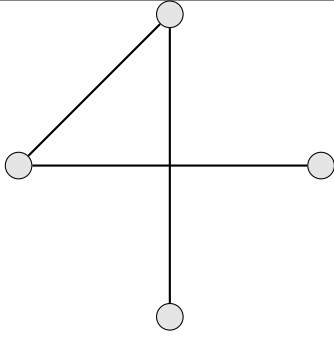
| | |
|--|---|
|  |  |
| 1.5850 | 2.0000 |

Table C.1: All trees on four unlabelled vertices and their maximum generalised roundness values.

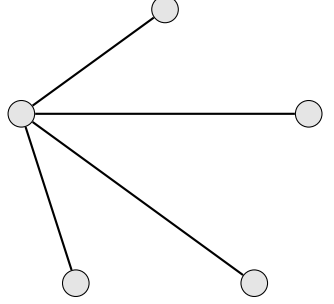
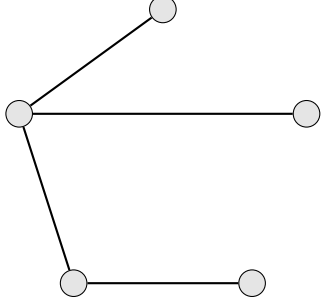
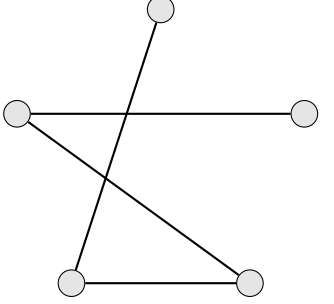
| | | |
|---|--|---|
|  |  |  |
| 1.4150 | 1.5761 | 2.000 |

Table C.2: All trees on five unlabelled vertices and their maximum generalised roundness values.

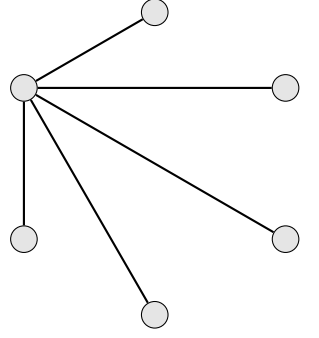
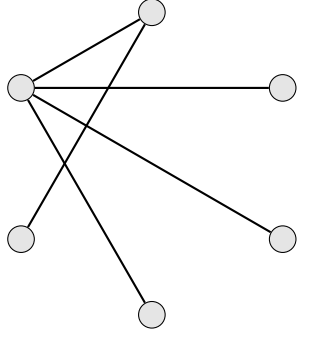
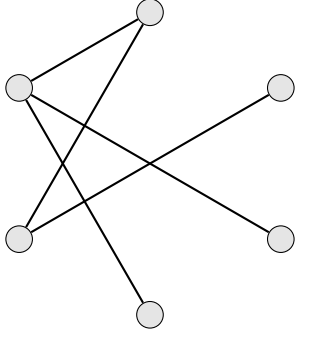
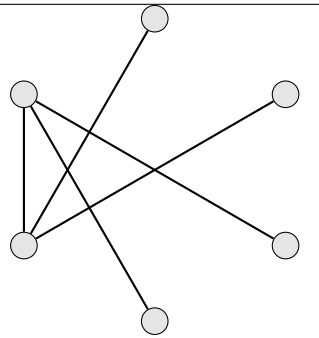
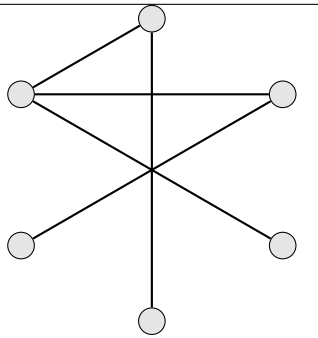
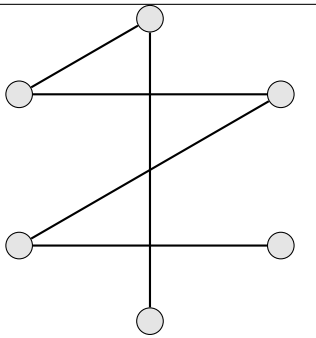
| | | |
|---|--|---|
|  |  |  |
| 1.3219 | 1.4080 | 1.5713 |
|  |  |  |
| 1.5166 | 1.5658 | 2.0000 |

Table C.3: All trees on six unlabelled vertices and their maximum generalised roundness values.

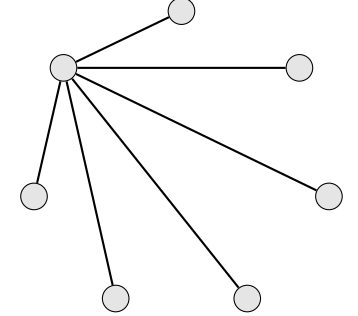
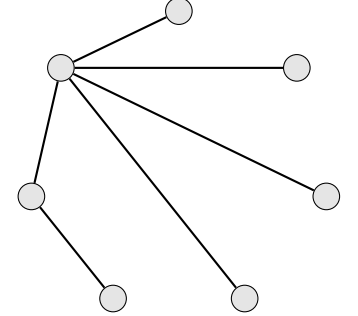
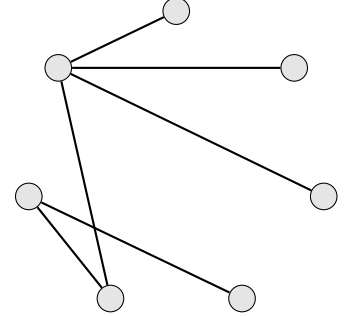
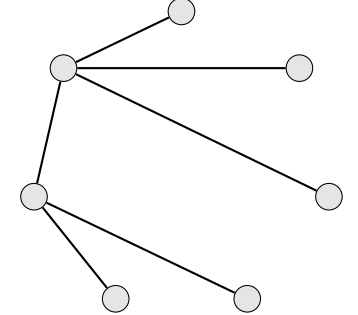
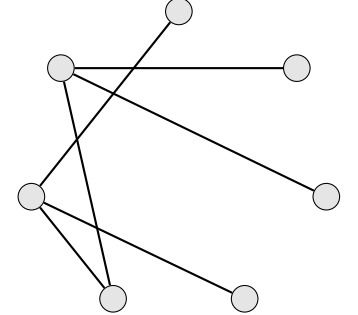
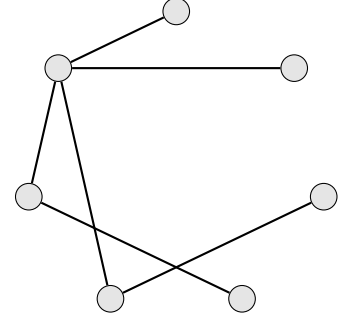
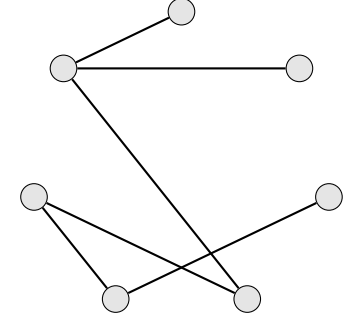
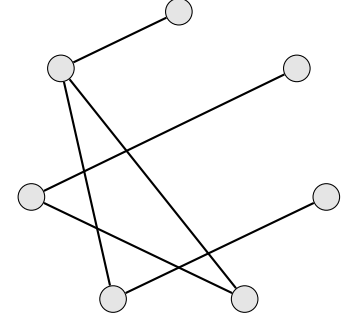
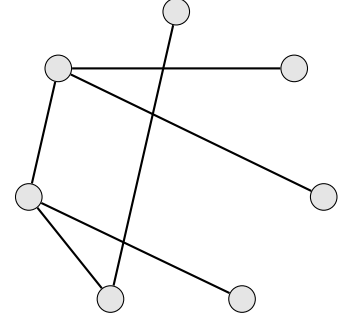
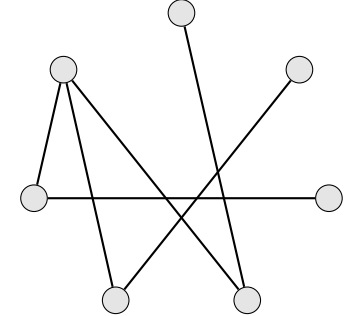
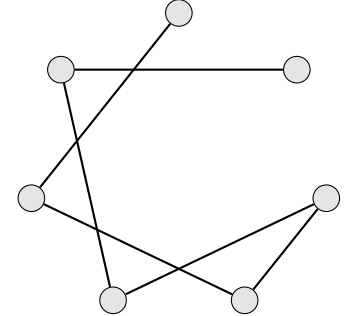
| | | |
|---|--|---|
|  |  |  |
| 1.2603 | 1.3167 | 1.4046 |
|  |  |  |
| 1.3906 | 1.5271 | 1.4004 |
|  |  |  |
| 1.5687 | 1.5603 | 1.5058 |
|  |  | |
| 1.5540 | 2.0000 | |

Table C.4: All trees on seven unlabelled vertices and their maximum generalised roundness values.

| | | |
|--------|--------|--------|
| | | |
| 1.2224 | 1.2590 | 1.3142 |
| | | |
| 1.3073 | 1.3954 | 1.3468 |
| | | |
| 1.3111 | 1.3854 | 1.4027 |
| | | |
| 1.3967 | 1.3824 | 1.5359 |

Table C.5: All trees on eight unlabelled vertices and their maximum generalised roundness values.

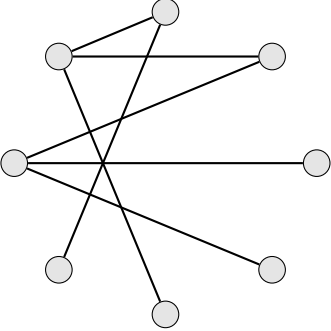
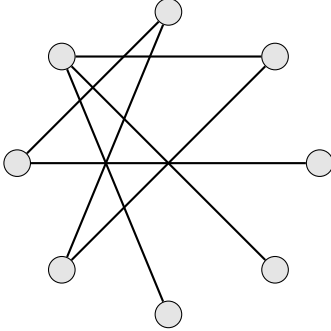
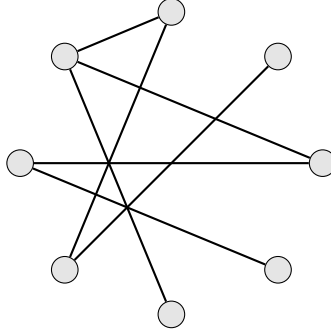
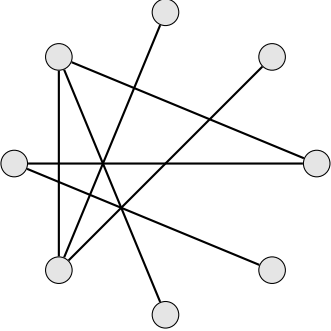
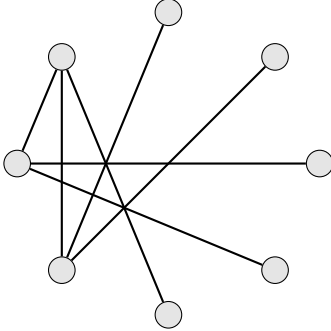
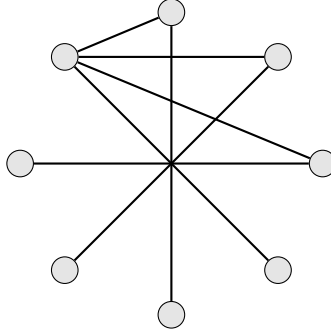
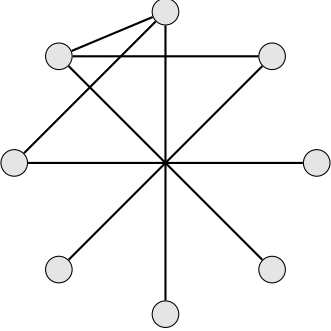
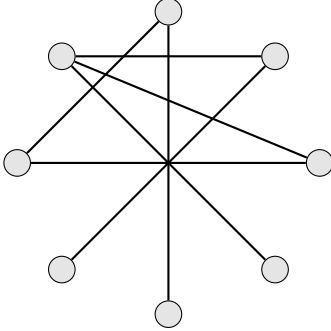
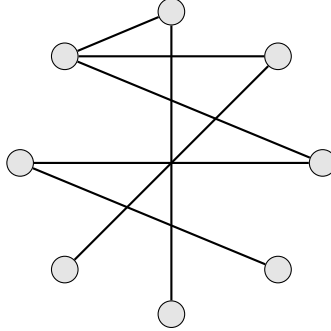
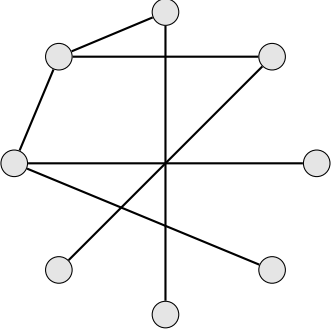
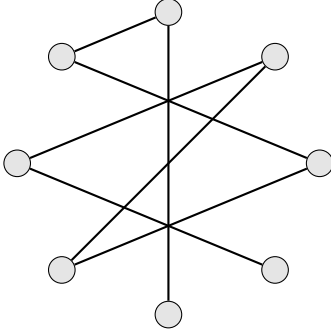
| | | |
|---|--|---|
|  |  |  |
| 1.5176 | 1.5671 | 1.5544 |
|  |  |  |
| 1.5003 | 1.4647 | 1.3922 |
|  |  |  |
| 1.4946 | 1.5574 | 1.5477 |
|  |  | |
| 1.4938 | 2.0000 | |

Table C.5: All trees on eight unlabelled vertices and their maximum generalised roundness values (continued).

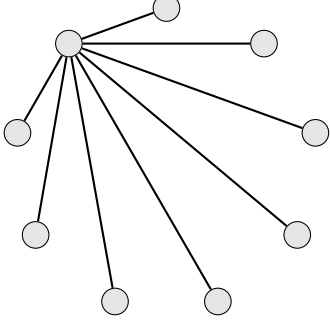
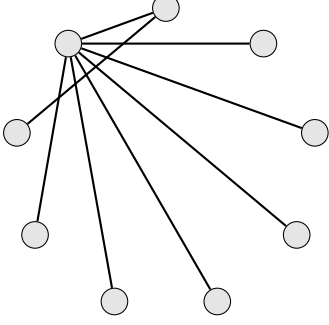
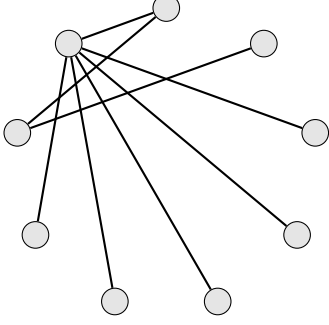
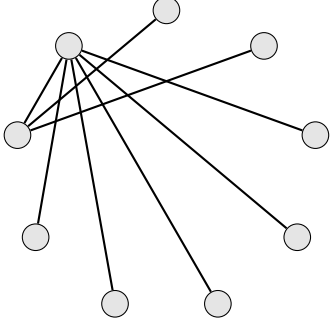
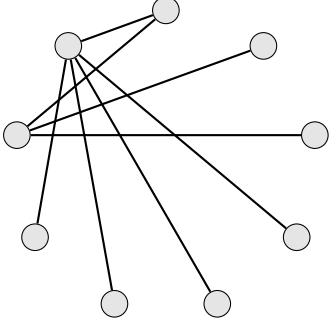
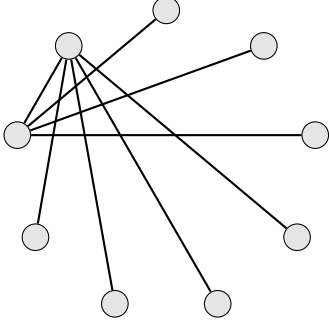
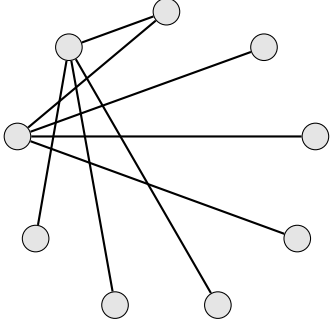
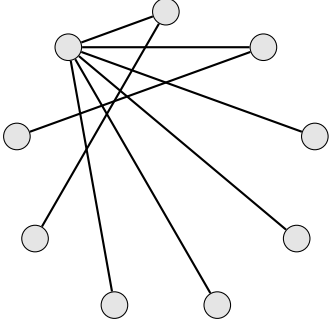
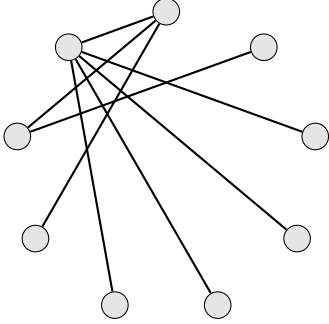
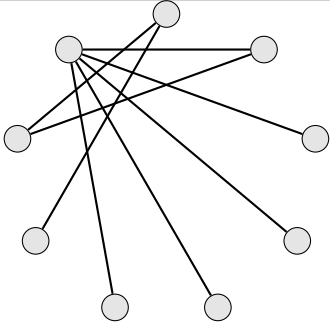
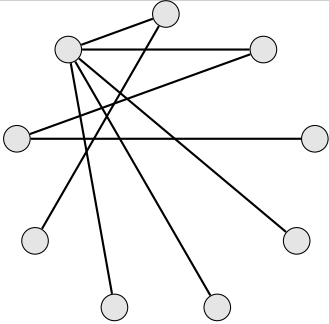
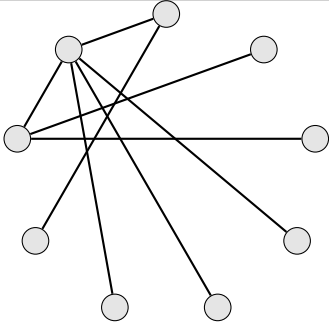
| | | |
|---|--|---|
|  |  |  |
| 1.1926 | 1.2193 | 1.2572 |
|  |  |  |
| 1.2529 | 1.3097 | 1.2895 |
|  |  |  |
| 1.3608 | 1.2543 | 1.3041 |
|  |  |  |
| 1.3129 | 1.3086 | 1.3015 |

Table C.6: All trees on nine unlabelled vertices and their maximum generalised roundness values.

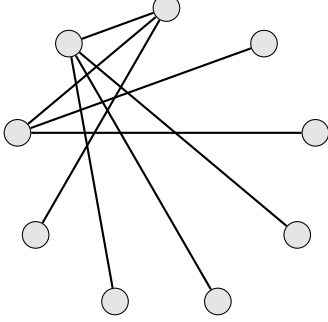
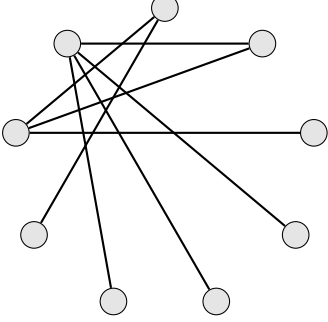
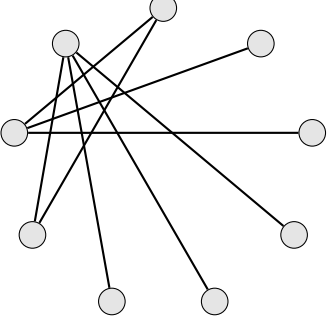
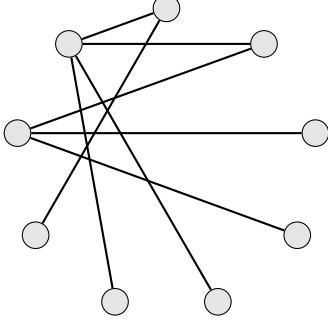
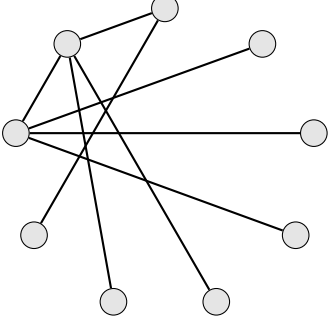
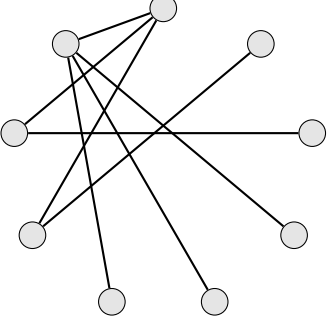
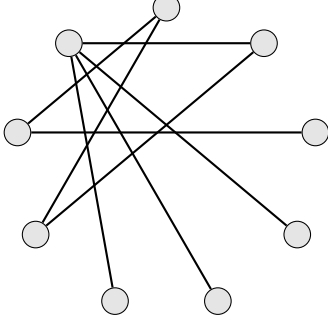
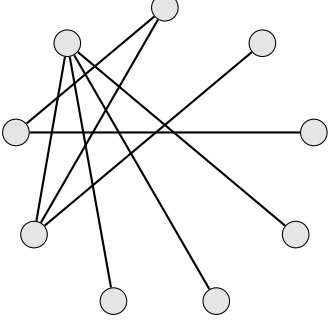
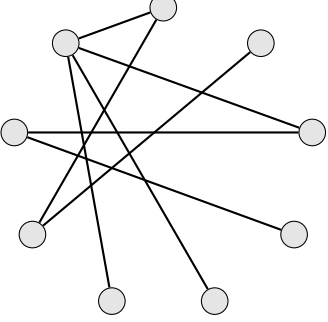
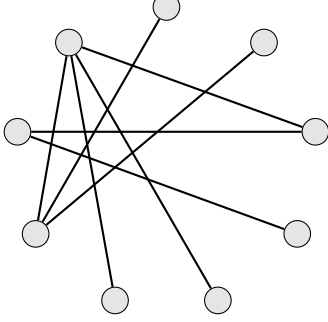
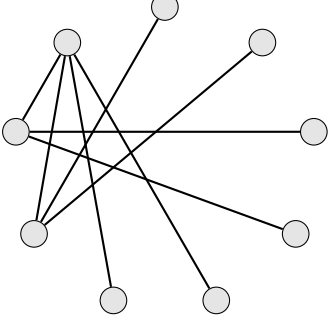
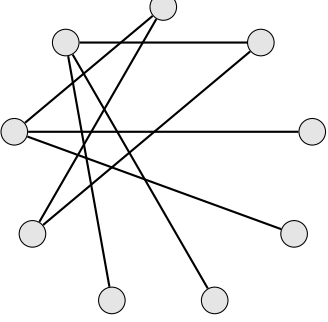
| | | |
|---|--|---|
|  |  |  |
| 1.3729 | 1.3922 | 1.3976 |
|  |  |  |
| 1.3872 | 1.3401 | 1.3796 |
|  |  |  |
| 1.4016 | 1.3827 | 1.3929 |
|  |  |  |
| 1.3784 | 1.3646 | 1.5408 |

Table C.6: All trees on nine unlabelled vertices and their maximum generalised roundness values (continued).

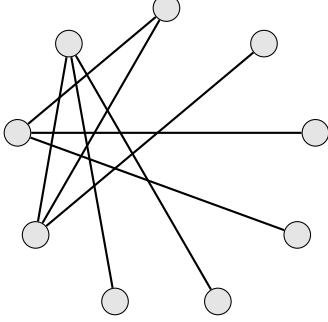
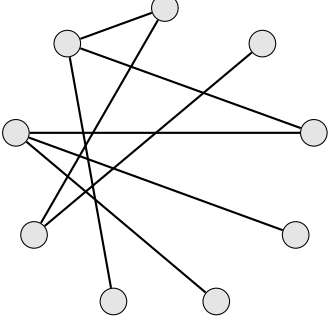
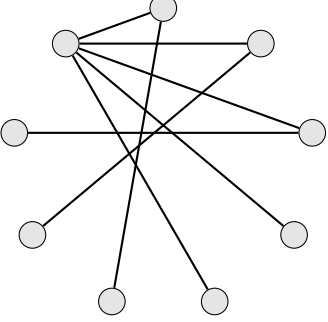
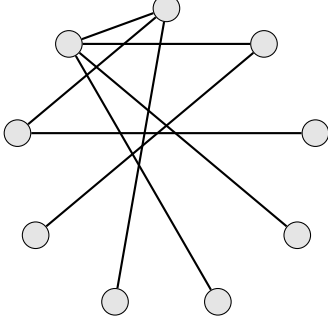
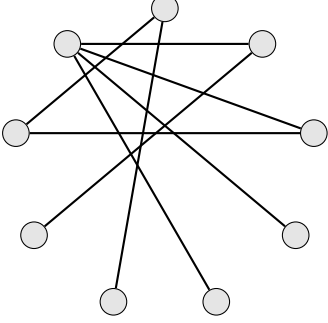
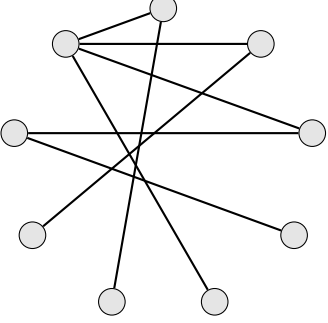
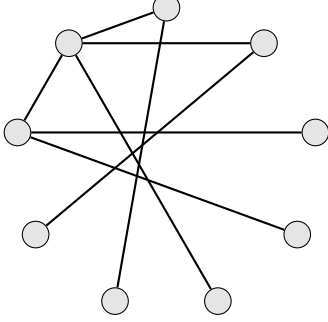
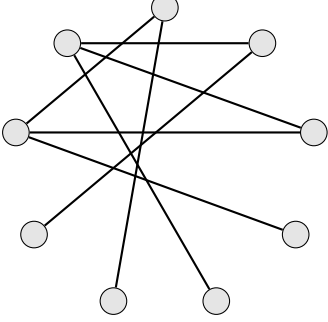
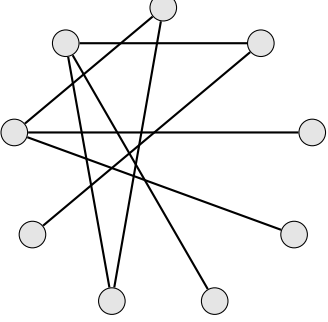
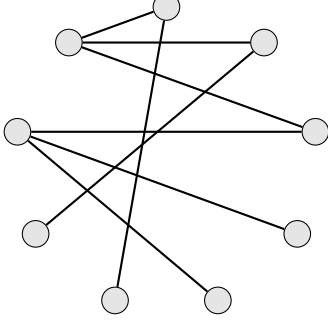
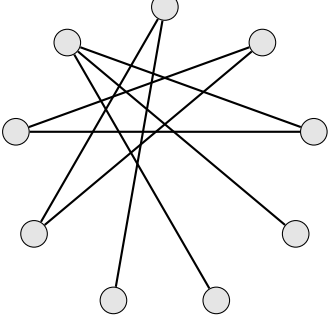
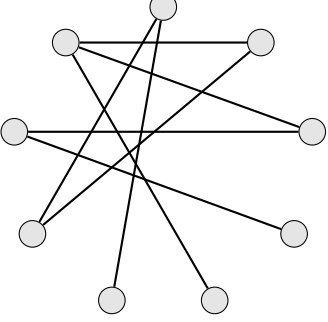
| | | |
|---|--|---|
|  |  |  |
| 1.4748 | 1.5122 | 1.3054 |
|  |  |  |
| 1.3770 | 1.3947 | 1.3882 |
|  |  |  |
| 1.3737 | 1.5079 | 1.5266 |
|  |  |  |
| 1.5064 | 1.5660 | 1.5512 |

Table C.6: All trees on nine unlabelled vertices and their maximum generalised roundness values (continued).

| | | |
|--------|--------|--------|
| | | |
| 1.4551 | 1.4971 | 1.5555 |
| | | |
| 1.4888 | 1.5410 | 1.4877 |
| | | |
| 1.4524 | 1.3835 | 1.4823 |
| | | |
| 1.5443 | 2.0000 | |

Table C.6: All trees on nine unlabelled vertices and their maximum generalised roundness values (continued).

APPENDIX D

Catalogue of graphs

The following tables list the maximum generalised roundness values of all unweighted connected graphs on $4 \leq n \leq 5$ vertices, as confirmed by both the MATLAB and Maple programs.

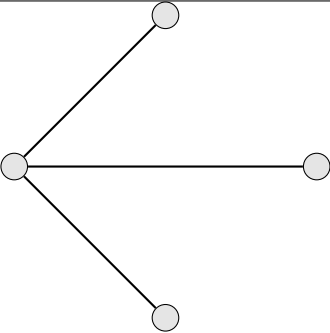
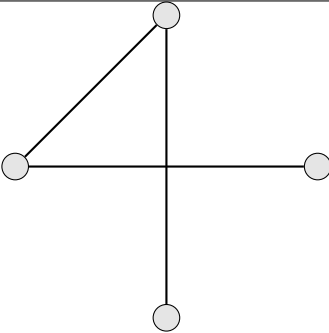
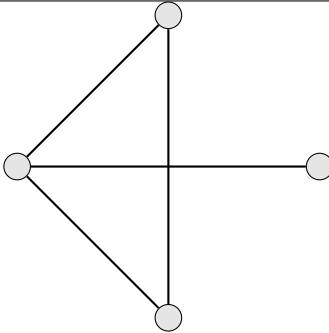
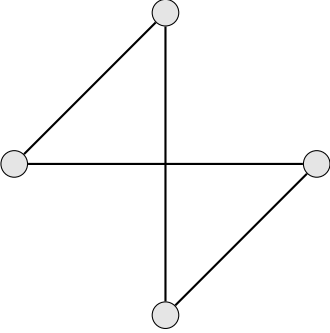
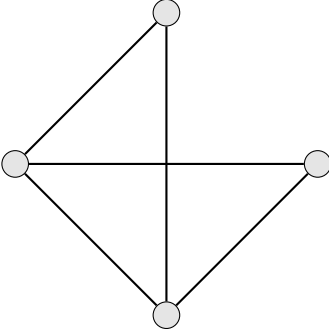
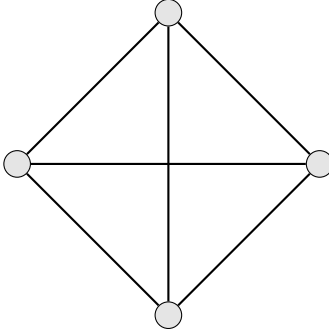
| | | |
|---|--|---|
|  |  |  |
| 1.5850 | 2.0000 | 1.9000 |
|  |  |  |
| 1.0000 | 1.5850 | ∞ |

Table D.1: All connected graphs on four unlabelled vertices and their maximum generalised roundness values.

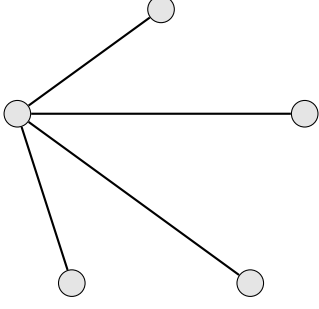
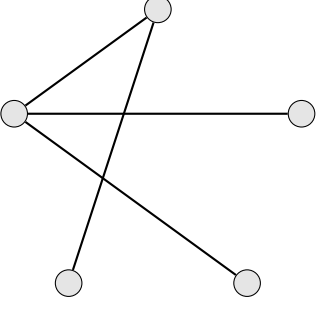
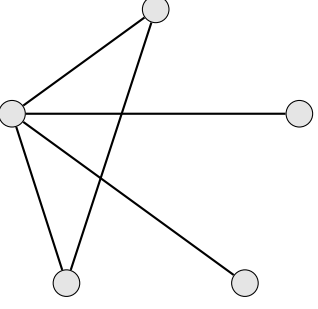
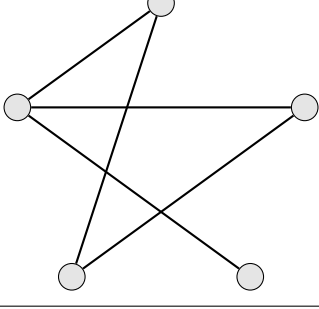
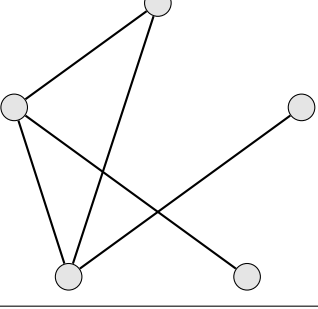
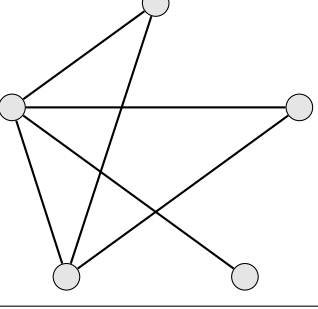
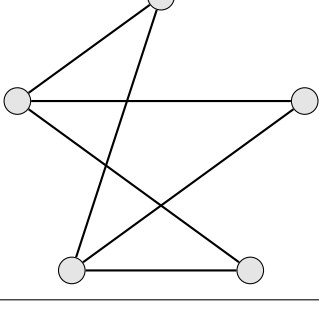
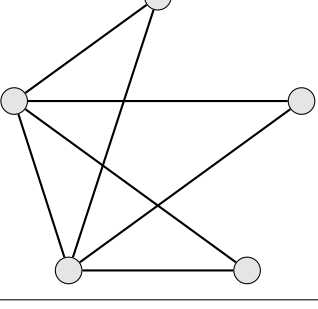
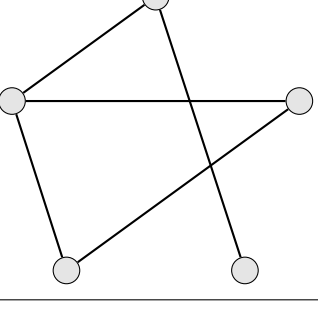
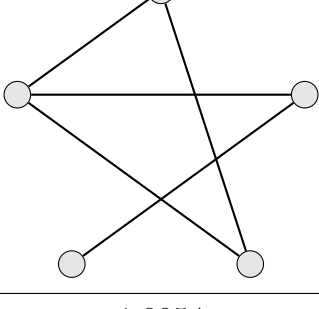
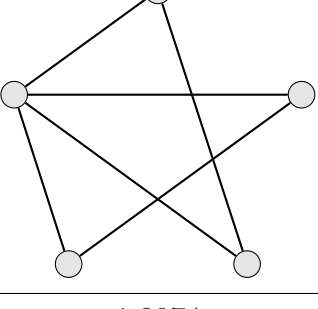
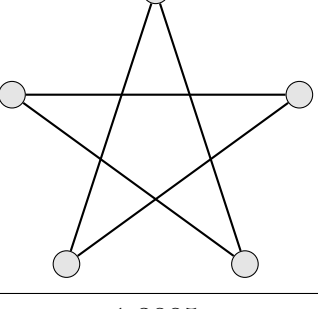
| | | |
|---|--|---|
|  |  |  |
| 1.4150 | 1.5761 | 1.5390 |
|  |  |  |
| 1.0000 | 1.8013 | 1.4886 |
|  |  |  |
| 0.7776 | 1.1699 | 2.0000 |
|  |  |  |
| 1.8954 | 1.8074 | 1.3885 |

Table D.2: All connected graphs on five unlabelled vertices and their maximum generalised roundness values.

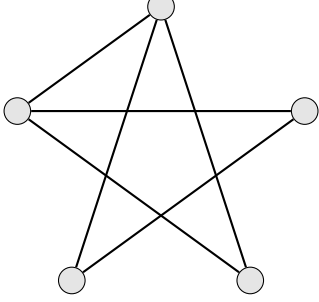
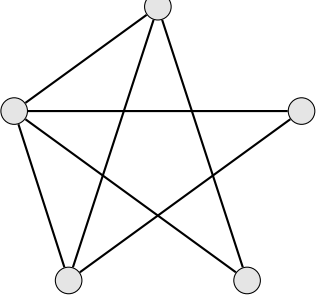
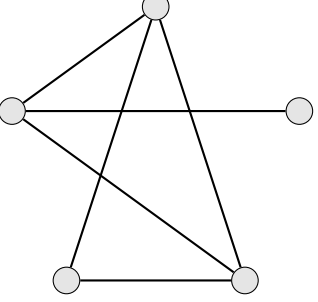
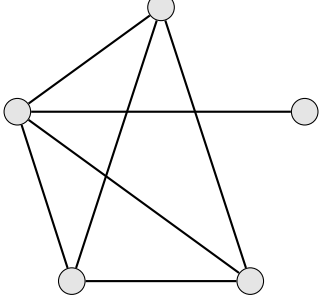
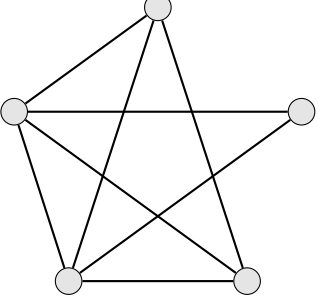
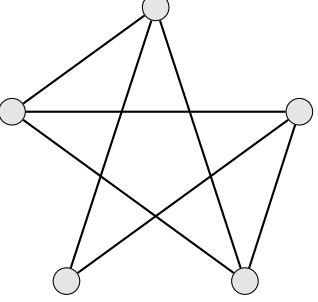
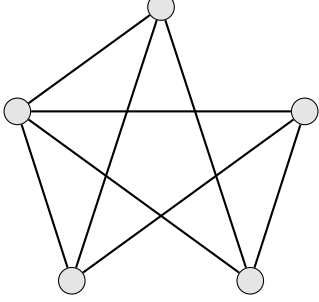
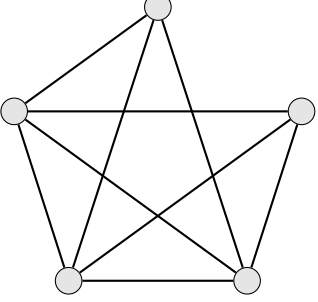
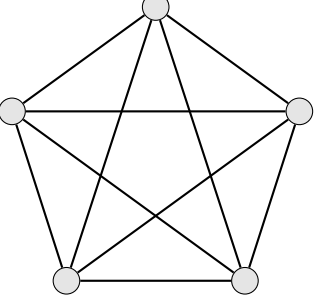
| | | |
|---|--|---|
|  |  |  |
| 1.0000 | 1.3885 | 1.5713 |
|  |  |  |
| 1.8612 | 1.4461 | 0.8918 |
|  |  |  |
| 1.0000 | 1.4150 | ∞ |

Table D.2: All connected graphs on five unlabelled vertices and their maximum generalised roundness values (continued).

References

- [1] Ailon, N. and Charikar, M., Fitting tree metrics: hierarchical clustering and phylogeny, *SIAM J. Comput.* **40**(5) (2011), 1275–1291.
- [2] Alon, N. et al., Ordinal embeddings of minimum relaxation: general properties, trees and ultrametrics, *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (2005), 650–659.
- [3] Australian Bureau of Statistics, *Census 2016, G51 Industry of employment by age by sex (SA2+)* 2017, ABS.Stat (beta). Findings based on use of ABS.Stat (beta) data.
- [4] Bartal, Y., On approximating arbitrary metrics by tree metrics, *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing* (1998), 161–168.
- [5] Bernardin, L. et al., *Maple Programming Guide*, Maplesoft: 2012.
- [6] Bernstein, S., Sur les fonctions absolument monotones, *Acta Math.* **52**(1) 1929, 1–66.
- [7] Blum, A., Hopcroft, J. and Kannan, R., *Foundations of Data Science*, Cambridge University Press: Cambridge, 2013.
- [8] Bochner, S., Vorlesungen über Fouriersche Integrale, *Monatshefte für Mathematik und Physik* **40**, 1933.
- [9] Bollobás, B., *Random Graphs*, Cambridge University Press: Cambridge, 2001.
- [10] Brent, R., *Algorithms for Minimization without Derivatives*, Prentice-Hall: Englewood Cliffs, 1973.
- [11] Bretagnolle, J., Dacunha-Castelle, D. and Krivine, J., Lois stables et espaces L^p , *Ann. Inst. H. Poincaré Sect. B (N.S.)* **2** (1965/1966), 231–259.
- [12] Bodirsky, M., Gröpl, C. and Kang, M., Generating labeled planar graphs uniformly at random, *Theoretical Computer Science* **379** (2007), 377–386.
- [13] Brinkmann, G., Coolsaet, K., Goedgebeur, J., Mélot, H., House of Graphs: a database of interesting graphs, *Discrete Appl. Math.* **161**(1–2) (2013), 311–314. Retrieved 31 March 2020 from <http://hog.grinvin.org>.

- [14] Brinkmann, G., Goedbegeur, J. and McKay, B., Fullerenes, Retrieved 15 April 2020 from <https://caagt.ugent.be/buckygen/>.
- [15] Brinkmann, G. and McKay, B., plantri and fullgen, Retrieved 8 April 2020 from <https://users.cecs.anu.edu.au/~bdm/plantri/>.
- [16] Caffarelli, E., Doust, I. and Weston, A., Metric trees of generalized roundness one, *Aequationes Mathematicae* **83**(3) (2012), 239–256.
- [17] Carlsson, G. and Mémoli, F., Characterization, stability and convergence of hierarchical clustering methods, *J. Machine Learning Research* **11** (2010), 1425–1470.
- [18] Cayley, A., On a theorem in the geometry of position, *Cambridge Math. J.* **II** (1841), 267–271.
- [19] Cayley, A., A theorem on trees, *Quart. J. Pure Appl. Math.* **23** (1889), 376–378.
- [20] Charikar, M. et al., Approximating a finite metric by a small number of tree metrics, *Proceedings of the 39th Annual Symposium on Foundations of Computer Science* (1998), 379–388.
- [21] Dahma, A. and Lennard, C., Generalized roundness of the Schatten class, \mathcal{C}_p , *J. Math. Anal. Appl.* **412**(2) (2014), 676–684.
- [22] Dekker, T., Finding a zero by means of successive linear interpolation, *Constructive Aspects of the Fundamental Theorem of Algebra*, 37–48, Wiley-Interscience: New York, 1969.
- [23] Deza, M. and Maehara, H., Metric transforms and Euclidean embeddings, *Trans. Amer. Math. Soc.* **317**(2) (1990), 661–671.
- [24] Deza, M. and Maehara, H., A few applications of negative-type inequalities, *Graphs Combin.* **10**(3) (1994), 255–262.
- [25] Diestel, R., *Graph Theory*, Springer-Verlag: Heidelberg, 2016.
- [26] Doust, I., Sánchez, S. and Weston, A., A direct proof that $\ell_\infty^{(3)}$ has generalized roundness zero, *Expo. Math.* **33**(2) (2015), 259–267.
- [27] Doust, I., Sánchez, S. and Weston, A., Asymptotic negative type properties of finite ultrametric spaces, *J. Math. Anal. Appl.* **446**(2) (2017), 1776–1793.
- [28] Doust, I. and Weston, A., Enhanced negative type for finite metric trees, *J. Funct. Anal.* **254**(9) (2008), 2336–2364.
- [29] Durand, E., Équations du type $F(x) = 0$; racines d’un polynôme., in *Tome I: Solutions numériques des équations algébriques*, Masson et Cie (editors) (1960).

- [30] Enflo, P., On the nonexistence of uniform homeomorphisms between L_p -spaces, *Ark. Mat.* **8** (1968), 103–105.
- [31] Enflo, P., On a problem of Smirnov, *Ark. Mat.* **8** (1969), 107–109.
- [32] Erdős, P., Harary, F. and Tutte, W., On the dimension of a graph, *Mathematika* **12** (1965), 118–122.
- [33] Erdős, P. and Rényi, A., On the evolution of random graphs, *Bull. Inst. Internat. Statist.* **38** (1961), 343–347.
- [34] Erdős, P. and Simonovits, M., On the chromatic number of geometric graphs, *Ars Combin.* **9** (1980), 229–246.
- [35] Fakcharoenphol, J., Rao, S. and Talwar, K., A tight bound on approximating arbitrary metrics by tree metrics, *Proceedings of the Thirty-Fifth Annual ACM Symposium on Theory of Computing* (2003), 448–455.
- [36] Faver, T. et al., Roundness properties of ultrametric spaces, *Glasgow Math. J.* **56** (2014), 519–535.
- [37] Frankl, N., Kupavskii, A. and Swanepoel, K., Embedding graphs in Euclidean space, *J. Combin. Theory Series A* **171** (2020), 105146, 14pp.
- [38] Gilbert, E., Random graphs, *Ann. Math. Statist.* **30** (1959), 1141–1144.
- [39] Giménez, O. and Noy, M., Asymptotic enumeration and limit laws of planar graphs, *J. Amer. Math. Soc.* **22**(2) (2009), 309–329.
- [40] Goldschmidt, C., Random minimum spanning trees — Mathematical Institute, Retrieved 20 September 2020 from <https://www.maths.ox.ac.uk/node/30217>.
- [41] Graham, R., On isometric embeddings of graphs, *Progress in graph theory*, 307–322, Academic Press: Toronto, 1984.
- [42] Graham, R., Isometric embeddings of graphs, *Selected topics in graph theory*, 133–150, Academic Press: Toronto, 1988.
- [43] Graham, R. and Winkler, P., Isometric embeddings of graphs, *Proc. Nat. Acad. Sci. U.S.A.* **81** (1984), 7259–7260.
- [44] Graham, R. and Winkler, P., On isometric embeddings of graphs, *Trans. Amer. Math. Soc.* **288**(2) (1984), 527–536.
- [45] Griffith, D., Generating random connected planar graphs, *GeoInformatica* **22**(4) (2018), 767–782.
- [46] Heck, A., *Introduction to Maple (3rd ed.)*, Springer-Verlag: New York, 2003.
- [47] Hjorth, P., Lisoněk, P., Markvorsen, S. and Thomassen, C., Finite metric spaces of strictly negative type, *Linear Algebra Appl.* **270** (1998), 255–273.

- [48] Kerner, I., Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen, *Numer. Math.* **8** (1966) 290–294.
- [49] Koldobsky, A., Generalised Lévy representation of norms and isometric embeddings into L_p -spaces, *Ann. Inst. H. Poincaré Probab. Statist.* **28**(3) (1992), 335–353.
- [50] Kruskal, J., On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Amer. Math. Soc.* **7**(1) (1956), 48–50.
- [51] Lennard, C., Tonge, A. and Weston, A., Generalized roundness and negative type, *Michigan Math. J.* **44**(1) (1997), 37–45.
- [52] Li, H. and Weston, A., Strict p -negative type of a metric space, *Positivity* **14**(3) (2010), 529–545.
- [53] Lindenstrauss, J., On nonlinear projections in Banach spaces, *Michigan Math. J.* **11** (1964), 263–287.
- [54] Mathias, M., Über positive Fourier-Integrale, *Mathematische Zeitschrift* **16** (1923), 103–125.
- [55] McDiarmid, C., Steger, A. and Welsh, D., Random planar graphs, *J. Combin. Theory Series B* **93**(2) (2005), 187–205.
- [56] McKay, B. and Piperno, A., Practical graph isomorphism, II, *J. Symbolic Comput.* **60** (2014), 94–112.
- [57] Menger, K., Untersuchungen über allgemeine Metrik, *Math. Ann.* **100** (1928), 175–163.
- [58] Menger, K., New Foundation of Euclidean Geometry, *Amer. J. Math.* **53**(4) (1931), 721–745.
- [59] Misiewicz, J., Positive definite functions on $\ell_\infty^{(n)}$, *Statist. and Prob. Letters* **8**(3) (1989), 255–260.
- [60] Moler, C., *Numerical Computing with MATLAB*, SIAM: Philadelphia, 2008. Available at https://www.mathworks.com/moler/index_ncm.html
- [61] Moon, J., *Counting Labelled Trees*, Canadian Mathematical Congress: Montreal, 1970.
- [62] Murugan, M., Supremal p -negative type of vertex transitive graphs, *J. Math. Anal. Appl.* **391**(2) (2012), 376–381.
- [63] Mütze, T., Sawada, J. and Williams, A., The Combinatorial Object Server. Retrieved 31 March 2020 from <http://combos.org>.
- [64] OEIS Foundation Inc. (2020), The On-Line Encyclopedia of Integer Sequences, Retrieved 31 March 2020 from <http://oeis.org>.

- [65] Otter, R., The Number of Trees, *Ann. of Math. Second Series* **49**(3) (1948), 583–599.
- [66] Palka, Z., Bipartite complete induced subgraphs of a random graph, *Random graphs '83 (Poznań)*, 209–219, North-Holland: Amsterdam, 1985.
- [67] Prim, R., Shortest connection networks and some generalizations, *Bell Sys. Tech. J.*, **36**(6) (1957) 1389–1401.
- [68] Prüfer, H., Neuer Beweis eines Satzes ber Permutationen, *Arch. Math. Phys.* **27** (1918), 742–744.
- [69] Robbins, H., A Remark on Stirling’s Formula, *Amer. Math. Monthly* **62**(1) (1955), 26–29.
- [70] Roth, R. and Winkler, P., Collapse of the metric hierarchy for bipartite graphs, *European J. Combin.* **7**(4) (1986), 371–375.
- [71] Sánchez, S., On the supremal p -negative type of finite metric spaces, *J. Math. Anal. Appl.* **389**(1) (2012), 98–107.
- [72] Schoenberg, I., Remarks to Maurice Fréchet’s article “Sur la définition axiomatique d’une classe d’espaces distanciés vectoriellement applicable sur l’espace de Hilbert”, *Ann. of Math.* **36**(3) (1935), 724–732.
- [73] Schoenberg, I., On certain metric spaces arising from Euclidean spaces by a change of metric and their imbedding in Hilbert space, *Ann. of Math.* **38**(4) (1937), 787–793.
- [74] Schoenberg, I., Metric spaces and positive definite functions, *Trans. Amer. Math. Soc.* **44**(3) (1938), 522–536.
- [75] Semple, C. and Steel, M., *Phylogenetics, Volume 24 of Oxford Lecture Series in Mathematics and its Applications*, Oxford University Press: Oxford, 2003.
- [76] Whittaker, E. and Watson, G., *A course of modern analysis: an introduction to the general theory of infinite processes and of analytic functions: with an account of the principal transcendental functions (4th ed.)*, Cambridge University Press: London, 1927.
- [77] Weber, G., Ohno-Machado, L. and Shieber, S., Representation in stochastic search for phylogenetic tree reconstruction, *J. Biomedical Informatics* **39** (2006), 43–50.
- [78] Weston, A., Optimal lower bound on the supremal strict p -negative type of a finite metric space, *Bull. Aust. Math. Soc.* **80**(3) (2009), 486–497.
- [79] Weston, A., On the generalized roundness of finite metric spaces, *J. Math. Anal. Appl.* **192**(2) (1995), 323–334.
- [80] Wilf, H., *generatingfunctionology (2nd ed.)*, Academic Press: San Diego, 1994.

- [81] Wilson, D. B., Generating random spanning trees more quickly than the cover time, *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing* (1996), 296–303.
- [82] Wolf, R., On the gap of finite metric spaces of p -negative type, *Linear Algebra Appl.* **436**(5) (2012), 1246–1257.
- [83] Wright, R., Richmond, B., Odlyzko, A. and McKay, B., Constant time generation of free trees, *SIAM J. Comput.* **15**(2) (1986), 540–548.
- [84] Yost, D., L_1 contains every two-dimensional normed space, *Ann. Polon. Math.* **49** (1988), 17–19.