Graph geometry from effective resistances

Karel Devriendt
Mansfield College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2022

To my parents,
Hilde & Koen

Acknowledgements

There are many people without whom this thesis would have never been possible. First of all, I would like to thank my supervisor Renaud Lambiotte. Thank you Renaud, for guiding me through my DPhil with your relentless enthusiasm, for the many stimulating discussions and invaluable advice, and for encouraging me to pursue my interests — wherever they lead me. I deeply admire your 'people first' attitude as a supervisor and academic, and hope to live up to your example wherever I end up.

While I did not have an official second supervisor, I want to thank Piet Van Mieghem for his mentorship before and during the early stages of my DPhil. Thank you Piet, for all the opportunities, support and advice throughout the years and, of course, for introducing me to effective resistances and Fiedler's work. The 'open questions' in the footnotes of your Networking class were the starting point for my interest in graph theory.

One of the great joys in research is the community, and I have had the pleasure to work with a wonderful group of people during my DPhil. Thank you Ben, Marc, Samuel, Bastian, Sanjukta, Michele, Marilena, Shazia, Alice and Alex for the pleasant collaborations and for everything I learned from you. Thank you also to my many other colleagues at Oxford, the Alan Turing Institute, the Networks Seminar and the students I taught and supervised for the lovely and stimulating work environment. Thank you Rodrigo, for welcoming me to the department and for your warm friendship, and Yu for making the period of writing this thesis more pleasant than I could have ever hoped for.

I am fortunate to have been surrounded by amazing friends and housemates wherever I was staying, in London, Leuven or Oxford. I will never forget the dinners, the parties, the conversations, the socially-distanced walks and all the other great times. A big thanks to Thomas, Alice, Bailey, Eboni and Katie for your friendship, your inspiration and for making our house a real home.

Due to the COVID-19 pandemic, these past few years have been extraordinary and difficult in many ways. Without the support of everyone around me and the courage and dedication of all the medical staff, the frontline workers, the teachers, parents, etc. who worked untiringly to make the world go round during these exceptional times, I could have never even dreamt of continuing my research. Thank you all, and in particular my temporary housemates in Leuven for their hospitality, support and friendship during the first few months of the pandemic.

Thank you mama en papa, for all the opportunities, love and support. For always guiding me (but never steering), for your continued interest in my work and for your eternal 'good luck' postcards. Thank you Bart en Annelies for being an amazing older brother and sister and thank you to my grandparents for their unwavering support and interest in my (academic) life. Thank you Bart en Leen and the rest of the family for all the good times.

And finally, thank you Enja. Thank you for being my closest companion in everything I do, for having shared with me all the steps of this process — from the fun and exciting parts to the boring and difficult ones —, for the inspiration and for your support and patience. Thank you for the innumerable other things which these pages are too narrow to contain.

Abstract

The main subject of this thesis is the effective resistance — a measure of distance between the nodes of a graph which reflects how close or well-connected two nodes are, taking into account all paths between them and their interconnections.

The effective resistance has many other properties, however, and the contribution of this thesis is twofold. First, we provide a self-contained and unified exposition of some important but not well-known results on the effective resistance: the Fiedler–Bapat identity between Laplacian and resistance matrices and the graph–simplex correspondence. Second, we introduce and study two new applications of the effective resistance: a distance-based measure of variance for distributions on the nodes of a graph, and two resistance-based graph invariants $\mathbf{p}$ and σ^2 which we interpret in the context of discrete curvature.

These contributions are unified in their geometric perspective on the effective resistance — ranging from the celebrated concept of resistance distance, to the lesser-known but far-reaching relation to simplex geometry and finally, a newly discovered relation to discrete curvature.

While the focus lies on developing the mathematical theory of effective resistances in a unified matrix-theoretic language, we also discuss many examples and some applications. In particular, the proposed measures of variance and covariance are readily applicable in practice to study functional data (signals, distributions, etc.) defined on graphs.

We hope that the exposition in this thesis will enable and stimulate further research on this fascinating subject and in particular on the theory of the resistance-based graph invariants $\mathbf{p}$ and σ^2, the full depths of which still remain to be explored.

Contents

Papers related to this thesis

Some of the content in this thesis appeared before in journal publications.

Parts of Chapters 2–4 appeared in the following paper:

[70] Karel Devriendt. *Effective resistance is more than distance: Laplacians, simplices and the Schur complement*, Linear Algebra and its Applications, 639:24-49, 2022.

Parts of Chapter 5 appeared in the following paper:

[73] Karel Devriendt, Samuel Martin-Gutierrez and Renaud Lambiotte. *Variance and covariance of distributions on graphs*, SIAM Review, 64(2):343-359, 2022.

Most of Chapter 6 appeared in the following paper:

[72] Karel Devriendt and Renaud Lambiotte. *Discrete curvature on graphs from the effective resistance*, J. Phys. Complex., 3(2):025008, 2022.

The following publications on nonlinear network dynamics resulted from work during my DPhil but are not included in this thesis:

[71] Karel Devriendt and Renaud Lambiotte. *Nonlinear network dynamics with consensus-dissensus bifurcation*, Journal of Nonlinear Science, 31:18, 2021.

[129] Marc Homs-Dones, Karel Devriendt and Renaud Lambiotte. *Nonlinear consensus on networks: Equilibria, effective resistance, and trees of motifs*, SIAM Journal on Applied Dynamical Systems, 20(3):1544-1570, 2021.

Statement of authorship & originality

Much of Chapters 2–4 is expository work. The presented results are either previously known or we include specific references if they are new contributions. Apart from the unified exposition and new proofs for several results, the original contributions in these chapters are as follows:

- In Chapter 2, the bounds in Propositions 2.10 and 2.16 are implicit in known results but, to our knowledge, have not yet appeared in this form.
- In Chapter 3, the following are original contributions: the characterization of resistance matrices in Theorem 3.9, the derivation of triangle excess matrices from the Fiedler–Bapat identity, Corollary 3.34 for the effective resistances in node-merged graphs, and the results in Section 3.4 on the resistance radius σ^2 as a set function.
- In Chapter 4, the following are original contributions: the matrix characterization of resistance metric spaces in Proposition 4.7, and Theorem 4.24 extending the graph–simplex correspondence to maps between graphs and simplices (which is implicit in Fiedler's work).

Chapter 5 contains original work done in collaboration with Renaud Lambiotte and Samuel Martin-Gutierrez and parts of this chapter also appear in the following thesis:

- Samuel Martin-Gutierrez. *Models of social interaction and data analysis of online user behavior during polarization processes*, PhD Thesis, Universidad Politécnica de Madrid, Madrid, 2020.

The initial research question behind Chapter 5 was formulated by Renaud Lambiotte and Samuel Martin-Gutierrez. In Section 5.2, the data collection, coding and preparation of the figures (Figures 5.3–5.6 and Table 5.1) was done by Samuel Martin-Gutierrez for the paper [73]. All other work and writing in the chapter are mine.

Chapter 6 contains original work done in collaboration with Renaud Lambiotte.

Chapter 1

Introduction

1.1 Graphs and networks

A graph, or network, is a simple mathematical object that consists of nodes and links that connect pairs of nodes. Typically these are drawn using dots and lines, giving rise to figures that probably look familiar to many, such as Figure 1.1.

Figure 1.1: A network of mathematical concepts and their relations (see Chapter 5).

But why do graphs look so familiar? In the preface of his now-famous textbook on graph theory, Frank Harary remarked that "*[i]t has become fashionable to mention that there are applications of graph theory to some areas of physics, chemistry, communication science, computer technology, electrical and civil engineering, architecture, operational research, genetics, psychology, sociology, economics, anthropology, and linguistics*" [123, p. v]. In other words, graphs are everywhere! This was in 1969, and five decades later such statements — often including an even longer list — are still fashionable. This variety of disciplines and applications, loosely related by their common methodology of describing a system's connections as a graph, has since then

developed into a new field of study called 'network science' [184], complete with its own conferences, journals, institutes and graduate programs.

In parallel with the growth of network science, there has been a continuous development of new concepts, tools and ideas on how to analyse graphs. We may note, for instance, that some of the tools developed in this interdisciplinary field have likely been instrumental in the response to the recent COVID-19 pandemic, and they could play a similar important role in the other systemic challenges we face in the next few decades. Partly, these tools are borrowed from graph theory and other disciplines such as dynamical systems, statistical physics, probability and linear algebra, but working with large, heterogeneous and generally messy real-world networks often requires the development of new ideas or a reconceptualization of old ones. The contributions of this thesis are an attempt at interfacing between these two sides: the mathematics of graph theory and the questions arising from applications. The first three chapters, Chapters 2–4, focus on developing the theory of 'effective resistances' in graphs while Chapters 5 and 6 pay more attention to showing examples, developing a conceptual understanding and proposing some new tools to study graphs and associated data.

As network scientists and engineers come up with new ways to use graphs, the mathematical theory of graphs has of course not been standing still. Harary's preface continues: "*The theory [of graphs] is also intimately related to many branches of mathematics, including group theory, matrix theory, numerical analysis, probability, topology, and combinatorics*" [123, p. v]. Again, this is even more so the case today with results like Szemerédi's regularity Lemma [150], to name but one, having percolated through many fields of mathematics. One particular branch that might be added to Harary's list and that will be considered in this thesis is *geometry*. Geometric graph theory can refer to "*a large, amorphous body of research*" according to [190, p. 1], including topics like straight-line embeddings of graphs [190] (the rigid cousin of topological embeddings [178]), geometric representations of graphs [164], random geometric graphs [195], distances and metrics on graphs [39, 102] and discrete curvature [26, 183, 199], with numerous related applications [27, 35, 121, 152, 187, 223, 255].

In other words, there is at present no one unified geometric theory of graphs, but the connections are myriad (see also [164, Ch. 20] and [27]). As suggested by the title, this thesis attempts to further add to this amorphous body of research on graph geometry. In particular, we discuss three geometric perspectives on graphs that follow from the 'effective resistance': we study the effective resistance as a metric between the nodes of a graph and discuss some applications, we study the geometric characterization of weighted graphs as hyperacute simplices and thirdly, we propose

a new notion of discrete curvature based on the effective resistance. The specifics of these perspectives are further detailed in the contributions section below and, of course, in the respective chapters.

1.2 The effective resistance

The second giveaway of the title is that this thesis deals with the concept of effective resistance — this will be the main lens through which we will look at graphs. The effective resistance is a measure of distance between the nodes of a graph and reflects how close or well-connected two nodes are based on the different paths that run between them and their interconnections. While originally defined as a tool for electrical circuit analysis and design, the effective resistance can be seen as a purely graph-theoretic concept with a rich theory independent of its original applied context.

Our focus will be on the matrix-theoretic and geometric properties of the effective resistance, but we stress that this is just one of the many different perspectives. The list below provides an overview of these perspectives, most of which will not be discussed any further.

- **Electrical circuits:** As noted, the effective resistance originated in the context of electrical circuit analysis and is discussed in most textbooks on this subject. Relevant references are the broader mathematical theory of electrical circuits [6, 79, 260], Kirchhoff's original analysis of electrical circuits [143, 144], and the algebraic (potential) theory of electrical circuits [23]. Some important results on the effective resistance that stem from this context are Thomson's principle and Rayleigh monotonicity (see e.g. [168]).

- **Random walks:** A standard reference work on the relation between electrical circuits, effective resistance and graphs is the monograph of Doyle and Snell [80] and we also recommend Tetali's work [236]. Other relevant references are the surveys [169, 192], the early contributions by Nash-Williams [224], and the work of Chandra et al. [46].

- **Topology and combinatorics:** In a first encounter with the effective resistance, one often learns that the effective resistance is calculated by repeatedly making small local changes to an electrical circuit — series, parallel and $Y - \Delta$ transformations[1] — until only a single resistor is left with the sought-after ef-

[1] Usually, removing degree-one nodes and self-loops are also counted among the electrical transformations [59].

fective resistance. This procedure works in the case of planar graphs [86] and a more general study of the combinatorics of these *electrical transformations* and their relation to topology can be found in [47, 59, 63]. See also the seminal work of Tutte et al. on square tilings of rectangles [37].

- **Spanning trees:** The relation between spanning trees and electrical circuits traces back to the original work of Kirchhoff, with the celebrated matrix–tree Theorem as an important exponent [22, 23]. The relation between effective resistance and (random) spanning trees (Theorem 2.13) is generalized in the far-reaching transfer-current Theorem [41] which in turn is part of the theory of determinantal point processes [167]. There is also a strong relation between spanning trees, effective resistances and computational problems on graphs [82, 171, 218]. An important practical consequence is that the effective resistance can be approximated efficiently, as discussed in Section 2.3.2.

- **Matrix theory:** Detailed references on the matrix theory of effective resistances will appear throughout the thesis. The basic references are Fiedler's and Bapat's books [11, 96] and further relevant works are [232, 264].

- **Metrics and geometry:** One of the celebrated properties of the effective resistance is that it is a metric between the nodes of a graph. This was discovered independently by Gvishiani & Gurvich [117, 119] and Klein & Randíc [146, 149] and is implicit in the 'no-amplification property' in electrical circuit analysis [214, Eq. 6], [236]. Geometrically, the effective resistance gives rise to a simplicial embedding of graphs which was discovered independently by Fiedler [92, 96] and Subak-Sharpe & Moore [179, 214]. Detailed references are given in Chapter 4.

- **Chemical graph theory:** In mathematical chemistry, the structure of chemical compounds is sometimes modeled using graphs, and various graph invariants are studied in relation to chemical properties. The effective resistance has a long history in this context, in particular with a focus on the *Kirchhoff index* (Definition 3.4) as a resistance-based graph invariant [18, 118, 149, 191, 240].

- **Robustness:** The effective resistance and in particular the Kirchhoff index have been linked to various notions of 'robustness' of graphs. The relation to "structural robustness" of networks is treated for instance in [85, 104] while the relation to "dynamical robustness" is considered in [243, 244, 265].

- **Generalizations:** The concept of effective resistance can be generalized in a number of ways. First, the classical theory of passive electrical circuits also comprises reactive elements (inductors and capacitors) in addition to resistors. In these circuits, the *effective impedance* describes a frequency-dependent transfer function between the (alternating) current and voltage between two points in the circuit [6, 78]. The effective resistance has furthermore been generalized to other structures such as directed graphs [266], simplicial complexes [151], cellular sheaves [122], infinite graphs [80, 237] and fractals [142]. As a metric, the effective resistance is generalized by strict negative type metrics as discussed in detail in Chapter 4.

While this variety of contexts and approaches may pose some problems in terms of translating and communicating new results between the different research communities involved, it mainly enriches the theory of effective resistances by making available a broad set of tools and concepts.

1.3 Contributions and outline

The contribution of this thesis is twofold. First, in Chapters 1–4 we provide a self-contained, unified and detailed introduction to the effective resistance and a number of important but not widely known results. Second, in Chapters 5 and 6 we discuss two new applications of effective resistances: we propose a measure of (co)variance for distributions on a graph and study its properties, and we study two graph-based invariants $\mathbf{p}$ and σ^2 and their relation to discrete curvature. The specific contributions and new results are summarized in the statement of originality.

The outline below details the most important results (both known and new) discussed in each of the chapters.

Chapter 2 starts by introducing the relevant mathematical objects. We pay particular attention to the (combinatorial) Laplacian matrix Q, its Moore–Penrose pseudoinverse $Q^\dagger$ and their properties. Definition (2.8) introduces the effective resistance ω_{ij} and we discuss a number of aspects: some intuitions, computational results, the effective resistance as a metric and finally the relation to random spanning trees.

Chapter 3 deals with the matrix theory of the effective resistance, based on the resistance matrix Ω, which contains all pairwise effective resistances. The main result is the Fiedler–Bapat identity (Theorem 3.6), which is an inverse matrix identity

between the Laplacian Q and resistance matrix Ω. The Fiedler–Bapat identity features the resistance-based invariants $\mathbf{p}$ and σ^2 and we derive some initial properties of these invariants. Theorem 3.9 is a new concise matrix characterization of resistance matrices. Next, we discuss how the Fiedler–Bapat identity together with the Schur complement lead to the notion of Kron reduction (Definition 3.14) and merged nodes (Definition 3.27). Finally, we discuss some properties of σ^2 as a set function defined on subsets of the nodes (Definition 3.37).

Chapter 4 deals with the effective resistance as a metric. We start with a proof of the fact that the effective resistance is a metric (Theorem 4.2), which leads to the notion of a resistance metric space $(\mathcal{N}, \omega)$. We discuss the classification of the resistance metric as a metric of strict negative type and show that every resistance metric space corresponds to the squared Euclidean distances between the vertices of a hyperacute simplex S (a simplex with nonobtuse dihedral angles) in Theorem 4.19. Equivalently, this is formulated as Fiedler's bijective graph–simplex correspondence, which associates a hyperacute simplex to each weighted graph. We discuss how this correspondence is also 'functorial' since maps between graphs are in correspondence with maps between simplices (Theorem 4.24). Finally, we introduce the resistive embedding φ, which maps functions defined on the nodes of a graph to the ambient Euclidean space of S. This determines a bijection between distributions on the nodes and the simplex S (Theorem 4.28).

Chapter 5 introduces a measure of variance var_d for distributions defined on the nodes of a graph (Definition 5.1), with respect to a distance d. This measure is illustrated in a number of examples and experiments on an empirical network. We then consider the maximum-variance problem: "$\max_{\mathbf{f}} \mathrm{var}_d(\mathbf{f})$ over all distributions $\mathbf{f}$". We first provide a characterization of maximum variance distributions $\mathbf{f}^\star$ for general distances d and show how the maximum-variance problem can be redefined as a problem over subsets of the metric space (equivalently, submatrices of the distance matrix) in Theorem 5.15. The problem simplifies for negative type metrics and for resistance metrics a greedy solution is possible (Theorem 5.23) and the solution is characterized in terms of $\mathbf{p}$ and σ^2. Finally, we interpret maximum variance distributions and their support $\mathrm{supp}(\mathbf{f}^\star)$ in terms of a certain notion of 'boundary' of a graph and discuss a number of examples.

Chapter 6 presents a discrete curvature interpretation of $\mathbf{p}$. We discuss some background on discrete curvature and present several arguments for the interpretation of $\mathbf{p}$, defined on the nodes, and a related notion κ defined on the links, as discrete curvatures. In particular, we discuss theoretical and numerical evidence for

their convergence in the case of Euclidean random geometric graphs and show their relation to existing notions of discrete curvature. We study an associated discrete Ricci flow — a differential equation describing an evolution $G(t)$ of graphs — and a simplified local flow which is related to Kron reduction (Proposition 6.21). Finally, we consider graphs with positive curvature $\mathbf{p} > 0$ and show that they have strong connectivity properties (Theorem 6.31) and that their resistance radius σ^2 has submodular properties (see expression (6.29)) as a set function (Theorem 6.33).

Appendix A discusses some technical results in linear algebra: matrix theory related to graphs, spectral results on symmetric matrices, and the Moore–Penrose pseudoinverse.

1.4 Notation and terminology

This thesis uses some non-standard notation and terminology. For instance, the standard terminology in graph theory is 'vertices and edges' while we use 'nodes and links' when describing a graph. The first reason for this choice is the appearance of simplices, where vertices and edges have a related but different meaning. A second more general reason is the context of this thesis at the interface between research communities with different standards and conventions. We further remark that our non-standard notation for the Schur complement is a result of trying to find a unified notation for Schur complements, Kron reductions and faces of a simplex.

This thesis also introduces several new names for resistance-related mathematical objects and results, such as relative resistance $c_{ij}\omega_{ij}$, resistance curvature $\mathbf{p}$, resistance radius σ^2 and resistive embeddings φ. We briefly explain the meaning of each name when they first appear in the text. Importantly, these names are introduced to improve the exposition and readability of this thesis but we do not necessarily advocate for them to become standard. While we believe that the objects and results with clearly distinguished properties merit a specific name, it might well be that some of their properties (which are yet to be discovered) suggest a different and more natural name. For instance, the relation of $\mathbf{p}$ to potential theory through the equilibrium characterization $\Omega\mathbf{p} = 2\sigma^2\mathbf{u}$ might lead to a more natural name.

While unwieldy long, an alternative name for the Fiedler–Bapat identity (Theorem 3.6) could have been the 'Fiedler–Bapat-Subak-Sharpe identity' to recognize the contributions of Subak-Sharpe to this identity and its consequences, including (i)

the Fiedler–Bapat identity was discovered independently by Subak-Sharpe and collaborators, see for instance [228], (ii) together with Moore, Subak-Sharpe discovered the simplex characterization of resistance matrices in [179, 213] and (iii) in much of his work, Subak-Sharpe advocated for the use of the resistance matrix and associated concepts (our $\mathbf{p}$ and σ^2) in electrical circuit theory [228–230] (even though to a certain degree misguided [258]). Yet other names might relate to the content of or the ideas behind the identity, and we may converge on a suitable name as the theory continues developing.

Chapter 2

Graphs, Laplacians and the effective resistance

This chapter introduces the mathematical objects that will be used in the rest of this thesis: graphs as the main structures of interest, Laplacians as their matrix representations and finally the effective resistance. We introduce the relevant properties of Laplacian matrices and discuss some general aspects of the effective resistance such as its metric property and the relation to spanning trees.

2.1 Weighted graphs

The main mathematical objects of interest in this thesis are graphs. A graph G consists of a set $\mathcal{N}$ of n nodes and a set $\mathcal{L}$ of m links, which consist of pairs of nodes[1]. We write $(i, j) \in \mathcal{L}$ for a link and $i \sim j$ if two nodes are linked. Without loss of generality, we can label the nodes of the graph by $1, 2, \ldots, n$ and use both $i \in \mathcal{N}$ or $1 \leq i \leq n$ as the index of a node. A subset of nodes will be denoted by $\mathcal{V}$ and its complement by $\mathcal{V}^c := \mathcal{N} \backslash \mathcal{V}$. In general, graphs may be directed and have self-

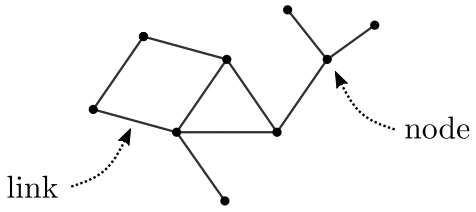


Figure 2.1: A connected graph with $n = 9$ nodes and $m = 10$ links.

[1]The standard terminology for nodes and links is 'vertices' and 'edges'. To avoid confusion with the geometric notion of vertices and edges in Chapter 4 and after, we use the terminology nodes and links instead.

loops or multilinks but here we will only consider graphs which are undirected (with $(i,j) \equiv (j,i)$) and without self-loops ((i,i) is not a link) or multilinks ($\mathcal{L}$ is a set).

A *weighted graph* is a graph together with a positive function $c : \mathcal{L} \to (0,\infty)$ on the links. For a link (i,j), the function value will be denoted by c_{ij} and is called the *weight* of the link. The weight of a link is used to represent or model the strength, affinity or similarity between its end nodes[2]. While most results in this thesis are derived for weighted graphs, they can be translated to unweighted graphs by choosing a constant weight function $c = 1$. In summary, a weighted graph is determined by the data $G = (\mathcal{N}, \mathcal{L}, c)$.

We will make use of a number of substructures in graphs. A *path* P is determined by an ordered set of nodes $i_1, i_2, \ldots, i_\ell$ such that all nodes are distinct and all consecutive nodes are linked as $i_k \sim i_{k+1}$ for $1 \leq k < \ell$; a path with $i_1 = i$ and $i_\ell = j$ is also called an $i - j$ path. A *connected component* (or simply, component) of G is a subgraph in which there exists a path between every pair of nodes; these components partition the set of nodes and the number of connected components is denoted by $\beta(G)$. Unless stated otherwise, we will generally assume that G *is a connected graph*, with $\beta(G) = 1$. A *cycle* is determined by an ordered set of vertices $i_1, i_2 \ldots, i_\ell$ such that all nodes are distinct and all consecutive nodes are linked as $i_k \sim i_{k+1}$ for $1 \leq k < \ell$ and $i_\ell \sim i_1$.

In some cases, a graph is 'almost' disconnected in the sense that there is a single node or link whose removal can disconnect the graph; these are called *cut nodes* and *cut links*, respectively. A connected graph without cut nodes is called *biconnected*[3] and a maximal subgraph without cut nodes is a biconnected component. More generally, a *cut* $C \subseteq \mathcal{L}$ is a set of links whose removal disconnects the graph[4] (or the component in which they are contained). Some examples of these substructures are shown in Figure 2.2.

For a node i, the nodes j with $j \sim i$ are called the *neighbours of* i and the number of neighbours is called the *combinatorial degree*, denoted by d_i. The total link weight

[2]While we exclude these limits from our definition of weighted graphs, a graph with $c_{ij} \to 0$ may intuitively be thought of as a graph with (i,j) removed while a graph with $c_{ij} \to \infty$ would correspond to a graph with i and j merged (in the sense introduced in Section 3.3.2).

[3]A biconnected graph is also called 2-connected.

[4]Often, the name 'cut' is used for minimal sets of links that disconnect a graph. These minimal cuts correspond to the links between a partition of the node set.

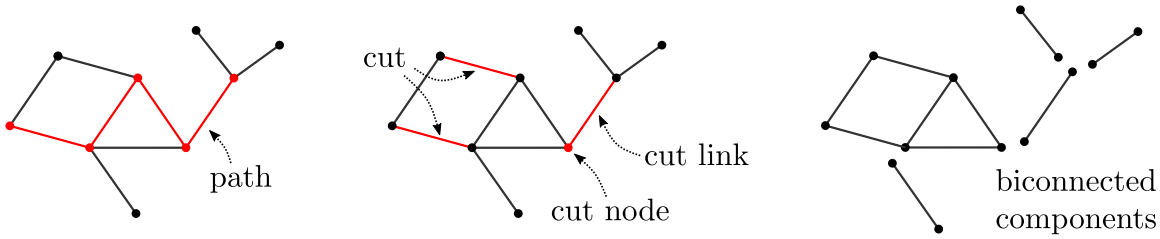


Figure 2.2: Some examples of substructures in a graph. From left to right: a path in a graph; a cut, a cut node and a cut link; the decomposition of the graph into its biconnected components.

of all links incident on a node is called the *degree*[5] and is denoted by

$$k_i := \sum_{j:j\sim i} c_{ij} \,.$$

The average degree is defined as $\langle k \rangle := \frac{1}{n} \sum_{i \in \mathcal{N}} k_i$. Some standard reference works on graph theory are [28, 77, 123, 242] and for graph theory in the context of applications, we refer to [184].

We now discuss a number of examples of graphs to illustrate the definitions and for further reference.

Example 2.1 (**tree graph**)**.** A *tree graph* T is a connected graph without cycles and with, as a result, $m = n - 1$ links. All links in a graph are cut links and all nodes are either leaf nodes (with combinatorial degree $d = 1$) or cut nodes. A *path graph* is a tree graph with exactly two leaf nodes.

Example 2.2 (**complete graph**)**.** The complete graph K_n is the graph on n nodes with a link between every pair of (distinct) nodes and thus $m = n(n-1)/2$ links. The complete graph has the most links and cycles of any graph on a given number of nodes. The graph K_2 is a two-node path and K_1 is a single node, also called the trivial graph. If a subset of nodes in a graph are fully connected, this is also called a *clique*; conversely a subset of nodes without any links between them is an *independent set*.

Example 2.3 (**cycle graph**)**.** A cycle graph consists of a single cycle. The nodes of a cycle graph can be labelled as $1, \ldots, n$ with links $i \sim i+1$ and $n \sim 1$.

[5] We choose to use the shorter name 'degree' for k_i since this will occur more often in this thesis, while d_i will be less relevant and we thus give it the longer name 'combinatorial degree'. The degree k_i is sometimes also called 'strength' or 'weighted degree'.

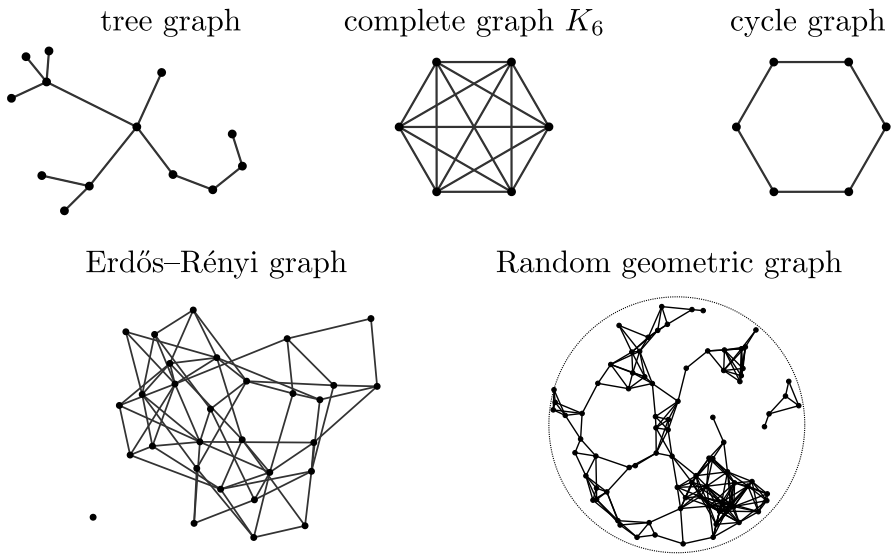


Figure 2.3: Some examples of graphs. The Erdős–Rényi random graph is constructed with $n = 30$ nodes and link probability $\rho = 0.16$ and the random geometric graph is constructed by sampling points from a disc.

Example 2.4 (**random graphs**)**.** A random graph is a graph that is constructed based on some probabilistic rule. For instance, an Erdős–Rényi graph (ER) is defined by fixing a set of n nodes and connecting each pair of nodes with some probability $0 \leq \rho \leq 1$ independently[6]; see for instance [29, 245]. Another example are random geometric graphs (RGGs) [195]. Here, the nodes correspond to randomly sampled points from some metric space M (for instance $M = \mathbb{R}^2$) and links are established between pairs of nodes whose distance (in M) is below a certain threshold. As shown for instance in Chapter 6, RGGs often inherit certain properties from the ambient space.

Figure 2.3 illustrates the different examples.

2.2 The Laplacian matrix

Among the numerous ways to study graphs, the approach in this thesis is to associate algebraic (matrices, vectors and scalar invariants) and geometric (metric spaces, simplices) objects to a graph, and use tools and ideas from linear algebra and geometry to study its properties. Many of these algebraic and geometric objects will be defined based on the Laplacian matrix of a graph.

[6]More precisely, this is the $G(n, \rho)$ ER model. Other ER models exist; for instance, $G(n, m)$ is a graph with n nodes and m links selected uniformly at random from all graphs with n nodes and m links.

An $n \times n$ matrix A can be related to a graph by associating each row and column with a node (i.e. by the labels $1 \leq i \leq n$). In this way, the entry $(A)_{ij}$ of the matrix is associated to the node pair $i, j \in \mathcal{N}$. See Appendix A.1 for more on matrix theory in the context of graphs. The *Laplacian matrix*[7] of a graph Q is defined as follows:

$$(Q)_{ij} := \begin{cases} -c_{ij} & \text{if } i \sim j \\ k_i & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Figure 2.4 below shows a graph and its associated Laplacian matrix.

$$\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{array} \begin{pmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 3 & -1 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 4 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 4 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -2 & 4 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

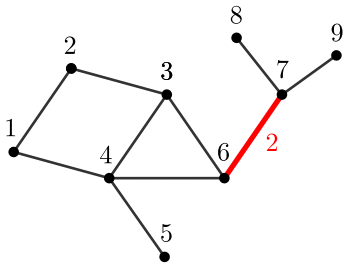


Figure 2.4: The Laplacian matrix of a graph. The highlighted link has weight $c_{67} = 2$ and thus $(Q)_{67} = -2$.

The diagonal entries of the Laplacian are given by the degrees of the nodes and the off-diagonal entries are determined by the links and their weights; clearly, the graph data $(\mathcal{N}, \mathcal{L}, c)$ can be reconstructed from the Laplacian matrix so specifying a graph G or a Laplacian Q is equivalent. The Laplacian matrix as a representation of graphs dates back to the work of Kirchhoff in electrical circuits [143, 144] and it is also called the admittance matrix, the combinatorial Laplacian or the Kirchhoff Laplacian. Some standard references on Laplacian matrices are [11, 38, 55, 58, 175, 177] and many books on (algebraic) graph theory discuss Laplacians [28, 107].

A number of properties follow immediately from definition (2.1): the Laplacian matrix Q is a symmetric matrix with nonpositive off-diagonal entries and with zero row (and column) sums, since

$$\sum_{j=1}^{n} (Q)_{ij} = k_i - \sum_{j:j\sim i} c_{ij} = k_i - k_i = 0 \text{ for all } i \in \mathcal{N}.$$

[7] There are different types of 'Laplacian matrices' associated to a graph such as the normalized Laplacian, the random-walk Laplacian and the combinatorial Laplacian. The matrix Q defined by equation (2.1) is usually known as the *combinatorial Laplacian* or *Kirchhoff Laplacian*. Since this is the only type of Laplacian that we will use in this thesis, we call Q "the Laplacian" of a graph.

In other words, $Q\mathbf{u} = 0$ or $\mathbf{u} \in \ker(Q)$, where $\mathbf{u} = (1, \ldots, 1)^T$ is the all-one vector. The connectedness of a graph furthermore implies that Q is an *irreducible* matrix — there exists no permutation P of the rows/columns such that PQP^T is a block diagonal matrix. From any matrix with these properties, a graph can be constructed with nodes corresponding to row indices and links corresponding to nonzero off-diagonal entries, which means that the above properties — *symmetry, nonpositive off-diagonal, zero row sum and irreducibility* — are necessary and sufficient conditions for a matrix to be a Laplacian.

Definition (2.1) furthermore implies the following expression for the Laplacian quadratic form (also called the *Dirichlet energy* [55, 58]):

$$\begin{aligned}
\mathbf{f}^T Q \mathbf{f} &= \sum_{i \in \mathcal{N}} f_i (k_i f_i - \sum_{j : j \sim i} c_{ij} f_j) \quad \text{(by definition of } Q\text{)} \\
&= \sum_{i \in \mathcal{N}} f_i \sum_{j : j \sim i} c_{ij} (f_i - f_j) \quad \text{(by definition of } k_i\text{)} \\
&= \sum_{i \sim j} c_{ij} \left[f_i (f_i - f_j) + f_j (f_j - f_i) \right] \\
&= \sum_{i \sim j} c_{ij} (f_i - f_j)^2. \qquad (2.2)
\end{aligned}$$

In other words, the quadratic form determined by the Laplacian matrix can be decomposed as a sum over the links. By nonnegativity of each term $c_{ij}(f_i - f_j)^2$, this quadratic form is nonnegative and can only be zero if $f_i = f_j$ for all $i \sim j$. By connectedness of the graph, this implies[8] that zero only occurs if $\mathbf{f}$ is a constant vector $\mathbf{f} \in \text{span}(\mathbf{u})$, or in other words

$$\mathbf{f}^T Q \mathbf{f} \geq 0 \text{ with equality if and only if } \mathbf{f} \in \text{span}(\mathbf{u}). \qquad (2.3)$$

This bound leads to a concise characterization of Laplacian matrices:

Proposition 2.5. *A symmetric matrix Q is the Laplacian matrix of a connected graph if and only if it has nonpositive off-diagonal entries $(Q)_{ij} \leq 0$ and kernel $\ker(Q) = \text{span}(\mathbf{u})$.*

Proof. *(forward direction)* Let Q be a Laplacian matrix. By definition, the off-diagonal entries of Q are nonpositive. Second, since the row sum of a Laplacian is zero we have $\text{span}(\mathbf{u}) \subseteq \ker(Q)$. Suppose for contradiction that this subset is strict.

[8] Take any two nodes i and j. By connectedness there is a path $i = i_1 \sim \cdots \sim i_\ell = j$ between them and since $f(i_k) = f(i_{k+1})$ for all links $1 \leq k < \ell$, we obtain $f_i = f_j$.

Then there must exist $\mathbf{f} \notin \text{span}(\mathbf{u})$ with $\mathbf{f} \in \ker(Q)$ and thus $\mathbf{f}^T Q \mathbf{f} = 0$ but this is in contradiction with (2.3) and we thus know that $\ker(Q) = \text{span}(\mathbf{u})$ as required.

(converse direction) Let Q be a symmetric matrix with nonpositive off-diagonal entries $(Q)_{ij} \leq 0$ and kernel $\ker(Q) = \text{span}(\mathbf{u})$; we write $i \sim j$ if $(Q)_{ij} \neq 0$. By definition of the kernel, we have $Q\mathbf{u} = 0$ and thus $(Q)_{ii} = -\sum_{j:j\sim i}(Q)_{ij}$, which means that Q satisfies the definition of a Laplacian matrix — we may construct a graph with nodes given by the row indices of Q and with links and their weights determined by the negative off-diagonal entries of Q — but it remains to prove that this graph is connected, i.e. that Q is irreducible. Suppose for contradiction that Q is reducible. Then there exists a partition of the nodes (row indices) into subsets $\mathcal{V}_1, \ldots, \mathcal{V}_\ell$ with $\ell > 1$ such that A is block-diagonal on these subsets and with $(Q)_{\mathcal{V}_k\mathcal{V}_k}$ irreducible and thus the Laplacian of a connected graph, for each partition $1 \leq k \leq \ell$. Thus for the vector $\mathbf{f}$ with entries $f_i = 1$ if $i \in \mathcal{V}_1$ and zero otherwise, we then find

$$Q\mathbf{f} = Q_{\mathcal{V}_1\mathcal{V}_1}\mathbf{u} = 0$$

since $Q_{\mathcal{V}_1\mathcal{V}_1}$ is a Laplacian. However, this is a contradiction since $Q\mathbf{f} = 0$ but $\mathbf{f} \notin \text{span}(\mathbf{u}) = \ker(Q)$. Thus Q must be irreducible and the Laplacian matrix of a connected graph as required. □

We may thus talk about Laplacian matrices without explicit reference to some underlying graph. The Laplacian matrix of a general graph (with $\beta(G) \geq 1$) is characterized similarly by nonpositive off-diagonal entries and a kernel spanned by piecewise constant vectors on a partition of the nodes (row indices). We note that the matrix $(I - \frac{\mathbf{u}\mathbf{u}^T}{n})$, where I is the identity matrix, is a projection matrix onto $\text{span}(\mathbf{u})^\perp = \ker(Q)^\perp$ and will appear frequently throughout this thesis.

Other matrix characterizations are that the Laplacian is a Z matrix (a matrix with nonpositive off-diagonal entries) and an M matrix (a positive semidefinite Z matrix, see below) [21]. Succinctly, *a Laplacian is a symmetric Z matrix with constant kernel.*

The Laplacian matrix is singular and cannot be inverted. Instead, we may consider the *Moore–Penrose pseudoinverse* of Q which is defined as the unique matrix $Q^\dagger$ that satisfies [196]

$$QQ^\dagger Q = Q, \quad Q^\dagger Q Q^\dagger = Q^\dagger \quad \text{and} \quad QQ^\dagger, Q^\dagger Q \text{ are symmetric.}$$

See Appendix A.3 for further results on the Moore–Penrose pseudoinverse. The Moore–Penrose pseudoinverse of a matrix always exists and is unique; we will refer to $Q^\dagger$ as *the pseudoinverse Laplacian* of a graph. From its definition, the pseudoinverse

Laplacian is a real symmetric matrix and, by $Q^\dagger = Q^\dagger Q Q^\dagger$, it shares the following important property with the Laplacian matrix

$$\mathbf{f}^T Q^\dagger \mathbf{f} \geq 0 \text{ with equality if and only if } \mathbf{f} \in \text{span}(\mathbf{u}), \tag{2.4}$$

and thus also $\ker(Q^\dagger) = \text{span}(\mathbf{u})$. The product between a Laplacian and its pseudoinverse satisfies the following property from which $Q^\dagger$ derives its name as a 'pseudoinverse' (see Appendix A.3):

Proposition 2.6. *The product of a Laplacian matrix and its Moore–Penrose pseudoinverse is the projection matrix onto* $\ker(Q)^\perp = \ker(Q^\dagger)^\perp = \text{span}(\mathbf{u})^\perp$*:*

$$QQ^\dagger = Q^\dagger Q = I - \frac{\mathbf{u}\mathbf{u}^T}{n}. \tag{2.5}$$

In other words, $Q^\dagger$ is the inverse of Q in the space $\ker(Q)^\perp$. We give one further useful expression for the Laplacian pseudoinverse in terms of a matrix inverse (see [104], or using (2.7) below):

$$Q^\dagger = \left(Q + \frac{\mathbf{u}\mathbf{u}^T}{n}\right)^{-1} - \frac{\mathbf{u}\mathbf{u}^T}{n}. \tag{2.6}$$

2.2.1 The Laplacian spectrum

The Laplacian matrix is one of the central objects in spectral graph theory, the subfield of graph theory concerned with the relation between spectral properties of various matrices associated with a graph[9] and graph-theoretic properties [38, 55, 65, 120, 175, 177, 247]. In the case of Laplacian matrices, one is interested in solutions to the eigenequation $Q\mathbf{z}_k = \mu_k \mathbf{z}_k$, where a solution μ_k is called an *eigenvalue* with corresponding *eigenvector* $\mathbf{z}_k \in \mathbb{R}^n$. As explained in Appendix A.2, the solutions to the eigenequation for symmetric matrices satisfy some special properties: the eigenvalues are real and there exists a basis of $\mathbb{R}^n$ consisting of eigenvectors, where eigenvectors corresponding to different eigenvalues are orthogonal[10]. As a result, we may write the following *eigendecomposition*[11] for the Laplacian matrix (see e.g. [130]):

$$Q = \sum_{k=1}^{n} \mu_k \mathbf{z}_k \mathbf{z}_k^T \text{ with } \mu_k \in \mathbb{R} \text{ and } \{\mathbf{z}_k\}_{k=1}^n \text{ an orthonormal basis for } \mathbb{R}^n,$$

[9] Aside from the Laplacian matrix, many results in spectral graph theory deal with the adjacency matrix; this is the $n \times n$ matrix A with entries $(A)_{ij} = 1$ if $i \sim j$ and zero otherwise. In this thesis, we do not study the adjacency matrix.

[10] In general two eigenvectors corresponding to the same eigenvalue may thus be non-orthogonal, but we will require them to be orthogonal such that the eigenvectors form an orthogonal basis.

[11] Often this eigendecomposition is written in matrix form as $Q = ZMZ^T$ with $Z = [\mathbf{z}_1 \ \dots \ \mathbf{z}_n]$ and $M = \text{diag}(\mu_1, \dots, \mu_n)$.

with the further choice of normalized eigenvectors ($\|\mathbf{z}_k\| = 1$) and eigenvectors corresponding to equal eigenvalues taken orthogonal. While the basis of eigenvectors is not necessarily unique, the multiset of eigenvalues $\{\mu_1, \ldots, \mu_n\}$ (with possible repeats) is unique and is called the *spectrum* of the Laplacian matrix[12].

Many results in spectral graph theory deal with how the Laplacian spectrum of a graph relates to its structural or graph-theoretic properties. We will not use many results from this rich field of study, but we may translate some of the results from the previous subsection in terms of the spectrum of Q. The quadratic form inequality (2.3) is equivalent to (see also [55]):

Property 2.7. *The Laplacian matrix is positive semidefinite (all eigenvalues are nonnegative) with a unique zero eigenvalue with corresponding constant eigenvector.*

Proof. Let Q be a Laplacian matrix with $\mathbf{z}_1, \ldots, \mathbf{z}_n$ an orthonormal basis of eigenvectors. Since $\ker(Q) = \operatorname{span}(\mathbf{u})$ we know that $Q\mathbf{z}_k = 0$ if and only if $\mathbf{z}_k \in \operatorname{span}(\mathbf{u})$ and thus 0 is a unique zero eigenvalue (of multiplicity 1) with corresponding constant eigenvector $\mathbf{z}_n$. Furthermore, by definition of the eigenvalues and (normalized) eigenvectors, we find that

$$\mu_k = \mu_k \mathbf{z}_k^T \mathbf{z}_k = \mathbf{z}_k^T Q \mathbf{z}_k \geq 0$$

and thus $\mu_k > 0$ as required for all $k \neq n$. □

We stress that this result is for Laplacian matrices as defined by Proposition 2.5 in correspondence with connected graphs; for disconnected graphs, the multiplicity of the zero eigenvalue is equal to the number of connected components $\beta(G)$. Following Property 2.7, the Laplacian eigenvalues can thus be ordered as $0 = \mu_n < \mu_{n-1} \leq \cdots \leq \mu_1$, and any corresponding basis of eigenvectors has $\mathbf{z}_n = \mathbf{u}/\sqrt{n}$ and $\mathbf{z}_k^T \mathbf{u} = 0$ for all $k \neq n$; we will call a vector $\mathbf{x}$ with $\mathbf{u}^T \mathbf{x} = 0$ a *zero-sum* vector. Following the same proof as for the Laplacian matrix, we find that the pseudoinverse Laplacian satisfies the same property:

Property 2.8. *The pseudoinverse Laplacian matrix $Q^\dagger$ is positive semidefinite (all eigenvalues are nonnegative) with a unique zero eigenvalue with corresponding constant eigenvector.*

The Laplacian and its pseudoinverse not only share the zero eigenvalue with corresponding constant eigenvector, but it can be checked that the nonzero eigenvalues of the pseudoinverse Laplacian are given by the inverse of the nonzero eigenvalues of

[12] Another definition of the spectrum is as the multiset of solutions to the equation $\det(Q - \mu I) = 0$.

the Laplacian and with corresponding eigenspaces. The pseudoinverse Laplacian can thus be found from the eigenvalues and eigenvectors of the Laplacian as follows:

$$Q^\dagger = \sum_{k=1}^{n-1} \mu_k^{-1} \mathbf{z}_k \mathbf{z}_k^T. \tag{2.7}$$

Importantly, the sum in (2.7) runs over $1 \leq k \leq n-1$ with the zero eigenvalue $\mu_n = 0$ excluded.

In this thesis, we will mainly make use of 'qualitative' results[13] on the Laplacian spectrum such as the dimension of the kernel and positive semidefiniteness of the Laplacian. Spectral graph theory often deals with more quantitative results. For example, the qualitative result 'if the second smallest Laplacian eigenvalue is positive, then the graph is connected' is made more quantitative in the Cheeger inequality (see e.g. [177]) and Fiedler's results on the algebraic connectivity [91], which relate the precise value of the second smallest Laplacian eigenvalue to how well-connected the graph is.

2.3 The effective resistance

We now arrive at the central definition of this thesis, the effective resistance:

Definition 2.9. Let G be a connected, weighted graph. The *effective resistance* ω_{ij} between a pair of nodes i, j is defined as

$$\omega_{ij} := (\mathbf{e}_i - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \text{ for all } i, j \in \mathcal{N}. \tag{2.8}$$

The vector $\mathbf{e}_i$ is the unit vector with $(\mathbf{e}_i)_k$ equal to 1 if $k = i$ and zero otherwise, and the effective resistance is thus equal to $\omega_{ij} = (Q^\dagger)_{ii} + (Q^\dagger)_{jj} - 2(Q^\dagger)_{ij}$. The effective resistance can also be seen as a function $\omega : \mathcal{N} \times \mathcal{N} \to \mathbb{R}$ (with $\omega(i,j) = \omega_{ij}$) on pairs of nodes and can be summarized in the $n \times n$ *resistance matrix* Ω with entries $(\Omega)_{ij} := \omega_{ij}$. This matrix is studied in detail in Chapter 3. For disconnected graphs, the effective resistance can be defined in the same way for pairs of nodes in the same component and with $\omega_{ij} = \infty$ for nodes in different components.

[13] One reason why the theory of effective resistances mainly relies on qualitative spectral results is, perhaps, that positivity of the link weights (off-diagonal nonnegativity of Q) plays such an important role while this sign information has no clear spectral fingerprint.

The name 'effective resistance' originates from the theory of (resistive) electrical circuits. If we force a unit current to flow from one point i in an electrical circuit to another point j, then the measured voltage difference between i and j is given by the effective resistance between these points. For two points that are electrically well-connected, such that the current experiences little resistance and easily flows from i to j without dissipating too much power, this voltage drop will be small (ω_{ij} small) whereas two isolated points will experience a much larger voltage drop (ω_{ij} large). Thus ω_{ij} is a **resistance** because resistance is defined as 'voltage/current' and it is **effective** because from the perspective of the two points, the whole circuit behaves effectively as if i and j were connected by a single resistor with resistance equal to ω_{ij} (see Figure 2.5). The units of resistance are Ohm with symbol 'Ω'.

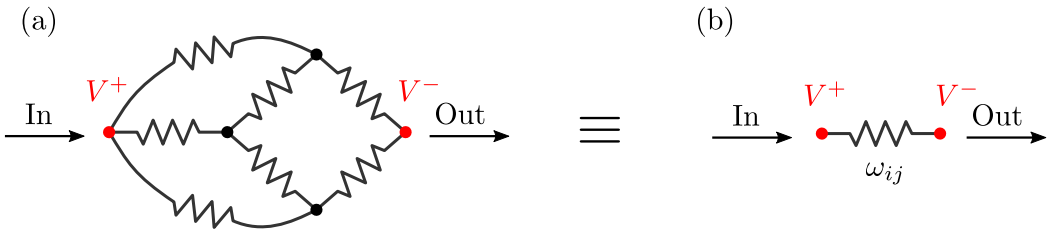


Figure 2.5: Circuit (a) is electrically equivalent to circuit (b) for electrical flows and potentials between the red nodes. In other words, the circuit between two nodes i and j can be replaced by a single resistor with the effective resistance ω_{ij} as resistance.

This definition-by-measurement of the effective resistance translates to the algebraic definition by mapping the electrical circuit onto a weighted graph: each resistor is modeled by a link and the weight of the link is given by the conductance of the resistor (conductance = 1/resistance), hence 'c' for link weights. With this translation, the linear system equations (Kirchhoff's and Ohm's laws) that relate the voltages $\mathbf{v}$ and external currents $\mathbf{x}$ at each node in the circuit are then expressed compactly using the Laplacian matrix as $Q\mathbf{v} = \mathbf{x}$; see for instance [23, 78]. A unit current from i to j is given by $\mathbf{x} = \mathbf{e}_i - \mathbf{e}_j$, and measuring the voltage difference $v_i - v_j$ can be written as $(\mathbf{e}_i - \mathbf{e}_j)^T\mathbf{v}$. Measuring the voltage difference retrieves the effective resistance as

$$\omega_{ij} = (v_i - v_j) = (\mathbf{e}_i - \mathbf{e}_j)^T\mathbf{v} = (\mathbf{e}_i - \mathbf{e}_j)^T Q^\dagger \mathbf{x} = (\mathbf{e}_i - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j).$$

The indeterminacy of the Laplacian equation in span($\mathbf{u}$) reflects that external currents must be net zero (by conservation of charge) and the fact that only differences between voltages are relevant. Importantly, there is a broad class of *passive linear systems* — for instance, thermodynamic, hydraulic and mechanical systems — whose system

equations are modeled by a Laplacian matrix [7]. In these settings the definition of effective resistance also makes sense.

2.3.1 Intuition: effective resistance and connectivity

While its definition is fairly simple — as a quadratic product with the pseudoinverse Laplacian — the interpretation of the effective resistance is not immediately obvious. In the following sections and throughout the rest of this thesis, we will encounter many different aspects of the effective resistance and develop various intuitions to understand what the resistance ω_{ij} says about the relation between the two nodes i and j in the graph. Section 2.3.4, for instance, discusses how the value $c_{ij}\omega_{ij}$ for a link (i, j) can be seen as a notion of *importance* of the link, which reflects whether the link is important or redundant for connectivity: a cut link will have maximal $c_{ij}\omega_{ij}$ while a link that is part of many triangles will tend to have small $c_{ij}\omega_{ij}$. In Section 2.3.3, we show that ω satisfies the properties of a distance (metric) on the graph.

In this section, we take a small detour to mention an additional result that may further guide the intuition. Proposition 2.10 below should not be seen as a technical result — no further theory will be based on it — but it may help in developing additional understanding of the effective resistance.

The proposed intuition is that the effective resistance can be thought of as reflecting *how well-connected two nodes are in a graph*. Here, 'connectivity' can be made more precise by using the notions of paths and cuts between i and j. A classical notion of connectivity for instance counts how many nodes or links need to be removed from a graph to disconnect two nodes. If a small cut exists, this implies a poor/small connectivity, while if many independent paths exist this implies a good/high connectivity. The effective resistance reflects a more integrated notion of connectivity, taking into account link weights and multiple sets of cuts and paths.

An $i-j$ path P in a graph is a path with end nodes i and j, and two paths P and P' are independent if they only share their end nodes. An $i-j$ cut C is a cut whose removal disconnects i from j, and two cuts C and C' are independent if they share no links. We recall that the weight c of a link should be thought of as an affinity. The weight of a path and a cut are defined as follows:

$$c(P)^{-1} := \sum_{(a,b)\in P} c_{ab}^{-1} \quad \text{and} \quad c(C) := \sum_{(a,b)\in C} c_{ab}.$$

In other words, a short path (with few links and/or large weights) will have a large weight, and a small cut (with few links and/or small weights) will have a small weight, and conversely. Intuitively, a short path between i and j adds to their connectivity while a small cut can be thought of as a bottleneck between i and j. The effective resistance can be bounded in terms of independent paths and cuts as follows:

Proposition 2.10. *The effective resistance between a pair of nodes i and j satisfies*

$$\sum_{C\in\mathcal{C}} c(C)^{-1} \leq \omega_{ij} \leq \Big(\sum_{P\in\mathcal{P}} c(P)\Big)^{-1}$$

for any set of independent $i-j$ paths $\mathcal{P}$ and independent $i-j$ cuts $\mathcal{C}$.

Proof. The lower-bound is known as the Nash-Williams criterion; see [169, §2.5],[224]. The upper-bound follows from the Rayleigh monotonicity principle, which says that removing any link from a graph will either leave the effective resistances unchanged or increase the effective resistances [169, §2.4]. The upper-bound then follows by constructing the new graph G' with all links removed except those in a path in $\mathcal{P}$. The effective resistance ω'_{ij} in this graph can be calculated as $\omega'^{-1}_{ij} = \sum_{P\in\mathcal{P}} c(P)$, for instance using the Kron reduction defined in Chapter 3. □

Proposition 2.10 thus says that the connectivity between two nodes is at least as good as can be expected from the independent paths between them (upper bound) and at most as good as can be expected from the independent cuts between them (lower bound). Figure 2.6 shows an example of the path and cut bounds for the effective resistance.

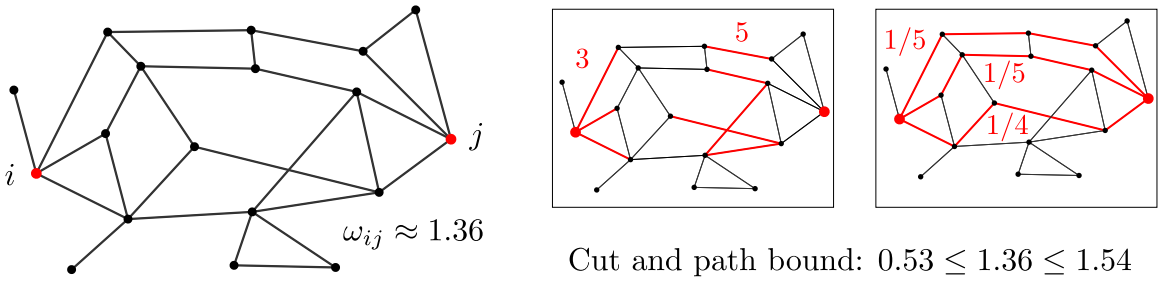


Figure 2.6: Example of the cut and path bound in a graph. The cuts and paths and their weights are indicated in red on the right.

2.3.2 Computational aspects

The definition of the effective resistance is constructive: for a given graph, one can formulate the Laplacian matrix, then calculate its pseudoinverse — for instance based

on formula (2.6) — and then determine the effective resistance as $\omega_{ij} = (Q^\dagger)_{ii} + (Q^\dagger)_{jj} - 2(Q^\dagger)_{ij}$. For example, the numerical experiments in this thesis were calculated in this way for graphs of up to $n = 5000$ nodes in a few minutes, using MATLAB on a computer with a 6-core Intel Core i9-8950HK processor with 32GB RAM.

For very large graphs, calculating the effective resistance straight from its definition becomes impractical since matrix inversion requires cubic time $O(n^3)$ in the worst case (recall that n is the number of nodes). Fortunately, there exist much more efficient solutions if an approximate solution for ω_{ij} suffices. Since computational problems involving Laplacian matrices are ubiquitous, for instance in numerical solutions of partial differential equations and applications in network science, there has been a sustained research effort to understand the computational complexity of these problems and develop efficient and practical algorithms [3, 53, 74, 194, 220, 221, 235, 249]. A breakthrough in this context was Spielman & Teng's result [220, 222] that an approximate solution $\tilde{\mathbf{x}}$ for the Laplacian problem $Q\mathbf{x} = \mathbf{b}$ can be calculated in near-linear time[14] in the number of links m. In the worst case of dense graphs with $m = O(n^2)$, this still takes quadratic time in n, but in the case of sparse graphs[15] with $m = O(n)$ it reduces to linear time in the number of nodes. These near-linear time Laplacian solvers can be used to approximate the effective resistance efficiently: if we let $\mathbf{b} = \mathbf{e}_i - \mathbf{e}_j$ and solve the Laplacian system $Q\mathbf{x} = \mathbf{b}$, we obtain an approximation for the effective resistance as $\tilde{\omega}_{ij} := (\mathbf{e}_i - \mathbf{e}_j)^T \tilde{\mathbf{x}}$.

Instead of solving a Laplacian system for each effective resistance separately, there exist tailor-made algorithms for approximating effective resistances. Chu et al. considered calculating the effective resistance of all links [53], and [219] built a data structure in near-linear time from which individual approximate effective resistances can then be calculated in $O(\log(n))$. The effective resistances can also be approximated by considering the graph locally, and [194] describes several local algorithms for approximating the effective resistance; these local algorithms are based on random walks, spanning trees or approximations of the pseudoinverse. Finally, in certain classes of graphs, the diagonal of the pseudoinverse Laplacian can be calculated in near-linear time in m [3], which is relevant for instance for calculating the resistance curvature (see Theorem 3.6 and Chapter 6).

[14] For precise statements on the time complexity, which includes additional factors in $\log(n)$ and the approximation error ϵ, we refer to the paper [220]. For more recent results, see for instance [53].

[15] We note that the term 'sparse graphs' is also used for graphs with a subquadratic number of links.

2.3.3 Resistance distance

A distance describes how far or close two things are and is often measured in terms of time, physical distance or more abstract units. Mathematically, distances are modeled by metrics and metric spaces:

Definition 2.11. A *metric space* (X, d) consists of a set X and a function $d : X \times X \to \mathbb{R}$ such that d satisfies for all $x, y, z \in X$:

$$\begin{aligned} &\text{(i)} \ d(x,y) = d(y,x), \\ &\text{(ii)} \ d(x,y) \geq 0, \\ &\text{(iii)} \ d(x,y) = 0 \Leftrightarrow x = y, \\ &\text{(iv)} \ d(x,y) + d(y,z) \geq d(x,z)\,. \end{aligned}$$

Condition (i) is a symmetry condition, conditions (ii) and (iii) say that a metric is nonnegative between two points and positive if they are distinct. The fourth condition (iv) is called the *triangle inequality* and reflects the property that the direct distance $d(x, z)$ between two points in X is never bigger than the distance $d(x, y)+d(y, z)$ when first passing through a third point. A function that satisfies these four conditions is called *a metric*. For example, by drawing a number of points on a piece of paper or a balloon, you can verify that the distances between the points indeed satisfy the conditions of a metric.

We will write $n = |X|$ for the number of elements of the metric space and generally assume a metric space to be finite. A metric space can be summarized by its distance matrix D, the $n \times n$ matrix with entries $(D)_{ij} := d(i, j)$ which is symmetric and zero-diagonal by definition. A function $\phi : X \to X'$ between two metric spaces (X, d) and (X', d') is called an *isometric embedding* if $d'(\phi(i), \phi(j)) = d(i, j)$ for all $i, j \in X$.

In the case of graphs, a metric $d : \mathcal{N} \times \mathcal{N} \to [0, \infty)$ reflects how 'far' two nodes are in the graph. The most commonly used distance in graph theory is the *shortest-path distance*, which counts the smallest number of links $|P|$ of any $i-j$ path P, or the smallest inverse path weight $1/c(P)$ in the case of weighted graphs. However, many other distances exist that reflect different notions of distance between the nodes; see for instance [39, 49, 75, 102, 103, 119, 149] and the examples in Chapter 4.

Another practical class of graph metrics follows from embeddings. Many data science and machine learning applications on graphs involve some form of *graph embedding* [35, 43, 121, 152, 250] where the nodes of a graph are embedded into some

other space $\phi : \mathcal{N} \to M$. These can for instance be embeddings 'by design' [57, 201] or learned, data-driven embeddings [116, 197, 253]. The downstream tasks such as clustering, classification or prediction are then performed on the embedded points $\phi(1), \ldots, \phi(n)$ with the advantage that M is typically a 'nicer' space to work in, often Euclidean space but also more general manifolds [36]. If the space M is a metric space (which is usually the case), then the graph has an induced metric $d := d_M \circ \phi$ as $d(i,j) = d_M(\phi(i), \phi(j))$, where d_M is the metric on M. In Chapters 4 and 5, we discuss how some very natural choices of M such as Euclidean, hyperbolic and spherical spaces result in metrics with desirable properties.

Metrics are relevant in this thesis because the effective resistance is a metric; for this reason, it is also called the *resistance distance*. Chapter 4 discusses properties of the resulting metric space $(\mathcal{N}, \omega)$ and introduces two characterizations: geometrically in terms of simplices and algebraically in terms of properties of the resistance matrix.

The resistance distance has a number of advantages and disadvantages when used in the context of applications or as a theoretical tool in the study of graphs. A first advantage is that ω is a robust measure of distance, where a small change of link weights dc will result in a small change of effective resistances $d\omega$ as a result of Rayleigh's monotonicity law; see for instance [79]. Other metrics such as the shortest-path distance may change discontinuously. Second, the rich theory of effective resistances means that any resistance-based measure or invariant is often amenable to a deep theoretical analysis; this is illustrated in Chapter 5 for the introduced graph variance. Third, the effective resistance can be approximated efficiently and locally in the common setting of sparse graphs, which means that it is available to use in practice. A first disadvantage is that the effective resistance is less intuitive to work with at first compared to other distances such as the shortest-path distance. Furthermore, the effective resistance is a non-local measure whose exact value depends on the whole graph in general (even though local approximations exist). This may be a disadvantage, for instance in the case of resistance curvature where a local measure would be desirable (see Chapter 6). Finally, a strong point of critique was formulated by von Luxburg et al. [251, 252], who showed that in certain classes of random graphs, the effective resistance between two nodes only depends on the degrees of these nodes in the first order $\omega_{ij} \approx d_i^{-1} + d_j^{-1}$. In these cases, the effective resistance becomes essentially independent of the precise graph structure and thus cannot provide a sensible measure of distance between nodes.

2.3.4 Relative resistance and spanning trees

The effective resistance is defined between any pair of nodes i, j in a weighted graph and, as noted, can be interpreted as a notion of connectivity and a distance between pairs of nodes. There is a second closely related definition with different properties:

Definition 2.12. The *relative resistance* of a link (i, j) is defined as $c_{ij}\omega_{ij}$.

In other words, the relative resistance is defined for each link in the graph (not between pairs of nodes) and is determined by the weight of the link times the effective resistance between its end nodes. Since the link weight c has the unit 'conductance = 1/resistance' in the context of electrical circuits, the relative resistance is thus equal to 'effective resistance divided by link resistance' which inspired our choice for the name "relative resistance".

The relative resistance is usually not considered as a separate quantity from the effective resistance — perhaps because often $c = 1$ — but we believe that its properties are distinct enough to merit an independent treatment.

The main results on relative resistances in this thesis will follow from their relation to random spanning trees of a graph. We recall that a *tree* is a connected graph with no cycles and thus $m = n - 1$ links; see [28] for basic properties of trees. A *spanning tree* T of a graph G is a connected subgraph on all the nodes which is a tree — i.e. T is a graph on the nodes of G with a subset of $n - 1$ links of G that do not form any cycles. The set of all spanning trees is denoted by $\mathcal{T}$ and we define a

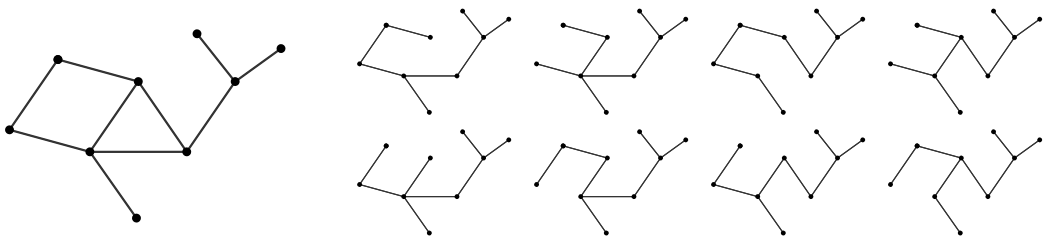


Figure 2.7: A graph (left) and its spanning trees (right). Note that the cut links appear in each spanning tree.

random spanning tree $\mathbf{T}$ as a random element (tree) in this set, with probability to equal any tree T given by

$$\Pr[\mathbf{T} = T] = \frac{\tau(T)}{\sum_{T' \in \mathcal{T}} \tau(T')} \quad \text{where} \quad \tau(T) = \prod_{(i,j) \in T} c_{ij}.$$

In other words, the probability that $\mathbf{T}$ is equal to some tree T is proportional to the product of the link weights of that tree. For unweighted graphs ($c = 1$) each tree is sampled with equal probability given by $1/(\#\text{ spanning trees})$. The relation[16] between the relative resistance and spanning trees is then given as follows:

Theorem 2.13. *The relative resistance of a link equals the probability that this link is contained in a random spanning tree, as* $c_{ij}\omega_{ij} = \Pr[(i,j) \in \mathbf{T}]$.

For a proof of Theorem 2.13 and more on spanning trees, see for instance [169, Ch. 4.2]. Theorem 2.13 is generalized by the transfer–current Theorem of Burton & Pemantle in [41], which in turn is part of the theory of determinantal point processes [167]. The relation between electrical currents and spanning trees dates back to Kirchhoff's work on electrical circuits in the 19th century [143].

Following Theorem 2.13, we can think of the relative resistance as a notion of 'link importance': if the relative resistance is large, this means that the link (i,j) appears in many spanning trees and thus that it is an important link, on average, for the connectivity between i and j. In the most extreme case, $\Pr[(i,j) \in \mathbf{T}] = 1$, the link is part of every spanning tree and must thus be a cut link (see also Proposition 2.16), which corresponds to a highly important link for the connectivity of i and j. Instead, whenever a link (i,j) is part of a triangle (i,j,k), the probability that (i,j) is part of a random spanning tree decreases since the three nodes can be connected locally by including the links (i,k) and (k,j) without (i,j). Thus the more triangles a link (i,j) is part of, the smaller its relative resistance will generally be[17]; this corresponds to (i,j) being less important for the connectivity of i and j. The link importance interpretation is further supported by the theory of graph sparsification, where sampling links from a graph with probability proportional to their relative resistances yields statistically representative sparse graph samples [221]; this is related to the concept of statistical load/leverage in numerical linear algebra [82].

We now show a number of results that follow from Theorem 2.13.

Property 2.14. *The relative resistance of a link satisfies* $0 < c_{ij}\omega_{ij} \leq 1$ *with equality in the upper bound if and only if the link is a cut link.*

Proof. By Theorem 2.13, the relative resistance of a link is a probability and thus $0 \leq c_{ij}\omega_{ij} \leq 1$.

[16] Alternatively, this relation could be taken as a definition of relative resistances with the advantage that it suggests possible routes for generalization, for instance to matroids [167].

[17] This of course also depends on the weights of the links in the triangles.

(lower bound strictness) Because the link weight c_{ij} is positive and $\omega_{ij} = (\mathbf{e}_i - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) > 0$ since $(\mathbf{e}_i - \mathbf{e}_j) \in \text{span}(\mathbf{u})^\perp$, the lower bound is strict.

(upper bound, forward) Next, we consider when equality occurs in the upper bound. If (i,j) is a cut link in G then any subgraph of G without (i,j) is disconnected. Hence, all connected subgraphs (including all spanning trees) of G must contain (i,j) and thus $c_{ij}\omega_{ij} = \Pr[(i,j) \in \mathbf{T}] = 1$. Hence, if (i,j) is a cut link then $c_{ij}\omega_{ij} = 1$.

(upper bound, converse) Conversely, let $c_{ij}\omega_{ij} = 1$ and thus $\Pr[(i,j) \in \mathbf{T}] = 1$ for some link (i,j) and suppose for contradiction that it is not a cut link. If we then consider the graph $G' = G \backslash (i,j)$ from which (i,j) has been removed, we arrive at a contradiction: since (i,j) is not a cut link, the graph G' is connected. Let T be a spanning tree of G'. Then we know that (a) T is also a spanning tree of G, and (b) T does not contain (i,j) because G' does not contain (i,j). But (a) and (b) are in contradiction with $\Pr[(i,j) \in \mathbf{T}] = 1$, which confirms that (i,j) must indeed be a cut link. This completes the proof. □

As a result, we also know that the effective resistance is bounded as follows:

$$0 < \omega_{ij} \leq c_{ij}^{-1} \text{ for all } i \sim j,$$

with equality $\omega_{ij} = c_{ij}^{-1}$ if and only if (i,j) is a cut link. Next, we find a number of results for the sum of relative resistances over subsets of links in a graph. A first classical result is *Foster's Theorem*, named after Ronald M. Foster [101].

Theorem 2.15 (Foster's Theorem)**.** *The sum over all relative resistances in a graph is equal to the number of connected components:* $\sum_{i \sim j} c_{ij}\omega_{ij} = \beta(G)$.

Proof. We show that the result holds for a connected graph ($\beta(G) = 1$). Since the relative resistances are determined independently in each connected component of a general graph, this implies Foster's Theorem in the general case.

Let G be a connected graph. Then the sum over relative resistances is equal to

$$\begin{aligned}
\sum_{(i,j)\in\mathcal{L}} c_{ij}\omega_{ij} &= \sum_{(i,j)\in\mathcal{L}} \Pr[(i,j) \in \mathbf{T}] \\
&= \sum_{(i,j)\in\mathcal{L}} \left(\sum_{T\in\mathcal{T}} \Pr[\mathbf{T} = T] \mathbf{1}_{\{(i,j)\in T\}} \right) \\
&= \sum_{T\in\mathcal{T}} \Pr[\mathbf{T} = T] \sum_{(i,j)\in\mathcal{L}} \mathbf{1}_{\{(i,j)\in T\}} \\
&= \sum_{T\in\mathcal{T}} \Pr[\mathbf{T} = T](n-1) \quad \text{(every spanning tree has } n-1 \text{ links)} \\
&= n - 1,
\end{aligned}$$

which completes the proof. □

A third result that follows from Theorem 2.13 are the following *cut bounds* for relative resistances[18]:

Proposition 2.16 (cut bounds)**.** *The relative resistances satisfy*

$$\sum_{(i,j)\in C} c_{ij}\omega_{ij} \geq 1 \text{ for any cut } C, \tag{2.9}$$

with equality if and only if C consists of a single cut link.

Proof. As in the proof of Foster's Theorem, we can derive that

$$\sum_{(i,j)\in C} c_{ij}\omega_{ij} = \sum_{T\in\mathcal{T}} \Pr[\mathbf{T} = T] \sum_{(i,j)\in C} \mathbf{1}_{\{(i,j)\in T\}}$$

for any cut C. If C is a cut in G, then any subgraph of G without links in C is disconnected. Hence, all connected subgraphs (and spanning trees) of G must contain at least one link in C. In other words $\sum_{(i,j)\in C} \mathbf{1}_{\{(i,j)\in T\}} \geq 1$ for every spanning tree T, which implies that $\sum_{(i,j)\in C} c_{ij}\omega_{ij} \geq 1$, as required. We now show the conditions for equality.

(equality, forward) If C consists of a single link, then by Property 2.14 we find $c_{ij}\omega_{ij} = 1$, as required. *(equality, converse)* Conversely, if $\sum_{(i,j)\in C} c_{ij}\omega_{ij} = 1$ then since every spanning tree must have at least one link in C (as shown above) it follows that every spanning tree must have exactly one link in C. We now show that this implies that C must consist of a single link, which is thus a cut link. If G is a tree graph then the relative resistance of each link is $c_{ij}\omega_{ij} = 1$ and the result follows; it thus remains to prove the result for non-tree graphs.

Let T be a spanning tree and let $\ell = T \cap C$ be the unique link of T in the cut C. Suppose for contradiction that $|C| > 1$. Then there exists another link $\ell' \in C$ and we can construct the graph $T + \ell'$, i.e. the tree T with link ℓ' included. This graph will have a unique cycle H (see for instance [28]) and since cycles have length at least 3, we can take a link $\ell'' \in H$ which is different from ℓ, ℓ'. But then the graph $T + \ell' - \ell''$ is a spanning tree of G (see [28]) that contains both ℓ and ℓ' in the cut C. Since this is in contradiction with every spanning tree having exactly one link in C, we know that $|C| = 1$ as required. □

While the result holds for any set of links whose removal disconnects the graph (a cut), the bound is of course most tight for 'minimal' cuts. As an example, the incident

[18] We were unable to determine whether these bounds are new, but they are implied, for instance, by the Nash-Williams inequality (i.e. the lower bound in Proposition 2.10).

links on a node are a cut and thus we find that $\sum_{j:j\sim i} c_{ij}\omega_{ij} \geq 1$ with equality if and only if i is connected by a single link. This will be relevant in Chapter 6.

We recall that a cycle in a graph is an ordered set of links $(i_1, j_1), \ldots, (i_\ell, j_\ell)$ where consecutive links have a node in common as $j_k = i_{k+1}$ for $1 \leq k < \ell$ and with $j_\ell = i_1$. The number of links in a cycle H is denoted by $|H|$. The cut bound has the following 'dual' cycle bound:

Proposition 2.17 (cycle bound)**.** *The relative resistances satisfy*

$$\sum_{(i,j)\in H} c_{ij}\omega_{ij} \leq |H| - 1 \text{ for any cycle } H.$$

Proof. As in the proof of Foster's Theorem, we can derive that

$$\sum_{(i,j)\in H} c_{ij}\omega_{ij} = \sum_{T\in\mathcal{T}} \Pr[\mathbf{T} = T] \sum_{(i,j)\in H} \mathbf{1}_{\{(i,j)\in T\}}$$

for any cycle H. If H is a cycle in G, then any subgraph of G that contains all links in H has at least one cycle. Hence, all cycle-free subgraphs (and thus spanning trees) can have at most $|H| - 1$ links in H. As a result, $\sum_{(i,j)\in H} \mathbf{1}_{\{(i,j)\in T\}} \leq |H| - 1$ for every spanning tree T, which implies the proposition. □

Chapter 3

The resistance matrix

This chapter develops the matrix theory of the resistance matrix Ω. While many interesting bounds and results may be found for individual effective or relative resistances, a whole new picture emerges when considering all effective resistances at once; this is where the resistance matrix Ω becomes relevant. The main result in this chapter is the Fiedler–Bapat identity (Theorem 3.6) which describes an inverse relation between extended Laplacian and resistance matrices. This simple identity in combination with the Schur complement leads to two interesting theories: (i) taking submatrices of the resistance matrix (i.e. metric subspaces) is equivalent to Kron reduction of graphs, while (ii) taking submatrices of the Laplacian matrix can be used to study the effect of merging nodes on the effective resistance. Kron reduction in particular plays an important role in the rest of this thesis. The Fiedler–Bapat identity also features the resistance curvature $\mathbf{p}$ and resistance radius σ^2, which are interpreted in the context of discrete curvature in Chapter 6.

Section 3.1 defines the resistance matrix and proves some basic properties.

In Section 3.2, we prove the Fiedler–Bapat identity and derive some of its consequences. In particular, we show Bapat's expression (3.8) for the inverse resistance matrix Ω^{-1} and we describe a new concise matrix characterization of resistance matrices in Theorem 3.9.

Section 3.3 introduces the Schur complement and some of its properties.

Section 3.3.1 shows a first application of combining the Fiedler–Bapat identity and the Schur complement: the Kron reduction of a graph $G \to G'$ is defined via the Schur complement of the Laplacian matrix $Q \to Q'$. Most importantly, (i) this operation maps Laplacians to Laplacians (Proposition 3.15), which means Kron reduction is well-defined, and (ii) the effective resistance is invariant under Kron reduction (Property 3.21). We discuss some examples, describe the structural effects of Kron

reduction and describe how the resistance curvature and radius change with the Kron reduction $\mathbf{p}, \sigma^2 \to \mathbf{p}', \sigma'^2$ in Proposition 3.25.

Section 3.3.2 shows the second application: the inverse of a Laplacian submatrix is related to the effective resistances in a graph with merged nodes (Definition 3.27). This results in new expressions for the effective resistance in node-merged graphs in Proposition 3.34, which are illustrated in some examples.

Finally, in Section 3.4, we describe some properties of the resistance radius σ^2. In particular, we study the resistance radius for all possible Kron reductions of a graph, interpreted as a set function defined on subsets of the nodes. We prove some bounds and show an inclusion–exclusion property (Proposition 3.40) from which it follows that the resistance radius decomposes over its biconnected components.

The main references for this chapter are the books of Fiedler and Bapat [11, 96]. For the Schur complement, we refer to the book of Zhang [267] and for Kron reduction to the survey [78].

3.1 The resistance matrix

We recall from the previous chapter the definition of the resistance matrix.

Definition 3.1. The *resistance matrix* Ω of a graph is the matrix containing all pairwise effective resistances: $(\Omega)_{ij} = \omega_{ij}$ for all $i, j \in \mathcal{N}$.

For a graph on n nodes, the resistance matrix is thus an $n \times n$ matrix and since the underlying graph is assumed to be connected, all entries in Ω are finite (see for instance Proposition 4.5). Furthermore, following the definition of the effective resistance as a quadratic product with the pseudoinverse Laplacian, we know that the resistance matrix satisfies the following properties:

Property 3.2. *The resistance matrix is a real symmetric matrix with zero diagonal and positive off-diagonal entries.*

Instead of defining the resistance matrix entrywise, it can also be expressed compactly in terms of the pseudoinverse Laplacian and its diagonal as:

$$\Omega = \mathbf{u}\boldsymbol{\zeta}^T + \boldsymbol{\zeta}\mathbf{u}^T - 2Q^\dagger \text{ with } \boldsymbol{\zeta} := \operatorname{diag}(Q^\dagger). \tag{3.1}$$

Equation (3.1) shows that the resistance matrix equals the pseudoinverse Laplacian matrix up to a rank two matrix addition[1]. As a first consequence of equation (3.1), we find that certain quadratic forms with the resistance matrix can be written in terms of the pseudoinverse Laplacian as

$$\mathbf{f}^T\Omega\mathbf{f} = -2\mathbf{f}^T Q^\dagger \mathbf{f} \text{ for all } \mathbf{f} \in \operatorname{span}(\mathbf{u})^\perp. \tag{3.2}$$

Since $\ker(Q^\dagger) = \operatorname{span}(\mathbf{u})$, this can be summarized as follows:

$$Q^\dagger = -\frac{1}{2}\left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right)\Omega\left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right). \tag{3.3}$$

This implies a characterization of the Laplacian pseudoinverse diagonal in terms of average effective resistances:

Proposition 3.3. *For $n > 1$, the pseudoinverse diagonal is positive and satisfies*

$$(Q^\dagger)_{ii} = \langle\omega_i\rangle - \langle\omega\rangle \text{ with } \langle\omega_i\rangle := \frac{1}{n}\sum_{j\in\mathcal{N}}\omega_{ij} \text{ and } \langle\omega\rangle := \frac{1}{2n^2}\sum_{i,j\in\mathcal{N}}\omega_{ij}. \tag{3.4}$$

Proof. Following equation (3.3) we find that the pseudoinverse diagonal entries can be written as

$$(Q^\dagger)_{ii} = -\frac{1}{2}\Omega_{ii} + \frac{1}{n}\mathbf{e}_i^T\Omega\mathbf{u} - \frac{1}{2n^2}\mathbf{u}^T\Omega\mathbf{u}.$$

With $(\Omega)_{ii} = \omega_{ii} = 0$ and introducing $\langle\omega_i\rangle$ and $\langle\omega\rangle$, this completes the proof. We note that for $n = 1$ we have $Q^\dagger = 0$. □

In other words, the pseudoinverse diagonal entries are related to how far a node i is on average to the rest of the nodes in the graph, in terms of the resistance distance. This 'average distance' idea was used in [203, 248] to propose the vector $\boldsymbol{\zeta}$ as a measure of importance or 'centrality' of a node in a graph. The expression for $\langle\omega\rangle$ is closely related to the so-called Kirchhoff index, which we define for further reference:

Definition 3.4. The *Kirchhoff index* of a graph is defined as $Kf := \frac{1}{2}\mathbf{u}^T\Omega\mathbf{u}$.

The Kirchhoff index was originally defined in the context of chemical graph theory in analogy with the Wiener index (see Example 5.4) and has been widely studied since [18, 118, 191]. The Kirchhoff index also appears in applications as a measure of robustness [85, 104, 243, 244, 265]. From (3.4) and (3.1), we have

$$Kf = n^2\langle\omega\rangle = n\operatorname{tr}(Q^\dagger) = n\sum_{k:\mu_k>0}\mu_k^{-1},$$

[1] If the pseudoinverse diagonal is constant, this is a rank one matrix addition. This happens for instance in the case of node transitive graphs (see Example 6.10).

with Laplacian eigenvalues $\mu_1, \ldots, \mu_n$, and we will encounter a number of other expressions in the rest of this thesis.

Expression (3.1) for the resistance matrix in terms of the pseudoinverse Laplacian, in combination with the Courant–Fischer–Weyl Theorem on eigenvalues (see Appendix A.2) leads to the following important spectral characterization of the resistance matrix:

Property 3.5. *For $n > 1$ the resistance matrix has 1 positive and $n - 1$ negative eigenvalues. In particular, the resistance matrix is invertible.*

Proof. First, since $n > 1$ and the resistance matrix is off-diagonally positive, we find

$$\mathbf{u}^T \Omega \mathbf{u} > 0,$$

which by Corollary A.3 of the Courant–Fischer–Weyl Theorem implies that Ω has at least one positive eigenvalue. Second, from (3.2) and properties of the pseudoinverse Laplacian quadratic form, we find

$$\mathbf{f}^T \Omega \mathbf{f} = -2\mathbf{f}^T Q^\dagger \mathbf{f} < 0 \text{ for all nonzero } \mathbf{f} \in \operatorname{span}(\mathbf{u})^\perp,$$

which again by Corollary A.3 of the Courant–Fischer–Weyl Theorem implies that Ω has at least $n - 1$ negative eigenvalues. As the total number of eigenvalues is n, this completes the proof. □

Fiedler calls a matrix with one positive eigenvalue and negative eigenvalues otherwise an *elliptic matrix* [96]. We note that since Ω has zero trace, its positive eigenvalue is equal to the sum of its negative eigenvalues; the eigenvalues of the resistance matrix are further studied for instance in [269]. The invertibility of Ω will be very important in the rest of this thesis. In Chapters 5–6, the sum of the inverse matrix $\mathbf{u}^T \Omega^{-1} \mathbf{u}$ and its row sums $\Omega^{-1}\mathbf{u}$ will play a central role, and in the remainder of this chapter, we will derive a matrix identity for Ω^{-1} in terms of the Laplacian matrix and discuss some of its consequences.

3.2 The Fiedler–Bapat identity

We now introduce the *Fiedler–Bapat identity*, an inverse matrix identity relating the (inverse) resistance matrix to the Laplacian matrix. This identity was derived by Miroslav Fiedler in the context of simplex geometry [96] (see Chapter 4) and by Subak-Sharpe et al. in electrical circuit theory [228], and is closely related to Bapat's expression for the inverse resistance matrix [11, §9.5],[10].

Theorem 3.6 (Fiedler–Bapat identity)**.** *The resistance and Laplacian matrix of a graph satisfy*

$$\begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega \end{pmatrix}^{-1} = -\frac{1}{2}\begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & Q \end{pmatrix}, \tag{3.5}$$

where $\mathbf{p} = \frac{1}{2}Q\boldsymbol{\zeta} + \frac{1}{n}\mathbf{u}$ *and* $\sigma^2 = \frac{1}{4}\boldsymbol{\zeta}^T Q\boldsymbol{\zeta} + \frac{1}{n}\mathbf{u}^T\boldsymbol{\zeta}$.

Proof. We first confirm that the 'extended resistance matrix' $\left(\begin{smallmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega \end{smallmatrix}\right)$ is invertible. Suppose for contradiction that the extended resistance matrix is singular instead. Then there exist $\alpha \in \mathbb{R}, \mathbf{f} \in \mathbb{R}^n$, not both equal to zero, such that

$$\begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega \end{pmatrix}\begin{pmatrix} -\alpha \\ \mathbf{f} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix} \Leftrightarrow \begin{cases} \mathbf{u}^T\mathbf{f} = 0 \\ \Omega\mathbf{f} = \alpha\mathbf{u}\,. \end{cases}$$

If $\alpha = 0$ then $\Omega\mathbf{f} = 0$, which by invertibility of the resistance matrix implies that $\mathbf{f} = 0$ as well. But α and $\mathbf{f}$ cannot both be zero, hence $\alpha \neq 0$. But then $\Omega\mathbf{f} = \alpha\mathbf{u}$ implies that $\mathbf{f} \neq 0$ (again since Ω is nonsingular) and from $\mathbf{u}^T\mathbf{f} = 0$ we arrive at the contradiction

$$0 = \alpha\mathbf{f}^T\mathbf{u} = \mathbf{f}^T\Omega\mathbf{f} = -2\mathbf{f}^T Q^\dagger\mathbf{f} < 0 \text{ since } \mathbf{f} \in \operatorname{span}(\mathbf{u})^\perp.$$

The extended resistance matrix is thus invertible and there exist unique $\sigma^2 \in \mathbb{R}, \mathbf{p} \in \mathbb{R}^n$ and $\widetilde{Q} \in \mathbb{R}^{n\times n}$ such that

$$\begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega \end{pmatrix}\begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & \widetilde{Q} \end{pmatrix} = \begin{pmatrix} -2 & \mathbf{0}^T \\ \mathbf{0} & -2I \end{pmatrix} \Leftrightarrow \begin{cases} \mathbf{u}^T\mathbf{p} = 1 \\ 2\sigma^2\mathbf{u} = \Omega\mathbf{p} \\ \Omega\widetilde{Q} = -2I + 2\mathbf{u}\mathbf{p}^T \\ \widetilde{Q}\mathbf{u} = 0\,. \end{cases}$$

We now show that these four equations determine $\sigma^2, \mathbf{p}$ as in the theorem and show that $\widetilde{Q}$ is equal to the Laplacian matrix Q corresponding to Ω. From the identities $2\sigma^2\mathbf{u} = \Omega\mathbf{p}$ and $\mathbf{u}^T\mathbf{p} = 1$ and expression (3.1) for the resistance matrix, we find

$$2\sigma^2\mathbf{u} = (\mathbf{u}\boldsymbol{\zeta}^T + \boldsymbol{\zeta}\mathbf{u}^T - 2Q^\dagger)\mathbf{p} \Leftrightarrow \left(2\sigma^2 - \boldsymbol{\zeta}^T\mathbf{p} - \frac{1}{n}\boldsymbol{\zeta}^T\mathbf{u}\right)\mathbf{u} = \left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right)\boldsymbol{\zeta} - 2Q^\dagger\mathbf{p}.$$

In the second expression, the left-hand side is a vector in $\operatorname{span}(\mathbf{u})$ and the right-hand side a vector in $\operatorname{span}(\mathbf{u})^\perp$, thus we know that both sides must be equal to zero, and thus

$$2\sigma^2 = (\mathbf{p} + \tfrac{1}{n}\mathbf{u})^T\boldsymbol{\zeta} \quad \text{and} \quad 2Q^\dagger\mathbf{p} = \left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right)\boldsymbol{\zeta} \tag{3.6}$$

Left-multiplying the second expression in (3.6) with the Laplacian matrix, we find

$$2\left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right)\mathbf{p} = Q\boldsymbol{\zeta} \Leftrightarrow \mathbf{p} = \frac{1}{2}Q\boldsymbol{\zeta} + \frac{\mathbf{u}}{n},$$

which confirms the formula for $\mathbf{p}$ in the theorem. Introducing the vector $\mathbf{p}$ in the first expression in (3.6) also confirms the formula for σ^2.

The matrix $\widetilde{Q}$ is determined by the remaining identities $\widetilde{Q}\mathbf{u} = 0$ and $\Omega\widetilde{Q} = -2I + 2\mathbf{u}\mathbf{p}^T$. Left-multiplying the second identity by the Laplacian matrix Q corresponding to the resistance matrix Ω, introducing expression (3.1) and invoking $\mathbf{u}^T\mathbf{p} = 1$, we find

$$Q(\mathbf{u}\boldsymbol{\zeta}^T + \boldsymbol{\zeta}\mathbf{u}^T - 2Q^\dagger)\widetilde{Q} = -2Q \Leftrightarrow \widetilde{Q} = Q,$$

which completes the proof. □

We will also call the matrices appearing in (3.5) the *extended* resistance and *extended* Laplacian matrix and use index '1' to denote the extra row/column. As we will discuss in Section 4.2, the resistance matrix is a squared Euclidean distance matrix. In this geometric setting, the extended resistance matrix is an example of a *Cayley-Menger matrix* which is a central tool in the theory of distance geometry [25].

The proof of Theorem 3.6 readily generalizes to any symmetric zero-diagonal matrix A with $\mathbf{u}^T A\mathbf{u} > 0$ and $\mathbf{f}^T A\mathbf{f} < 0$ for $\mathbf{f} \in \text{span}(\mathbf{u})^\perp$; these are squared Euclidean distance matrices of simplices (see Chapter 4), which was Fiedler's original context. More generally, for any symmetric invertible matrix A with $\mathbf{u}^T A^{-1}\mathbf{u} \neq 0$ the extended matrix $\left(\begin{smallmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & A \end{smallmatrix}\right)$ is invertible[2] and similar matrices/vectors/scalars to $Q/\mathbf{p}/\sigma^2$ can be defined based on the inverse extended matrix A.

The Fiedler–Bapat identity features two new resistance-based quantities: $\mathbf{p}$ and σ^2.

We will call $\mathbf{p}$ the *resistance curvature*. This is a vector defined on the nodes of the graph[3]. To guide the intuition, one may think that p_i is related to how 'clustered' the neighbourhood of i is. If there are a lot of triangles around i, then p_i will be large (and positive) whereas if there are few triangles around i, then p_i will be small and possibly negative. More precise interpretations are provided in Chapter 5, where we relate $\mathbf{p}$ to the "boundary of a graph" and in Chapter 6, where we interpret $\mathbf{p}$ in the context of discrete curvature; the latter underlies our choice for the terminology "resistance curvature". Several alternative expressions for $\mathbf{p}$ are presented throughout this thesis and are summarized in Section 6.2.3.

We will call σ^2 the *resistance radius*. This is a scalar invariant defined for a graph. To guide the intuition, one may think of σ^2 as a notion of 'size' for the graph, similar

[2]Follow the same proof as in Theorem 3.6, and for $\alpha \neq 0$, use the contradiction $0 = \alpha^{-1}\mathbf{u}^T\mathbf{f} = \mathbf{u}^T A^{-1}\mathbf{u} \neq 0$.

[3]As further explained in Appendix A.1, a vector $\mathbf{f} \in \mathbb{R}^n$ can also be interpreted as a function $f : \mathcal{N} \to \mathbb{R}$, where the vector entries f_i are given by the function values $f(i)$ on the nodes.

to the radius of a metric space but related to the effective resistance between all pairs of nodes. More precisely, in Section 5.3.1 we show that the resistance radius σ^2 is determined up to a factor two by the largest effective resistance in a graph. In Section 3.4 we discuss the resistance radius as a 'set function' on subsets of nodes, in Chapter 5 we relate σ^2 to the concept of variance on graphs (this motivated the notation 'σ^2') and in Chapter 6, we discuss the relation to discrete curvature.

The vector $\mathbf{p}$ and scalar σ^2 have been studied before in the context of distance geometry [111], simplex geometry [96], electrical circuit theory [228] and graph theory [11]. Our main contributions to the theory of these resistance-based invariants are presented in Chapters 5–6.

As shown already in the proof of the Fiedler–Bapat identity, the resistance curvature and radius satisfy the following identities which may be taken as alternative definitions:

Corollary 3.7. *The Fiedler–Bapat identity implies that*

$$\mathbf{u}^T\mathbf{p} = 1 \quad \text{and} \quad \Omega\mathbf{p} = 2\sigma^2\mathbf{u} \quad \text{and} \quad Q\Omega = -2I + 2\mathbf{p}\mathbf{u}^T \tag{3.7}$$

Proof. From the Fiedler–Bapat identity, we find:

$$\begin{pmatrix} -2 & 0^T \\ 0 & -2I \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega \end{pmatrix} \begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & Q \end{pmatrix} = \begin{pmatrix} -2\mathbf{u}^T\mathbf{p} & \mathbf{u}^TQ \\ 4\sigma^2\mathbf{u} - 2\Omega\mathbf{p} & -2\mathbf{u}\mathbf{p}^T + \Omega Q \end{pmatrix}.$$

□

These identities are also noted in [11, 96, 228]. The identity for $Q\Omega$ will be particularly useful and for further reference, we will call this the $Q\Omega$ *identity*. A second corollary of the Fiedler–Bapat identity is the following expression for the inverse resistance matrix:

Corollary 3.8. *The inverse resistance matrix is equal to*

$$\Omega^{-1} = -\frac{1}{2}Q + \frac{\mathbf{p}\mathbf{p}^T}{2\sigma^2}. \tag{3.8}$$

Proof. Following Corollary 3.7, we know that

$$\begin{cases} Q\Omega = -2I + 2\mathbf{p}\mathbf{u}^T \\ \Omega\mathbf{p} = 2\sigma^2\mathbf{u} \end{cases} \Leftrightarrow \begin{cases} Q = -2\Omega^{-1} + 2\mathbf{p}\mathbf{u}^T\Omega^{-1} \\ \Omega^{-1}\mathbf{u} = \mathbf{p}/(2\sigma^2) \end{cases} \Leftrightarrow \Omega^{-1} = -\frac{1}{2}Q + \frac{\mathbf{p}\mathbf{p}^T}{2\sigma^2},$$

which completes the proof. In Section 3.3, we show an alternative proof based on the Schur complement. □

Expression (3.8) was derived for tree graphs[4] by Graham & Lovász in [111], and was later generalized to the resistance matrix of general weighted graphs by Bapat in [10]. Corollary 3.8 also allows to express the Laplacian matrix in terms of the inverse resistance matrix as follows:

$$Q = -2\Omega^{-1} + 2\frac{\Omega^{-1}\mathbf{u}\mathbf{u}^T\Omega^{-1}}{\mathbf{u}^T\Omega^{-1}\mathbf{u}}. \tag{3.9}$$

This expression can be used to provide a matrix characterization of resistance matrices, without explicit reference to an underlying graph:

Theorem 3.9. *Let Ω be a real symmetric matrix. Then Ω is a resistance matrix if and only if it is invertible and satisfies*

$$(\mathbf{u}^T\Omega^{-1}\mathbf{u})(\Omega^{-1})_{ij} \geq (\Omega^{-1}\mathbf{u}\mathbf{u}^T\Omega^{-1})_{ij} \text{ for all } i \neq j. \tag{3.10}$$

Proof. *(forward direction)* If Ω is a resistance matrix, then it is invertible by Property 3.5 and satisfies condition (3.10) by definition of the Laplacian in terms of the inverse resistance matrix and positivity of link weights:

$$\left(\Omega^{-1} - \frac{\Omega^{-1}\mathbf{u}\mathbf{u}^T\Omega^{-1}}{\mathbf{u}^T\Omega^{-1}\mathbf{u}}\right)_{ij} = -\frac{1}{2}(Q)_{ij} \geq 0 \text{ for all } i \neq j.$$

(converse direction) Let Ω be a real symmetric invertible matrix that satisfies condition (3.10). We first show that $\mathbf{u}^T\Omega^{-1}\mathbf{u} \neq 0$. Suppose for contradiction that $\mathbf{u}^T\Omega^{-1}\mathbf{u} = 0$. Then from condition (3.10) we know that $(\Omega^{-1}\mathbf{u})_i(\Omega^{-1}\mathbf{u})_j \leq 0$ for all $i \neq j$, which is only possible if $\Omega^{-1}\mathbf{u} = 0$. But this is in contradiction with Ω being invertible, and thus we must have $\mathbf{u}^T\Omega^{-1}\mathbf{u} \neq 0$ and we can define the matrix $X := -2\Omega^{-1} + 2\frac{\Omega^{-1}\mathbf{u}\mathbf{u}^T\Omega^{-1}}{\mathbf{u}^T\Omega^{-1}\mathbf{u}}$. This matrix X is real, symmetric, has nonpositive off-diagonal entries (by condition (3.10)) and satisfies $X\mathbf{u} = 0$ by construction. Moreover, from

$$\operatorname{rank}(2\Omega^{-1}) \leq \operatorname{rank}(X) + \operatorname{rank}\left(2\tfrac{\Omega^{-1}\mathbf{u}\mathbf{u}^T\Omega^{-1}}{\mathbf{u}^T\Omega^{-1}\mathbf{u}}\right) \Rightarrow \operatorname{rank}(X) \geq n - 1,$$

we know that X has rank $n-1$ and thus $\ker(X) = \mathbf{u}$. All together, we thus find that X is a Laplacian matrix and thus, by construction of X from Ω, we know that Ω is a resistance matrix, as required. This completes the proof. □

This characterization is of course equivalent to other characterizations of the effective resistance matrix, such as the characterizations of Subak-Sharpe & Moore [179], Weihrauch [259], Kigami's definition based on resistance forms [142, §2.3] and the

[4]For tree graphs, the shortest-path distance matrix is equal to the resistance matrix; see Proposition 4.5.

different equivalences of Fiedler in [92, 94]. As far as we know, however, the concise matrix characterization in terms of the inverse resistance matrix in Theorem 3.9 is new.

The condition (3.10) is particularly relevant given the important role of the scalar $\mathbf{u}^T\Omega^{-1}\mathbf{u}$ and vector $\Omega^{-1}\mathbf{u}$ in this thesis. Following Corollary 3.8, we find that they are related to the resistance curvature and radius as follows:

Corollary 3.10. *The Fiedler–Bapat identity implies that*

$$\mathbf{p} = \frac{\Omega^{-1}\mathbf{u}}{\mathbf{u}^T\Omega^{-1}\mathbf{u}}, \qquad \sigma^2 = \frac{1}{2}\left(\mathbf{u}^T\Omega^{-1}\mathbf{u}\right)^{-1} \quad \text{and} \quad \sigma^2 = \frac{1}{2}\mathbf{p}^T\Omega\mathbf{p}$$

Proof. Right-multiplying the inverse resistance matrix (3.8) with the all-one vector $\mathbf{u}$, we find $\Omega^{-1}\mathbf{u} = \mathbf{p}/(2\sigma^2)$ and further left-multiplying with $\mathbf{u}^T$ yields $\mathbf{u}^T\Omega^{-1}\mathbf{u} = 1/(2\sigma^2)$. The relation between the resistance curvature and radius follows more directly from $\mathbf{u}^T\mathbf{p} = 1$ and $\Omega\mathbf{p} = 2\sigma^2\mathbf{u}$ in Corollary 3.7. This completes the proof. □

These expressions also appear in [13, 111, 228] and further alternative expression for $\mathbf{p}$ and σ^2 are summarized in Section 6.2.3.

In addition to relating the resistance and Laplacian matrix and introducing the resistance curvature and radius, the Fiedler–Bapat identity is also the starting point of a rich theory of submatrices of the resistance and Laplacian matrices. We introduce this theory in the following sections, starting with the Schur complement.

3.3 The Schur complement

The Schur complement, named after Issai Schur by Emilie Haynsworth in [125], is an algebraic operation defined on matrices. For an $n \times n$ matrix A with rows/columns indexed by the set $\mathcal{N}$ and block decomposition $A = \left(\begin{smallmatrix} A_{\mathcal{V}\mathcal{V}} & A_{\mathcal{V}\mathcal{V}^c} \\ A_{\mathcal{V}^c\mathcal{V}} & A_{\mathcal{V}^c\mathcal{V}^c} \end{smallmatrix}\right)$, the Schur complement is defined as follows:

Definition 3.11 (Schur complement)**.** The *Schur complement* of a matrix A with respect to a nonempty subset (of row/column indices) $\mathcal{V}$ is defined as

$$A/\mathcal{V}^c := A_{\mathcal{V}\mathcal{V}} - A_{\mathcal{V}\mathcal{V}^c}(A_{\mathcal{V}^c\mathcal{V}^c})^{-1}A_{\mathcal{V}^c\mathcal{V}}$$

if the submatrix $A_{\mathcal{V}^c\mathcal{V}^c}$ is invertible.

If the subset $\mathcal{V} = \mathcal{N}$ is the full set of row/column indices, we have $A/\emptyset = A$. The Schur complement $A/\mathcal{V}^c$ with respect to $\mathcal{V}$ is also called the $\mathcal{V}$-Schur complement of

A and for a single node v, the v^c-Schur complement $A/\{v\}$ is called an *elementary* Schur complement.

The Schur complement of a matrix appears when studying subsystems of a linear system of equations. For the system

$$A\mathbf{x} = \mathbf{b} \text{ with } \mathbf{b}_{\mathcal{V}^c} = 0, \tag{3.11}$$

the Schur complement can be used to find the solution $\mathbf{x}$ restricted to $\mathcal{V}$ based on the subsystem $(A/\mathcal{V}^c)\mathbf{x}_{\mathcal{V}} = \mathbf{b}_{\mathcal{V}}$ (see Property 3.13). Equation (3.11) is also called a *boundary value problem* since the problem is specified on a subset $\mathcal{V}$, 'the boundary'. The Schur complement is a well-studied matrix operation and we will only introduce those results that are relevant for the context of resistance matrices and Laplacians. For an overview of the history and theory of the Schur complement — such as its relation to Gaussian elimination, determinants, matrix inertia and its many closure results — we refer to the excellent book [267] and references therein.

A first relevant property of the Schur complement is the so-called quotient property [62],[267, Thm. 1.3], which implies the following:

Property 3.12 (composition property)**.** *Any Schur complement can be composed by a sequence of elementary Schur complements, as*

$$A/\mathcal{V}^c = \big((A/\{v_k\})/\dots\big)/\{v_1\} \text{ with } \mathcal{V}^c = \{v_1, \dots, v_k\} \tag{3.12}$$

whenever the Schur complement exists.

The "composition" terminology refers to the perspective where the Schur complement is seen as an operation $(\star)/\mathcal{V}^c$ acting on matrices; the quotient property [267, Thm. 1.3] says that these operations can be composed as $(\star)/\mathcal{W} \circ (\star)/\mathcal{V} = ((\star)/\mathcal{V})/\mathcal{W} = (\star)/\{\mathcal{V} \cup \mathcal{W}\}$ for two disjoint node sets $\mathcal{V}, \mathcal{W}$. In particular, this implies that any Schur complement can be composed by elementary Schur complements as in (3.12).

The composition property is useful when elementary Schur complements are well-understood for a class of matrices, since it allows one to generalize this understanding (via composition) to general Schur complements. This is illustrated in the proof of Proposition 3.15 on the Schur complement of Laplacian matrices, and lies at the basis of several closure properties of the Schur complement. We note that the order of the elements of $\mathcal{V}^c$ in (3.12) is arbitrary.

A second relevant result relates submatrices of the matrix inverse to the Schur complement:

Property 3.13. *Let A be an invertible matrix. Then*

$$(A/\mathcal{V}^c)^{-1} = (A^{-1})_{\mathcal{V}\mathcal{V}} \tag{3.13}$$

whenever the Schur complement exists.

We refer to [267, Thm. 1.2] for a proof. To illustrate the utility of Property 3.13 in combination with the Fiedler–Bapat identity, we give an alternative proof of Corollary 3.8 on the inverse resistance matrix:

Proof of Corollary 3.8 First, from the Fiedler–Bapat identity, we know that

$$\left[\begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & Q \end{pmatrix}^{-1}\right]_{1^c1^c} = \left[-\frac{1}{2}\begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega \end{pmatrix}\right]_{1^c1^c} = -\frac{1}{2}\Omega,$$

where 1^c denotes the set of all rows/columns except the first. Second, by the Schur complement property $(A^{-1})_{1^c1^c} = (A/\{1\})^{-1}$, we find that

$$\left[\begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & Q \end{pmatrix}^{-1}\right]_{1^c1^c} = \left(Q - \frac{\mathbf{p}\mathbf{p}^T}{\sigma^2}\right)^{-1}.$$

Equating both expressions proves the corollary. □

In the following sections, we show two classes of results results that follow by combining the Schur complement and the Fiedler–Bapat identity. First, we show how taking the Schur complement of the extended Laplacian matrix leads to the concept of Kron reduction. Second, we show how taking the Schur complement of the extended resistance matrix leads to the concept of merging nodes.

3.3.1 Kron reduction

In the context of graphs and the effective resistance, the Schur complement plays an important role because it leads to the so-called Kron reduction of graphs (named after Gabriel Kron):

Definition 3.14 (Kron reduction)**.** The *$\mathcal{V}$-Kron reduction* of a graph G is the graph with Laplacian matrix $Q' = Q/\mathcal{V}^c$.

We also call this *the Kron reduction of G with respect to $\mathcal{V}$* and write[5] $G' = G/\mathcal{V}^c$, following the notation for the Schur complement. For $\mathcal{V} = \mathcal{N}$, we have $G' = G/\emptyset =$

[5]The notation G/ℓ for a link $\ell \in \mathcal{L}$ is sometimes used to denote merging the two end nodes of ℓ. This is a different operation from Kron reduction and we will only use the notation $G/\star$ in the context of Kron reduction.

G, and we furthermore require that $\mathcal{V} \neq \emptyset$ is a nonempty subset. The v^c-Kron reduction $G/\{v\}$ with respect to all but a single node is also called an *elementary* Kron reduction. A survey of the properties of Kron reduction and its applications to electrical networks is given in [78].

From the definition, it follows that the Kron reduction is a graph defined on the node set $\mathcal{V}$ and that it is uniquely defined if it exists. It is less immediate, however, whether Kron reduction is always well-defined since two things could go wrong: first, the Schur complement can be undefined if the Laplacian submatrix $Q_{\mathcal{V}^c\mathcal{V}^c}$ is singular and second, the matrix $Q/\mathcal{V}^c$ might not satisfy all properties of a Laplacian matrix, e.g. not have zero row sum, nonpositive off-diagonal entries, etc. However, we have:

Proposition 3.15 (closure)**.** *Let Q be a Laplacian matrix and $\emptyset \neq \mathcal{V} \subseteq \mathcal{N}$. Then the Schur complement $Q/\mathcal{V}^c$ is always well-defined and a Laplacian matrix.*

Proof. *(well-defined)* Proposition 3.26 (later) shows that the matrix $Q_{\mathcal{V}^c\mathcal{V}^c}$ is positive definite and thus, in particular, that it is invertible. This follows from Proposition 2.7 for the Laplacian matrix. Consequently, the Schur complement is always well-defined.

(Laplacian) Next we show that the Schur complement of a Laplacian matrix is again a Laplacian. By the composition property 3.12 of the Schur complement, it suffices to consider elementary Schur complements, with $\mathcal{V}^c = \{x\}$, for which the Schur complement equals

$$Q' := Q/\{x\} = Q_{x^cx^c} - \frac{\mathbf{q}\mathbf{q}^T}{k_x}, \tag{3.14}$$

where $\mathbf{q} := -Q_{x^cx} = Q_{x^cx^c}\mathbf{u}$ is the $(n-1) \times 1$ column vector containing the link weights to all other nodes; we note that $\mathbf{u}^T\mathbf{q} = k_x$.

We now show that the matrix Q' satisfies all properties of a Laplacian matrix. By construction, Q' is a real symmetric $(n-1) \times (n-1)$ matrix with nonpositive off-diagonal entries, since $(Q')_{ij} = -c_{ij} - c_{ix}c_{xj}/k_x \leq 0$. It thus remains to show that $\ker(Q) = \operatorname{span}(\mathbf{u})$. From (3.14) it follows that $Q'\mathbf{u} = 0$ and thus that $\operatorname{span}(\mathbf{u}) \subseteq \ker(Q')$. Suppose for contradiction that this subset is strict. Then there must exist $\mathbf{f} \notin \operatorname{span}(\mathbf{u})$ with $Q'\mathbf{f} = 0$. But then we find that

$$Q'\mathbf{f} = 0 \Leftrightarrow Q_{x^cx^c}\mathbf{f} - \frac{\mathbf{q}\mathbf{q}^T}{k_x}\mathbf{f} = 0 \Leftrightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} k_x & -\mathbf{q}^T \\ -\mathbf{q} & Q_{x^cx^c} \end{pmatrix} \begin{pmatrix} \frac{\mathbf{q}^T\mathbf{f}}{k_x} \\ \mathbf{f} \end{pmatrix} = Q \begin{pmatrix} \frac{\mathbf{q}^T\mathbf{f}}{k_x} \\ \mathbf{f} \end{pmatrix}$$

which is in contradiction with $\ker(Q) = \operatorname{span}(\mathbf{u})$ since $\mathbf{f}$ is not a constant vector and so $\begin{pmatrix} \mathbf{q}^T\mathbf{f}/k_x \\ \mathbf{f} \end{pmatrix}$ is not in $\operatorname{span}(\mathbf{u})$. This contradiction confirms that $\ker(Q') = \operatorname{span}(\mathbf{u})$ which by Proposition 2.5 thus completes the proof that Q' is a Laplacian matrix. □

In other words, the class of Laplacian matrices is *closed under Schur complement*. In Chapter 4, we discuss how this corresponds to closure of associated geometric objects. Proposition 3.15 establishes that the Kron reduction $G/\mathcal{V}^c$ is a well-defined and unique graph for each nonempty subset $\mathcal{V}$.

To develop an understanding of the properties of Kron reduction on graphs, expression (3.14) for the elementary Schur complement of a Laplacian is a good starting point. This expression shows that Kron reduction is a local operation: the elementary Kron reduction $G' = G/\{x\}$ is equal to the graph $G\backslash\{x\}$ in which x is removed and in which the neighbours N_x of x are fully connected; by expression (3.14), the link weights and degrees are given by:

$$c'_{ij} = c_{ij} + \frac{c_{ix}c_{xj}}{k_x} \quad \text{and} \quad k'_i = k_i - \frac{c_{ix}^2}{k_x} \text{ for all } i, j \in N_x.$$

This 'locality' is illustrated in Figure 3.1, which shows that only the neighbourhood of the removed node(s) changes. We now give a number of examples.

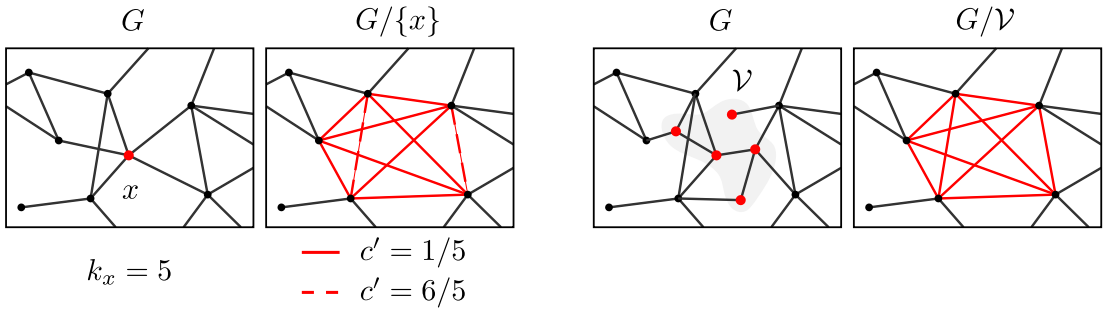


Figure 3.1: Example of an elementary Kron reduction (left) and a general Kron reduction (right); the new or changed links in the Kron reduction graphs are shown in red. These figures illustrate that the graph change due to Kron reduction is *local*.

Example 3.16. Figure 3.2 shows an example of Kron reduction. The composition property applied to this example says that $G/\{x, y\} = (G/\{x\})/\{y\}$.

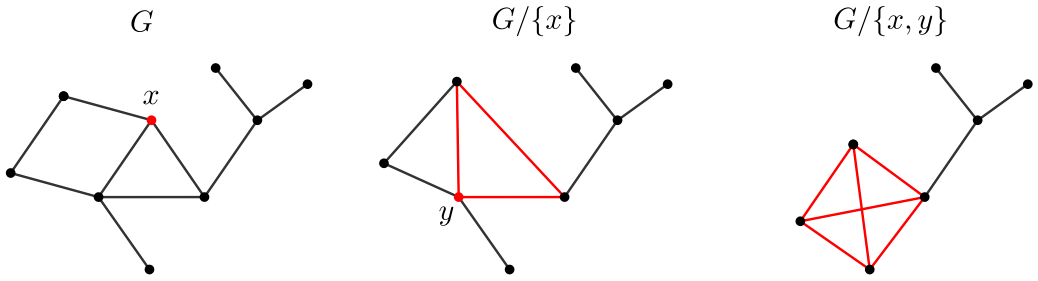


Figure 3.2: Example of Kron reduction (link weights are not given).

Example 3.17 (**cut nodes**)**.** If G is a graph and v a node with combinatorial degree one, also called a pendant node, then the Kron reduction $G/\{v\}$ is equal to $G\backslash\{v\}$, i.e. with the node removed. More generally, if x is a cut node that disconnects the graph in two components supported on A, B (i.e. A, x, B is a partition of the node set), then the Kron reduction with respect to A^c is equal to the graph with A removed: $G/A = G\backslash A$. This follows from the definition of the Schur complement, where we find that $\mathbf{e}_i^T Q_{B'A}(Q_{AA})^{-1}Q_{AB'}\mathbf{e}_j = 0$ for all $i \neq j \in B' = B \cup \{x\}$, and thus that $Q/A = Q_{BB}$.

Example 3.18 (**bipartite projection**)**.** Let G be a bipartite graph with a partition A, B of the nodes such that there are no links between pairs of nodes in A and pairs of nodes in B, and thus $\mathcal{L} \subseteq A \times B$. The Kron reduction $G' = G/B$ is a graph on the nodes A with links $c'_{ij} = \sum_{b\in B} c_{ib}c_{bj}/k_b$ between $i, j \in A$. This graph G' is also known as the *projection* of G onto the node set A; see for instance [270].

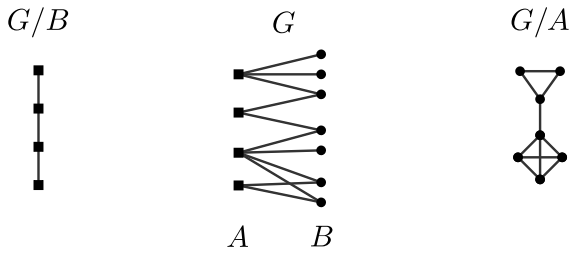


Figure 3.3: Example of bipartite projection (link weights are not given).

These examples give some initial insight in the Kron reduction, but for general graphs and general sets $\mathcal{V}$, it is not immediately clear what the graph G' looks like structurally. The following result relates the structure of G and G' (see [78, Thm. 3.4], [95, Thm. 1.10]):

Proposition 3.19. *Let $G' = G/\mathcal{V}^c$. Then a link (i, j) exists between two nodes in G' if and only if either (i, j) is a link in G or there exists an $i - j$ path in G, all nodes of which lie in $\mathcal{V}^c$.*

The weighted version of Proposition 3.19 says that $c'_{ij} \geq c_{ij}$ (where $c_{ij} = 0$ if there is no link) with equality unless there is a path between i and j contained in $\mathcal{V}^c$. An immediate consequence of Proposition 3.19 is that Kron reduction preserves cut nodes:

Proposition 3.20. *Let G be a graph and G' any Kron reduction. Then x is a cut node in G' if and only if x is a cut node in G.*

Proof. *(forward direction)* Let $G' = G/\mathcal{V}^c$ with $x \in \mathcal{V}$. If x is a cut node that disconnects the graph in two components supported on A, B, then all paths from $a \in A$ to $b \in B$ must pass through x. Since $x \notin \mathcal{V}^c$, there are no $a - b$ paths with all nodes in $\mathcal{V}^c$ and by Proposition 3.19 it then follows that there are no links between $A' = A \backslash \mathcal{V}^c$ and $B' = B \backslash \mathcal{V}^c$ in G' and thus that all paths from $a \in A'$ to $b \in B'$ must pass through x; equivalently, x is a cut node between A', B' in G'.

(converse direction) The graph G must have at least three nodes for the converse to apply. Let G be a graph with Kron reduction G' in which x is a cut node between a and b. Suppose for contradiction that x is not a cut node between a and b in G. Then there exists some $a - b$ path P in G which does not go through x. Then we can define the Kron reduction $G'' = G/\{a, x, b\}^c$ in which a and b are linked, by Proposition 3.19 and the fact that $P \subseteq \{a, x, b\}^c$. In particular, x is not a cut node in G''. However, this is in contradiction with the forward direction of the proof, which says that x must be a cut node in G'' because G'' is a Kron reduction of G' and x is a cut node in G'. This completes the converse direction and thus the proof. □

In particular, this result implies that the number of cut nodes (and thus the number of biconnected components) is nonincreasing with respect to Kron reduction.

While the Kron reduction is mathematically consistent and well-defined based on the properties of the Schur complement of the Laplacian matrix, its definition is until this point still 'unmotivated'. We now show why Kron reduction is a natural definition in the context of effective resistances:

Property 3.21. *Let G' be a Kron reduction of G. Then $\omega'_{ij} = \omega_{ij}$ for all pairs of nodes in G'.*

Proof. Let $G' = G/\mathcal{V}^c$. From the Fiedler–Bapat identity, we find that

$$\left[\begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & Q \end{pmatrix}^{-1}\right]_{\mathcal{V}_+\mathcal{V}_+} = \left[-\frac{1}{2}\begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega \end{pmatrix}\right]_{\mathcal{V}_+\mathcal{V}_+} = -\frac{1}{2}\begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega_{\mathcal{V}\mathcal{V}} \end{pmatrix}.$$

where $\mathcal{V}_+ = \mathcal{V} \cup \{1\}$ indicates the index set $\mathcal{V}$ including the first row/column index of

the extended matrix. Next, from the Schur complement property 3.13 we know that

$$\begin{aligned}
\left(\left[\begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & Q \end{pmatrix}^{-1}\right]_{\mathcal{V}_+\mathcal{V}_+}\right)^{-1} &= \begin{pmatrix} 4\sigma^2 & -2\mathbf{p}^T \\ -2\mathbf{p} & Q \end{pmatrix} / \mathcal{V}^c \\
&= \begin{pmatrix} 4\sigma^2 & -2\mathbf{p}_{\mathcal{V}}^T \\ -2\mathbf{p}_{\mathcal{V}} & Q_{\mathcal{V}\mathcal{V}} \end{pmatrix} - \begin{pmatrix} -2\mathbf{p}_{\mathcal{V}^c}^T \\ Q_{\mathcal{V}\mathcal{V}^c} \end{pmatrix} (Q_{\mathcal{V}^c\mathcal{V}^c})^{-1} \begin{pmatrix} -2\mathbf{p}_{\mathcal{V}^c} & Q_{\mathcal{V}^c\mathcal{V}} \end{pmatrix} \\
&= \begin{pmatrix} 4\sigma'^2 & -2\mathbf{p}'^T \\ -2\mathbf{p}' & Q/\mathcal{V}^c \end{pmatrix} \\
&= \left[-\frac{1}{2}\begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega' \end{pmatrix}\right]^{-1}. \quad \text{(Fiedler–Bapat for } G'\text{)}.
\end{aligned}$$

As a result, we find that $\Omega' = \Omega_{\mathcal{V}\mathcal{V}}$ and thus that $\omega'_{ij} = \omega_{ij}$ for all $i, j \in \mathcal{V}$. The relation between $\mathbf{p}', \mathbf{p}$ and σ'^2, σ^2 in Proposition 3.25 follows by equating the second and third expression on the right-hand side in the derivation above. □

In other words, Kron reduction preserves the effective resistance. An alternative route to defining the Kron reduction is thus via the effective resistance:

Definition 3.22 (Kron reduction 2)**.** The *$\mathcal{V}$-Kron reduction* of a graph G is the graph with node set $\mathcal{V}$ and the same effective resistances as G.

Starting from this definition, it is again not immediately clear whether this produces a well-defined graph. By this definition, the Kron reduction is a graph on the node set $\mathcal{V}$ and it is uniquely defined by the submatrix $\Omega_{\mathcal{V}\mathcal{V}}$ — i.e. if it is a resistance matrix, then there is a unique corresponding Laplacian obtained via the Fiedler–Bapat identity — but this submatrix could in principle fail to be a proper resistance matrix. However, following the relation between Kron reduction and effective resistances, we find that all is well:

Proposition 3.23 (closure)**.** *Any (nonempty) submatrix of a resistance matrix is again a resistance matrix.*

Proof. By Property 3.21, the Kron reduction $G' = G/\mathcal{V}^c$ has resistance matrix $\Omega' = \Omega_{\mathcal{V}\mathcal{V}}$ and thus $\Omega_{\mathcal{V}\mathcal{V}}$ is the resistance matrix of G'. □

The class of resistance matrices is *closed with respect to submatrices*. This closure result is not a generic property for distance matrices. Figure 3.4 for instance shows a graph for which a submatrix of its shortest-path distance matrix cannot be realized as the distance matrix of a graph:

Following Property 3.21, we discuss another Kron reduction example.

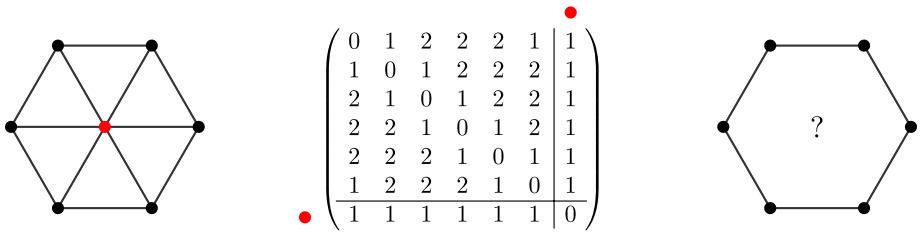


Figure 3.4: A graph (left) and its shortest-path distance matrix (middle). The distance submatrix excluding the red node (last row/column) has no realization as an unweighted graph: the adjacency matrix must satisfy $i \sim j$ if and only if $d(i,j) = 1$ which results in a cycle graph in which, for instance, $d(1,4) = 2$ is incorrect.

Example 3.24 (**two nodes**)**.** Let G be a graph with at least two nodes i, j. The Kron reduction $G' = G/\{i,j\}^c$ is a graph on two nodes i, j with effective resistance ω_{ij} and with link weight $c'_{ij} = \omega_{ij}^{-1}$ since (i,j) is a cut link in G'. If (i,j) is also a cut link in G, then by the weighted version of Proposition 3.19, it follows that $c'_{ij} = c_{ij}$ and thus that $c_{ij} = \omega_{ij}^{-1}$ for cut links. This shows an alternative proof for equality in the upper bound of Property 2.14.

As a final result, we show how Kron reduction influences the resistance curvature $\mathbf{p}$ and resistance radius σ^2. We find the following relation (see also [96, Thm. 3.2.3]):

Proposition 3.25. *Let G' be the $\mathcal{V}$-Kron reduction of a graph G. Then*

$$\sigma'^2 = \sigma^2 - \mathbf{p}_{\mathcal{V}^c}^T (Q_{\mathcal{V}^c\mathcal{V}^c})^{-1} \mathbf{p}_{\mathcal{V}^c} \quad \text{and} \quad \mathbf{p}' = \mathbf{p}_{\mathcal{V}} - Q_{\mathcal{V}\mathcal{V}^c}(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}\mathbf{p}_{\mathcal{V}^c}. \tag{3.15}$$

The proof of this proposition was included in the proof of Property 3.21. For elementary Kron reductions $G' = G/\{x\}$, this implies the much simpler formulas:

$$\sigma'^2 = \sigma^2 - \frac{p_x^2}{k_x} \quad \text{and} \quad p'_i = p_i + \frac{c_{ix}}{k_x} p_x. \tag{3.16}$$

These relations are central for the results on positively curved graphs in Chapter 6.

3.3.2 Merging nodes

Similar to how the submatrices of a resistance matrix lead to Kron reductions of a graph, the submatrices of a Laplacian have an interpretation in terms of effective resistances. A *Laplacian submatrix* is a matrix $Q_{\mathcal{V}^c\mathcal{V}^c}$ for some nonempty strict subset $\emptyset \neq \mathcal{V}^c \subset \mathcal{N}$ of nodes[6]. In the definition of Kron reduction, we already implicitly used the following property:

[6]We choose to consider Laplacian submatrices with respect to $\mathcal{V}^c$ and not $\mathcal{V}$, because this choice corresponds to merging the group of nodes $\mathcal{V}$ and is more consistent with the notation for Kron reductions.

Property 3.26. *A Laplacian submatrix is positive definite and thus invertible.*

Proof. By construction, a Laplacian submatrix is real and symmetric. Any quadratic product with $Q_{\mathcal{V}^c\mathcal{V}^c}$ corresponds to a quadratic product with the Laplacian matrix, as follows

$$\mathbf{f}^T Q_{\mathcal{V}^c\mathcal{V}^c}\mathbf{f} = \mathbf{g}^T Q\mathbf{g} \geq 0, \text{ with } \mathbf{g}_{\mathcal{V}^c} = \mathbf{f} \text{ and } \mathbf{g}_{\mathcal{V}} = 0,$$

which implies that $Q_{\mathcal{V}^c\mathcal{V}^c}$ is positive semidefinite (see Proposition A.2). Next, suppose for contradiction that the Laplacian submatrix is singular. Then there exists some nonzero vector $\mathbf{f} \in \ker(Q_{\mathcal{V}^c\mathcal{V}^c})$ such that the quadratic form $\mathbf{f}^T Q_{\mathcal{V}^c\mathcal{V}^c}\mathbf{f} = 0$ is zero and thus also $\mathbf{g}^T Q\mathbf{g} = 0$. But this is in contradiction with $\ker(Q) = \operatorname{span}(\mathbf{u})$ since $\mathbf{g}$ is not a constant vector. This proves that $Q_{\mathcal{V}^c\mathcal{V}^c}$ is positive definite, as required. A more direct proof using matrix theory uses the fact that Laplacian submatrices are symmetric strictly diagonally dominant matrices, which are positive definite. □

Since Laplacian submatrices are invertible, the Laplacian Schur complement and thus Kron reduction are well-defined. As a simple example, for $\mathcal{V}^c = \{i\}$, we find that $Q_{ii} = k_i$ is invertible, as expected. To further describe the inverse of Laplacian submatrices, we introduce the following graph operation:

Definition 3.27. Let $\mathcal{V}$ be a nonempty strict subset of nodes in a graph G. The $\mathcal{V} - v$ *merged graph* G' is obtained by merging the set of nodes $\mathcal{V}$ in G into a single node v with link weights $c'_{vi} = \sum_{j\in\mathcal{V}} c_{ij}$ for all $i \in \mathcal{V}^c$. The Laplacian matrices Q' and Q are related by

$$Q' := HQH^T,$$

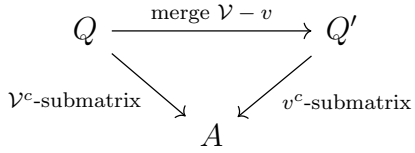
where H is the $(n - |\mathcal{V}| + 1) \times n$ matrix H with row indices $\mathcal{V}^c \cup \{v\}$ and column indices $\mathcal{N}$ and with entries $(H)_{ij} = 1$ if $i = j \in \mathcal{V}^c$ and $(H)_{vj} = 1$ for $j \in \mathcal{V}$ and zeroes otherwise.

In other words, a $\mathcal{V} - v$ merged graph Laplacian is obtained by adding up the rows/columns in $\mathcal{V}$ and assigning index v to the new row/column. Figure 3.5 below illustrates this construction.

The reason why node-merged graphs are relevant for Laplacian submatrices is the following observation: if G' is the $\mathcal{V} - v$ merged graph of G, then

$$Q_{\mathcal{V}^c\mathcal{V}^c} = Q'_{v^cv^c}.$$

Schematically, this observation can be summarized by saying that the diagram



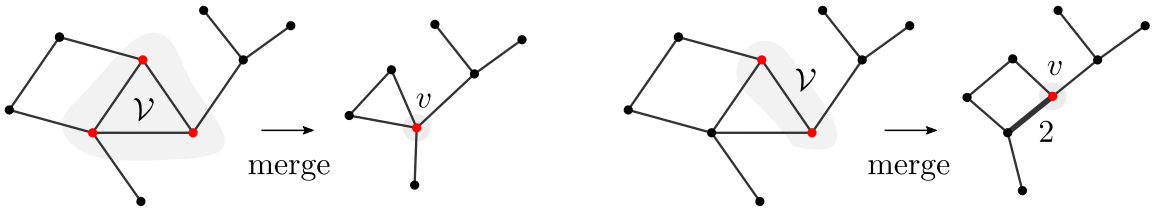


Figure 3.5: Two examples of node-merged graphs. Note the weight $c = 2$ in the right merged graph.

commutes for all Laplacians Q and all nonempty strict subsets of nodes $\mathcal{V}^c$, i.e. first taking the $\mathcal{V} - v$ merged graph $Q \to Q'$ and then taking the v^c submatrix is the same as taking the $\mathcal{V}^c$ submatrix of Q. In combination with Schur complements and the Fiedler–Bapat identity, this observation leads to the following results for the inverse of Laplacian submatrices (see also [94]):

Proposition 3.28. *The inverse Laplacian submatrix has entries*

$$[(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}]_{ij} = \frac{1}{2}\left(\omega'_{iv} + \omega'_{vj} - \omega'_{ij}\right)$$

with effective resistances ω' of the $\mathcal{V} - v$ merged graph G'.

Proof. Let G be a graph with a nonempty strict subset of nodes $\mathcal{V}^c$ and G' the $\mathcal{V} - v$ merged graph such that $Q_{\mathcal{V}^c\mathcal{V}^c} = Q'_{v^cv^c}$. The Fiedler–Bapat identity for the merged graph G' is equal to

$$\begin{pmatrix} 0 & 1 & \mathbf{u}^T \\ 1 & 0 & \Omega'_{vv^c} \\ \mathbf{u} & \Omega'_{v^cv} & \Omega'_{v^cv^c} \end{pmatrix}^{-1} = -\frac{1}{2}\begin{pmatrix} 4\sigma'^2 & -2p'_v & -2\mathbf{p}'^T_{v^c} \\ -2p'_v & Q'_{vv} & Q'_{vv^c} \\ -2\mathbf{p}'_{v^c} & Q'_{v^cv} & Q'_{v^cv^c} \end{pmatrix}.$$

By the Schur complement property $(A^{-1})_{v^cv^c} = (A/\{v\})^{-1}$ applied to the extended resistance matrix, we then find that

$$\left[-\frac{1}{2}Q'_{v^cv^c}\right]^{-1} = \begin{pmatrix} 0 & 1 & \mathbf{u}^T \\ 1 & 0 & \Omega'_{vv^c} \\ \mathbf{u} & \Omega'_{v^cv} & \Omega'_{v^cv^c} \end{pmatrix}/\{1, v\} = \Omega'_{v^cv^c} - \begin{pmatrix} \mathbf{u} & \Omega'_{v^cv} \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1}\begin{pmatrix} \mathbf{u}^T \\ \Omega'_{vv^c} \end{pmatrix}$$

Introducing the Laplacian submatrix $Q'_{v^cv^c} = Q_{\mathcal{V}^c\mathcal{V}^c}$ of G and expanding the matrix product in the right-hand side, this yields

$$(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1} = \frac{1}{2}\left(\mathbf{u}\Omega'_{vv^c} + \Omega'_{v^cv}\mathbf{u}^T - \Omega'_{v^cv^c}\right), \tag{3.17}$$

which completes the proof. □

In other words, the inverse Laplacian submatrix contains half the *triangle excess*[7] $\frac{1}{2}(\omega'_{iv}+\omega'_{vj}-\omega'_{ij})$ of the effective resistances off-diagonally, and the effective resistances ω'_{iv} from v to the rest of the nodes on the diagonal. We will also refer to this as a triangle excess matrix. Making use of the triangle inequality property of the effective resistance (see Chapter 4), this implies the following result:

Property 3.29. *The inverse Laplacian submatrix $(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}$ is a block-diagonal matrix with positive blocks for the connected components of $G\backslash\mathcal{V}$.*

Proof. Let $A_1, \dots, A_\ell$ be the node sets of the connected components of $G\backslash\mathcal{V}$. Then $Q_{\mathcal{V}^c\mathcal{V}^c}$ is block-diagonal on $A_1, \dots, A_\ell$ and thus its inverse $(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}$ is as well. Moreover, in the $\mathcal{V}-v$ merged graph G', the node v is a cut node between the same sets. By the triangle equality for effective resistances — Theorem 4.2 and Proposition 4.3 in Chapter 4, which say that $\omega_{ix}+\omega_{xj} \geq \omega_{ij}$ with equality if and only if x is a cut node between i and j — we then know that

$$\omega'_{iv}+\omega'_{vj}-\omega'_{ij}=0 \text{ if } i\in A_k \neq A_m \ni j \quad \text{and} \quad \omega'_{iv}+\omega'_{vj}-\omega'_{ij}>0 \text{ if } i,j\in A_k,$$

which confirms by Proposition 3.28 that $(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}$ is block diagonal on $A_1, \dots, A_\ell$ with positive entries within the blocks. □

Proposition 3.28 and Property 3.29 describe the properties of inverse Laplacian submatrices. If we shift the perspective to describing effective resistances in node-merged graphs, we can also reformulate Proposition 3.28 as follows:

Corollary 3.30. *The effective resistances in a $\mathcal{V}-v$ merged graph G' are given by*

$$\omega'_{ij}=(\mathbf{e}_i-\mathbf{e}_j)^T(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}(\mathbf{e}_i-\mathbf{e}_j) \quad \text{and} \quad \omega'_{iv}=\mathbf{e}_i^T(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}\mathbf{e}_i$$

for all $i,j\in\mathcal{V}^c$.

Proof. This follows immediately from Proposition 3.28 and in particular from the matrix expression (3.17) of the Laplacian submatrix inverse. □

The resistance matrix of the $\mathcal{V}-v$ merged graph G' can thus be summarized as

$$\Omega'=\begin{pmatrix}0 & \mathbf{q}^T\\ \mathbf{q} & \mathbf{q}\mathbf{u}^T+\mathbf{u}\mathbf{q}^T-2(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}\end{pmatrix} \text{ with } \mathbf{q}=\operatorname{diag}\left((Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}\right)$$

and we find a matrix expression similar to (3.1):

$$\Omega'_{v^cv^c}=\mathbf{q}\mathbf{u}^T+\mathbf{u}\mathbf{q}^T-2(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1} \text{ with } \mathbf{q}=\operatorname{diag}\left((Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}\right). \tag{3.18}$$

[7] This is also called the *Gromov product* of i and j with respect to v.

In particular, Corollary 3.30 is valid if the set $\mathcal{V} = \{v\}$ is a single node, for which the merged-node graph is simply $G' = G$. We give two example applications of this observation:

Example 3.31 (**effective resistances**)**.** Corollary 3.30 with the merged set $\mathcal{V}$ equal to a single node suggests two alternative definitions for the effective resistance:

$$\omega_{ij} = (\mathbf{e}_i - \mathbf{e}_j)^T (Q_{x^c x^c})^{-1} (\mathbf{e}_i - \mathbf{e}_j) \text{ for any } x \neq i, j \quad \text{and} \quad \omega_{ij} = [(Q_{i^c i^c})^{-1}]_{jj}$$

for any two nodes i, j. This definition can alternatively be derived by so-called 'grounding' of a node when solving the electrical equations; the voltage of some node x is then assumed to be zero, which eliminates the equation corresponding to the x^{th} row/column of the Laplacian matrix, and makes the Laplacian system invertible; see e.g. [78]. Finally, following (3.18), the elementary resistance submatrices satisfy:

$$\Omega_{i^c i^c} = \mathbf{q}\mathbf{u}^T + \mathbf{u}\mathbf{q}^T - 2(Q_{i^c i^c})^{-1} \text{ with } \mathbf{q} = \operatorname{diag}\left((Q_{i^c i^c})^{-1}\right).$$

Example 3.32 (**average resistance distance**)**.** From Corollary 3.30 with the merged set $\mathcal{V}$ equal to a single node, it follows that the average effective resistance from node i to the rest of the nodes (including i itself), and the Laplacian pseudoinverse diagonal entries are equal to

$$\langle \omega_i \rangle = \frac{1}{n} \operatorname{tr}\left((Q_{i^c i^c})^{-1}\right) \quad \text{and} \quad (Q^\dagger)_{ii} = \frac{1}{n^2} \mathbf{u}^T (Q_{i^c i^c})^{-1} \mathbf{u}.$$

Related to these expressions, we note that $\det(Q_{i^c i^c})$ is equal to the (weighted) number of spanning trees of a graph (independent of i); see for instance [11, §4.3],[37, 242]. This fact be used to prove Theorem 2.13.

While the inverse Laplacian submatrix $(Q_{x^c x^c})^{-1}$ is known to be a triangle excess matrix of the effective resistances in G (see for instance [78, Lem. 3.11]), the extension to merged graphs does not seem to be known as far as we know. In what follows, we make use of the Fiedler–Bapat identity in combination with the Schur complement to find new expressions for the effective resistances in merged graphs in terms of the original graph. We find the following result:

Proposition 3.33. *The inverse Laplacian submatrix is equal to*

$$(Q_{\mathcal{V}^c \mathcal{V}^c})^{-1} = -\frac{1}{2}\Omega_{\mathcal{V}^c \mathcal{V}^c} - \frac{1}{4} \begin{pmatrix} \mathbf{u} & \Omega_{\mathcal{V}^c \mathcal{V}} \end{pmatrix} \begin{pmatrix} 4\sigma''^2 & -2\mathbf{p}''^T \\ -2\mathbf{p}'' & Q/\mathcal{V}^c \end{pmatrix} \begin{pmatrix} \mathbf{u}^T \\ \Omega_{\mathcal{V}\mathcal{V}^c} \end{pmatrix}$$

with $\mathbf{p}''$ *and* σ''^2 *defined with respect to* $Q'' = Q/\mathcal{V}^c$.

Proof. Let G be a graph with a nonempty strict subset of nodes $\mathcal{V}^c$. The Fiedler–Bapat identity for G in block-matrix form reads

$$\begin{pmatrix} 0 & \mathbf{u}^T & \mathbf{u}^T \\ \mathbf{u} & \Omega_{\mathcal{V}\mathcal{V}} & \Omega_{\mathcal{V}\mathcal{V}^c} \\ \mathbf{u} & \Omega_{\mathcal{V}^c\mathcal{V}} & \Omega_{\mathcal{V}^c\mathcal{V}^c} \end{pmatrix}^{-1} = -\frac{1}{2} \begin{pmatrix} 4\sigma^2 & -2\mathbf{p}_{\mathcal{V}}^T & -2\mathbf{p}_{\mathcal{V}^c}^T \\ -2\mathbf{p}_{\mathcal{V}} & Q_{\mathcal{V}\mathcal{V}} & Q_{\mathcal{V}\mathcal{V}^c} \\ -2\mathbf{p}_{\mathcal{V}^c} & Q_{\mathcal{V}^c\mathcal{V}} & Q_{\mathcal{V}^c\mathcal{V}^c} \end{pmatrix}.$$

By the Schur complement property $(A^{-1})_{\mathcal{V}^c\mathcal{V}^c} = (A/\mathcal{V})^{-1}$ applied to the extended resistance matrix, we then find that

$$\begin{aligned} \left[-\frac{1}{2}Q_{\mathcal{V}^c\mathcal{V}^c}\right]^{-1} &= \begin{pmatrix} 0 & \mathbf{u}^T & \mathbf{u}^T \\ \mathbf{u} & \Omega_{\mathcal{V}\mathcal{V}} & \Omega_{\mathcal{V}\mathcal{V}^c} \\ \mathbf{u} & \Omega_{\mathcal{V}^c\mathcal{V}} & \Omega_{\mathcal{V}^c\mathcal{V}^c} \end{pmatrix} / (\mathcal{V} \cup \{1\}) \\ &= \Omega_{\mathcal{V}^c\mathcal{V}^c} - \begin{pmatrix} \mathbf{u} & \Omega_{\mathcal{V}^c\mathcal{V}} \end{pmatrix} \begin{pmatrix} 0 & \mathbf{u}^T \\ \mathbf{u} & \Omega_{\mathcal{V}\mathcal{V}} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{u}^T \\ \Omega_{\mathcal{V}\mathcal{V}^c} \end{pmatrix} \\ &= \Omega_{\mathcal{V}^c\mathcal{V}^c} + \frac{1}{2} \begin{pmatrix} \mathbf{u} & \Omega_{\mathcal{V}^c\mathcal{V}} \end{pmatrix} \begin{pmatrix} 4\sigma''^2 & -2\mathbf{p}''^T \\ -2\mathbf{p}'' & Q/\mathcal{V}^c \end{pmatrix} \begin{pmatrix} \mathbf{u}^T \\ \Omega_{\mathcal{V}\mathcal{V}^c} \end{pmatrix}, \\ \therefore (Q_{\mathcal{V}^c\mathcal{V}^c})^{-1} &= -\frac{1}{2}\Omega_{\mathcal{V}^c\mathcal{V}^c} - \frac{1}{4} \begin{pmatrix} \mathbf{u} & \Omega_{\mathcal{V}^c\mathcal{V}} \end{pmatrix} \begin{pmatrix} 4\sigma''^2 & -2\mathbf{p}''^T \\ -2\mathbf{p}'' & Q/\mathcal{V}^c \end{pmatrix} \begin{pmatrix} \mathbf{u}^T \\ \Omega_{\mathcal{V}\mathcal{V}^c} \end{pmatrix} \end{aligned}$$

where the third equality follows from the Fiedler–Bapat identity for $G/\mathcal{V}^c = G''$. □

We remark that the proof of Proposition 3.33 amounts to applying the Fiedler–Bapat identity twice, in combination with the Schur complement. Combining Corollary 3.30 and Proposition 3.33 for the effective resistances in a merged graph, we thus find:

Corollary 3.34. *The effective resistances in a $\mathcal{V} - v$ merged graph G' are given by*

$$\begin{aligned} \omega'_{ij} &= \omega_{ij} - \frac{1}{4} \sum_{a \sim'' b} c''_{ab} (\omega_{ia} + \omega_{jb} - \omega_{ib} - \omega_{ja})^2 \\ \omega'_{iv} &= -\sigma''^2 + \sum_{a \in \mathcal{V}} \omega_{ia} p''_a - \frac{1}{4} \sum_{a \sim'' b} c''_{ab} (\omega_{ia} - \omega_{ib})^2 \end{aligned}$$

where we use $a \sim'' b$ to denote that a and b are connected in $G'' = G/\mathcal{V}^c$ with weight c''_{ab}.

Proof. The result follows by combining the expression for effective resistances ω' in terms of the inverse Laplacian submatrix in Corollary 3.30 and the expression for the inverse Laplacian submatrix in Proposition 3.33; the quadratic forms determined by the Laplacian $Q/\mathcal{V}^c$ are expanded as $\mathbf{f}^T(Q/\mathcal{V}^c)\mathbf{f} = \sum_{a \sim'' b} c''_{ab}(f_i - f_j)^2$. □

We note that $(\omega_{ia} + \omega_{jb} - \omega_{ib} - \omega_{ja})^2 = [(\mathbf{e}_i - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_a - \mathbf{e}_b)]^2$ by expression (3.2). To illustrate the corollary, we calculate the effective resistances in a graph where a pair of nodes are merged:

Example 3.35 (**merging link**)**.** If the merged set consists of just two nodes a, b, we find that the $\{a, b\} - v$ merged graph G' has effective resistances

$$\omega'_{ij} = \omega_{ij} - \frac{(\omega_{ia} + \omega_{jb} - \omega_{ib} - \omega_{ja})^2}{4\omega_{ab}}$$
$$\omega'_{iv} = \frac{1}{2}(\omega_{ia} + \omega_{ib}) - \frac{1}{4}\left(\omega_{ab} + \frac{(\omega_{ia} - \omega_{ib})^2}{\omega_{ab}}\right),$$

for $i, j \neq a, b$. This follows since the Kron reduction $G'' = G/\{a, b\}^c$ is a single link (a, b) with link weight $c''_{ab} = \omega_{ab}^{-1}$, and as shown in Example 6.7 this graph has resistance curvature $\mathbf{p}'' = (1/2, 1/2)^T$ and resistance radius $\sigma''^2 = \omega_{ab}/4$. From the Cauchy–Schwarz inequality (or the maximum principle, Lemma 4.1), we find that $(\omega_{ia} + \omega_{jb} - \omega_{ib} - \omega_{ja})^2 \leq \omega_{ij}\omega_{ab}$ and thus $\omega'_{ij} \geq \frac{3}{4}\omega_{ij}$. By the triangle inequality (Theorem 4.2), we furthermore know that $|\omega_{ia} - \omega_{ib}| \leq \omega_{ab}$ which leads to $\omega'_{iv} \geq \frac{1}{2}(\omega_{ia} + \omega_{ib} - \omega_{ab})$.

We will not make any further use of the concept of merged nodes in the rest of this thesis. We note, however, that this concept is relevant when extending the definition of weighted graphs to take into account infinite link weights; a graph with $c_{ij} \to \infty$ is equivalent to the $\{i, j\} - v$ merged graph. This is also called 'shorting' in electrical circuit theory.

Finally, Laplacian submatrices can be characterized as a class of matrices without explicitly making reference to their being the submatrix of a Laplacian:

Property 3.36. *A matrix A is a Laplacian submatrix if and only if it can be written as $A = Q + V$ where Q is a Laplacian matrix and V a nonnegative diagonal matrix not equal to the zero matrix.*

Proof. Let $V = \operatorname{diag}(\mathbf{v})$. Then any matrix $A = Q + V$ is a submatrix of $\left(\begin{smallmatrix} \mathbf{u}^T\mathbf{v} & -\mathbf{v}^T \\ -\mathbf{v} & A \end{smallmatrix}\right)$ which is a Laplacian matrix. □

These matrices are also called Schrödinger matrices (operators) [17], loopy Laplacians [78], weakly diagonally-dominant M matrices [94], bottleneck matrices [145] or Stieltjes matrices [19]. Laplacian submatrices and their determinants play an important role in the context of spanning trees [11, §4.3].

3.4 The resistance radius

As introduced in the Fiedler–Bapat identity, the resistance radius is a resistance-based graph invariant. Following Corollary 3.10, the resistance radius can be defined

concisely as

$$\sigma^2 = \frac{1}{2}\left(\mathbf{u}^T\Omega^{-1}\mathbf{u}\right)^{-1}.$$

Section 6.2.3 discusses several alternative definitions and we note in particular (i) its relation to the resistance curvature by $\sigma^2 = \frac{1}{2}\mathbf{p}^T\Omega\mathbf{p}$, which will be further discussed in Chapter 5, (ii) the simple expression $\sigma^2 = \frac{1}{2}\sum_{j\sim i} c_{ij}(\omega_{ix} - \omega_{jx})^2$ for any x, as derived in Section 6.2.3, and (iii) its geometric interpretation as the (squared) radius of the simplex associated to a graph in Chapter 4: this last interpretation underlies the name 'resistance radius'. While the resistance radius appeared before in the work of Fiedler, Bapat and Subak-Sharpe, it has not yet been studied in detail and, as Fiedler noted, "*[i]t would be desirable to find the interpretation ... of* q_{00} *[*$4\sigma^2$*] in the graph-theoretical model.*" [96, p. 158].

Here we study how the resistance radius relates between different Kron reductions of a graph by considering the resistance radius as a function on subsets of the nodes. A function on a family of subsets is called a *set function* and appears for instance in combinatorics and measure theory.

Definition 3.37. The *resistance radius (set) function* is defined by

$$\sigma^2(\mathcal{V}) := \frac{1}{2}\left(\mathbf{u}^T[\Omega_{\mathcal{V}\mathcal{V}}]^{-1}\mathbf{u}\right)^{-1} \text{ for all } \mathcal{V} \subseteq \mathcal{N}, \tag{3.19}$$

and we define $\sigma^2(\emptyset) = \sigma^2(\{i\}) = 0$ for all $i \in \mathcal{N}$.

In other words, $\sigma^2(\mathcal{V})$ is the resistance radius of the $\mathcal{V}$-Kron reduction graph $G/\mathcal{V}^c$. As a quick example, one can check that this definition leads to $\sigma^2(\{i,j\}) = \omega_{ij}/4$ for any two nodes i, j based on the resistance matrix $\left(\begin{smallmatrix} 0 & \omega_{ij} \\ \omega_{ij} & 0 \end{smallmatrix}\right)$. This means that the resistance radius function fully characterizes the resistance matrix (consider all node pairs) and thus the graph. Before continuing, we recall Proposition 3.25, which says that the resistance radius of the Kron reduction of a graph is

$$\sigma^2(\mathcal{V}) = \sigma^2(\mathcal{N}) - \mathbf{p}_{\mathcal{V}^c}^T(Q_{\mathcal{V}^c\mathcal{V}^c})^{-1}\mathbf{p}_{\mathcal{V}^c} \text{ and thus } \sigma^2(\mathcal{N}\backslash\{x\}) = \sigma^2(\mathcal{N}) - \frac{p_x^2}{k_x}.$$

A family of sets carries a natural partial order given by the subset relation $A \leq B \Leftrightarrow A \subseteq B$ and a set function that respects this partial order, i.e. as $A \subseteq B \Rightarrow f(A) \leq f(B)$, is called a nondecreasing function. Following Proposition 3.25, we find:

Property 3.38. *The resistance radius* σ^2 *is a nondecreasing set function.*

Proof. Let $\emptyset \neq \mathcal{V} \subset \mathcal{N}$ be a nonempty subset of nodes. Then Proposition 3.25 says that

$$\sigma^2(\mathcal{V}) = \sigma^2(\mathcal{N}) - \mathbf{p}_{\mathcal{V}^c}^T (Q_{\mathcal{V}^c\mathcal{V}^c})^{-1} \mathbf{p}_{\mathcal{V}^c}.$$

Since $Q_{\mathcal{V}^c\mathcal{V}^c}$ is positive definite by Property 3.26, the second term is nonnegative ($\mathbf{p}_{\mathcal{V}^c}$ could be zero) and thus we find that $\sigma^2(\mathcal{V}) \leq \sigma^2(\mathcal{N})$. □

This property immediately implies that the range of the resistance radius function is determined by the smallest set ($\emptyset$) and the largest set ($\mathcal{N}$) as follows:

Property 3.39. *The resistance radius is nonnegative and satisfies* $0 \leq \sigma^2(\mathcal{V}) \leq \sigma^2(\mathcal{N})$ *with equality in the lower bound if and only if* $|\mathcal{V}| \leq 1$ *and equality in the upper bound if and only if* $\mathcal{V} = \mathcal{N}$ *or* $p_i = 0$ *for all* $i \in \mathcal{N} \setminus \mathcal{V}$.

Proof. The nondecreasing property of σ^2 implies the range $0 = \sigma(\emptyset) \leq \sigma^2(\mathcal{V}) \leq \sigma^2(\mathcal{N})$. Next, let $\mathcal{V}$ be any set with at least two nodes $|\mathcal{V}| \geq 2$ and pick two nodes i, j in $\mathcal{V}$. Then $\sigma^2(\mathcal{V}) \geq \sigma^2(\{i, j\}) = \omega_{ij}/4 > 0$ by the nondecreasing property and positivity of the effective resistance. Thus, zero resistance radius is achieved if and only if $|\mathcal{V}| \leq 1$ (in which case it holds by definition). Second, by Proposition 3.25 for the resistance radius in Kron reduction graphs, we have

$$\sigma^2(\mathcal{V}) = \sigma^2(\mathcal{N}) - \mathbf{p}_{\mathcal{V}^c}^T (Q_{\mathcal{V}^c\mathcal{V}^c})^{-1} \mathbf{p}_{\mathcal{V}^c}$$

for any $\emptyset \neq \mathcal{V} \subset \mathcal{N}$. Since $Q_{\mathcal{V}^c\mathcal{V}^c}$ is positive definite, the quadratic form is zero, i.e. $\sigma^2(\mathcal{V}) = \sigma^2(\mathcal{N})$ if and only if $\mathbf{p}_{\mathcal{V}^c} = 0$. This completes the proof. □

Another simple lower-bound follows directly from Property 3.38: $\sigma^2(\mathcal{V}) \geq \frac{1}{4} \max_{i,j \in \mathcal{V}} \omega_{ij}$ (see also Proposition 5.17). A second class of results deals with pairs of subsets that are not necessarily related by the subset partial order. A set function f satisfies the *inclusion–exclusion principle* if

$$f(A \cup B) = f(A) + f(B) - f(A \cap B) \text{ for all } A, B \subseteq \mathcal{N}.$$

The inclusion–exclusion property lies at the basis of many results in combinatorics and probability theory and formalizes an intuitive property of notions like the cardinality or size of an object [54, Ch. 4]. While σ^2 does not satisfy inclusion–exclusion in general[8], we find the following partial result:

[8] A counterexample follows from the unit-weight complete graph K_3 with nodes $\mathcal{N} = \{1, 2, 3\}$. Let $A = \{1, 2\}$ and $B = \{2, 3\}$. For these sets, we calculate $\sigma^2(A \cup B) = 2/9$ and $\sigma^2(A) = \sigma^2(B) = 1/6$ and $\sigma^2(A \cap B) = 0$, and thus obtain the strict inequality:

$$\sigma^2(A \cup B) = 2/9 < 2/6 = \sigma^2(A) + \sigma^2(B) - \sigma^2(A \cap B).$$

Proposition 3.40. *Let $A, B \subseteq \mathcal{N}$ be two node subsets such that removing $A \cap B$ disconnects A from B. Then*

$$\sigma^2(A \cup B) = \sigma^2(A) + \sigma^2(B) - \sigma^2(A \cap B).$$

Proof. Let $X, Y, Z \subseteq \mathcal{N}$ be three disjoint subsets of nodes such that removing Y disconnects X from Z, and define $A = X \cup Y$ and $B = Z \cup Y$ such that $A \cap B = Y$. If $X = 0$ (or $Z = 0$) then $A \subseteq B$ (or $B \subseteq A$) and the proposition holds trivially, so we may assume that X, Z are non-empty. Furthermore, Y is non-empty since the graph is connected.

Following Property 3.29 on the inverse Laplacian submatrix, we know that since Y is a cut node set between X and Z, the matrix $(Q_{Y^cY^c})^{-1}$ is block diagonal and equal to

$$(Q_{Y^cY^c})^{-1} = \begin{pmatrix} (Q_{XX})^{-1} & 0 \\ 0 & (Q_{ZZ})^{-1} \end{pmatrix}. \tag{3.20}$$

Furthermore, any path between two nodes $x, x' \in X$ that goes through Z also goes through Y (since Y is an $X - Z$ cut) and thus there exist no $x - x'$ path with all nodes in Z. By the weighted version of Proposition 3.19, this means that the $(X \cup Y)$-Kron reduction $G' = G/Z$ of G does not change any links or link weights between nodes in X. Moreover, since X has no neighbours in Z, we have $Q_{XZ} = 0$ and thus we know by Proposition 3.25 that the resistance curvature of nodes in X are left unchanged in G. To summarize, we have $Q'_{XX} = Q_{XX}$ and $p'_x = p_x$ for all $x \in X$.

We can derive the same results for the nodes in Z with respect to the $(Z \cup Y)$-Kron reduction $G'' = G/X$ and we thus have

$$Q'_{XX} = Q_{XX},\ Q''_{ZZ} = Q_{ZZ} \quad \text{and} \quad p'_x = p_x,\ p''_z = p_z \text{ for all } x \in X, z \in Z. \tag{3.21}$$

Now we apply Proposition 3.25 for the resistance curvature in Kron reduction graphs, and find:

$$\begin{aligned}
\sigma^2(X \cup Y \cup Z) - \sigma^2(Y) &= \begin{pmatrix} \mathbf{p}_X \\ \mathbf{p}_Z \end{pmatrix}^T (Q_{Y^cY^c})^{-1} \begin{pmatrix} \mathbf{p}_X \\ \mathbf{p}_Z \end{pmatrix} \\
&= \mathbf{p}_X^T (Q_{XX})^{-1} \mathbf{p}_X + \mathbf{p}_Z^T (Q_{ZZ})^{-1} \mathbf{p}_Z \text{ (by eq. (3.20))} \\
\sigma^2(X \cup Y) - \sigma^2(Y) &= \mathbf{p}'^T_X (Q'_{XX})^{-1} \mathbf{p}'_X \\
&= \mathbf{p}_X^T (Q_{XX})^{-1} \mathbf{p}_X \text{ (by eq. (3.21))} \\
\sigma^2(Y \cup Z) - \sigma^2(Y) &= \mathbf{p}''^T_Z (Q''_{ZZ})^{-1} \mathbf{p}''_Z \\
&= \mathbf{p}_Z^T (Q_{ZZ})^{-1} \mathbf{p}_Z \text{ (by eq. (3.21))}
\end{aligned}$$

Summing these three expressions and introducing the sets $A = X \cup Y$ and $B = Y \cup Z$ retrieves the inclusion–exclusion property and completes the proof. □

We note that Proposition 3.40 is significantly strengthened in the case of graphs with $\mathbf{p} \geq 0$, as shown in Theorem 6.33 in Chapter 6.

By a proper choice of node sets, the resistance radius of a graph can thus be calculated as a sum over the resistance radii of smaller graphs. For example, if two sets A, B overlap in a single cut node x, then the inclusion–exclusion principle implies that the resistance radius can be decomposed as

$$\sigma^2(A \cup B) = \sigma^2(A) + \sigma^2(B) \text{ if } A \cap B = \{x\} \text{ is an } A - B \text{ cut node}$$

since $\sigma^2(\{x\}) = 0$ by definition. Recursively removing cut nodes from the remaining connected components then directly leads to the following result.

Corollary 3.41. *The resistance radius of a graph is the sum of the resistance radii of its biconnected components.*

We recall that a biconnected component is a maximal subgraph without cut nodes. We now consider a number of examples.

Example 3.42 (**tree graphs**)**.** In tree graphs, each link is a biconnected component with resistance radius $\sigma^2(\{i,j\}) = \frac{1}{4}\omega_{ij} = \frac{1}{4}c_{ij}^{-1}$. By Corollary 3.41, we thus have:

Proposition 3.43. *The resistance radius of a tree is equal to* $\sigma^2(\mathcal{N}) = \sum_{i\sim j}(4c_{ij})^{-1}$*; for unweighted trees* ($c = 1$)*, this equals* $\sigma^2(\mathcal{N}) = \frac{1}{4}(n-1)$.

In other words, the resistance radius is equal for all (unweighted, connected) trees on the same number of nodes.

Example 3.44 (**node transitive graphs**)**.** In Example 6.10, we will discuss node transitive graphs, which are highly symmetric graphs in which the nodes are 'indistinguishable'. Proposition 6.11 shows that the resistance curvature is constant as a result of this symmetry, and equal to $\mathbf{p} = \mathbf{u}/n$. Consequently, the resistance radius of a node transitive graph is

$$\sigma^2(\mathcal{N}) = \frac{1}{2n^2}\mathbf{u}^T\Omega\mathbf{u} = \langle\omega\rangle = \frac{1}{n^2}Kf,$$

where we recall that Kf is the Kirchhoff index of the graph, whose exact value is known for many graphs. For instance, for the complete graph, we find $\sigma^2(\mathcal{N}) = (n-1)/n^2$ and for cycle graphs we find $\sigma^2(\mathcal{N}) = (n^2-1)/(12n)$, see [166]. This shows that the resistance radius can span several orders of magnitude, from $O(n^{-1})$ for the complete graph to $O(n)$ for cycles and tree graphs.

To conclude, we remark that the resistance radius is formally related to the concept of *magnitude*. Tom Leinster introduced magnitude as an Euler characteristic-like invariant[9] for a variety of mathematical objects, including metric spaces [154, 156]. In the finite case, magnitude can be defined based on a similarity matrix Z, which is a matrix with unit diagonal and nonnegative off-diagonal. For instance, for a metric space (X, d), the similarity matrix is constructed as[10] $(Z_t)_{ij} = \exp(-d(i,j)t)$ at scale t. The magnitude is defined based on this similarity matrix follows: if the equation $Z\mathbf{w} = \mathbf{u}$ has a solution $\mathbf{w}$, then this is called the weighting vector and $|Z| := \mathbf{u}^T\mathbf{w}$ is called the magnitude. For invertible similarity matrices, we thus have $|Z| = \mathbf{u}^T Z^{-1}\mathbf{u}$ in a form similar to the (inverse) resistance radius. We recall that for an invertible similarity matrix Z with nonzero magnitude, we may define the matrix $-2Z^{-1} + 2|Z|^{-1}Z^{-1}\mathbf{u}\mathbf{u}^T Z^{-1}$, which satisfies a Fiedler–Bapat-like identity with the similarity matrix. Section 5.3.1 will discuss some relations between the resistance radius and 'diversity', which is a concept related to magnitude.

[9]The formal relation between σ^2 and magnitude in combination with the latter's interpretation as an Euler characteristic-like invariant was our original inspiration to study the inclusion–exclusion property of the resistance radius function, with Proposition 3.40 and Theorem 6.33 as a result.

[10]This relates similarity matrices to distance matrices, as discussed in Chapter 5, since for small t, we have $Z_t \approx \mathbf{u}\mathbf{u}^T - Dt$ for distance matrix D.

Chapter 4

Resistance distance and simplices

This chapter deals with the metric aspect of effective resistances, following the interpretation of ω as the resistance distance. As mentioned in the introduction (Chapter 1), this is one of its most appreciated properties and underlies many applications. Less commonly known but perhaps even more far-reaching is that this metric perspective is the starting point for a geometric characterization of the effective resistance, and by extension weighted graphs, in terms of 'hyperacute simplices' in $\mathbb{R}^{n-1}$ (see Definition 4.17). This results in a geometric perspective on graphs and effective resistances where, in Fiedler's words, "*every geometric invariant of the simplex is at the same time an invariant of the graph*" [93, p. 73] and such that, according to Subak-Sharpe, "*all the arsenal of knowledge, acquired over many years, which we posses [sic] on Euclidean spaces may therefore be utilised in the solution of the resistive n-port problem [related to characterizing properties of the resistance matrix]*" [213, Abs.].

Section 4.1 starts by giving a proof of the metric property of ω based on the maximum principle, and discusses some related results.

In Section 4.2, we classify the resistance distance as a strict negative type metric (Proposition 4.9), which is a common class of metrics that includes, for instance, subsets of hyperbolic and spherical spaces. This classification positions the effective resistance in a 'hierarchy' of metric spaces (Proposition 4.22), where for each step upwards in the hierarchy, some constraints are relaxed.

In addition to this metric classification, Section 4.2.2 discusses the geometric characterization of the resistance distance as the squared Euclidean distance between the vertices of a hyperacute simplex S (Theorem 4.19).

Section 4.3.1 shifts the focus back to graphs and introduces Fiedler's graph–simplex correspondence which describes a bijection $G \leftrightarrow S$ between weighted graphs

and hyperacute simplices in Theorem 4.23; even stronger, this correspondence is 'functorial' and associates Kron relations $G \to G'$ with face relations $S \to S'$ (Theorem 4.24). As a first application, we give a new geometric proof of the metric property of the effective resistance based on subset-closure of hyperacute simplices (Theorem 4.20). As a second application, we introduce the resistive embedding φ, which maps functions on the graph to the ambient space of the simplex.

The main reference for this chapter is Fiedler's work on simplices and their relation to graphs [96]. For the resistance distance, see [119, 149] and for (strict) negative type metrics and distance geometry, the standard reference is Blumenthal's book [25]; see also [160, 174, 210].

4.1 Resistance is distance

In order to prove that the effective resistance is indeed a metric, we start by establishing the *maximum principle* as an auxiliary lemma:

Lemma 4.1 (maximum principle)**.** $(\mathbf{e}_y - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \leq \omega_{ij}$ for all $x, y, i, j \in \mathcal{N}$.

Proof. The lemma holds with equality if $i = j$ and if $(y, x) = (i, j)$, thus we may further assume that $i \neq j$ and that $n \geq 3$. We will first show that the lemma holds for $y = i$.

($y = i$) Let G be a graph with Laplacian Q and a pair of nodes $i \neq j$, and define the set of nodes

$$\mathcal{X} := \left\{ x \in \mathcal{N} : (\mathbf{e}_i - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \geq (\mathbf{e}_i - \mathbf{e}_z)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \text{ for all } z \in \mathcal{N} \right\}.$$

We note that this definition can be rewritten as $(\mathbf{e}_z - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \geq 0$ for all $x \in \mathcal{X}$ and all z, with equality if and only if $z \in \mathcal{X}$. Let $x \in \mathcal{X}$. Then either $z \in \mathcal{X}$ for all neighbours $z \sim x$, or x has at least one neighbour $x \sim z' \notin \mathcal{X}$ and then

$$\sum_{z : z \sim x} c_{xz} (\mathbf{e}_z - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) > 0,$$

since link weights are positive. With the definition of the Laplacian matrix, this can be rewritten as

$$-\mathbf{e}_x^T Q Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) > 0 \Leftrightarrow -\mathbf{e}_x^T \left(I - \tfrac{\mathbf{u}\mathbf{u}^T}{n} \right) (\mathbf{e}_i - \mathbf{e}_j) = -\mathbf{e}_x^T (\mathbf{e}_i - \mathbf{e}_j) > 0 \Leftrightarrow x = j.$$

In other words, if not all neighbours of $x \in \mathcal{X}$ are in $\mathcal{X}$, then $x = j$. The set $\mathcal{X}$ cannot be empty by its definition and it cannot contain all the nodes — this would

mean that $i \in \mathcal{X}$, but then we have the contradiction $0 = (\mathbf{e}_i - \mathbf{e}_i)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \geq (\mathbf{e}_i - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) > 0$. Since G is connected, there must be at least one node with not all neighbours in $\mathcal{X}$ and thus $j \in \mathcal{X}$ such that we get (by definition of $\mathcal{X}$)

$$\omega_{ij} = (\mathbf{e}_i - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \geq (\mathbf{e}_i - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \text{ for all } x \in \mathcal{N},$$

and thus that the Lemma holds for $y = i$. *(general y)* If $y \neq i$, we can write

$$\begin{aligned}&(\mathbf{e}_y - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \\ &= (\mathbf{e}_y - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) + (\mathbf{e}_j - \mathbf{e}_i)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) + (\mathbf{e}_i - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \leq \omega_{ij}\end{aligned}$$

where the inequality follows from the case where $y = i$. This concludes the proof. □

The proof shows that the maximum principle follows from basic properties of the Laplacian and its pseudoinverse and, crucially, positivity of the link weights. The name of the lemma refers to the related maximum principle of harmonic functions (solutions to the Laplace equation) [164, Ch. 4], [32, Ch. 9].

The maximum principle has the following physical interpretation in terms of currents and voltages: if we consider a unit current flow from i to j then the voltages in the network are given by the vector $Q^\dagger(\mathbf{e}_i - \mathbf{e}_j)$; this is also called a 'harmonic function'. The maximum principle then says that the voltage difference between i and j is larger than the voltage difference between any two other nodes x and y. In electrical circuit theory, this principle is also called the no-amplification property; see [179, Eq. 6]. From the maximum principle, it follows that ω is a metric, as first shown in [119, 149].

Theorem 4.2. *The effective resistance is a metric.*

Proof. Recall from Definition 2.11 that a metric is symmetric, nonnegative and zero only between identical elements and satisfies the triangle inequality. By its definition, the effective resistance ω_{ij} is symmetric and by properties of the pseudoinverse Laplacian quadratic form, it is nonnegative and equal to zero if and only if $i = j$. It thus remains to show the triangle inequality. For any $i, j, x \in \mathcal{N}$, we can write

$$\omega_{ij} = (\mathbf{e}_i - \mathbf{e}_x)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) + (\mathbf{e}_x - \mathbf{e}_j)^T Q^\dagger (\mathbf{e}_i - \mathbf{e}_j) \leq \omega_{ix} + \omega_{xj}$$

following the maximum principle. This completes the proof. □

The effective resistance as a metric is also called the *resistance distance*. We now proceed with showing two further results related to the resistance distance: when triangle equalities occur and a comparison with the shortest-path distance.

Lemma 4.1 implies that the triangle inequality holds for effective resistances, but it does not explain in which cases this inequality is strict or when it holds with equality. The forward direction of the following result was shown in [149]:

Proposition 4.3 (triangle equality)**.** *$\omega_{ij} = \omega_{ix} + \omega_{xj}$ if and only if x is a cut node between i and j.*

Proof. *(forward direction)* Let G be a graph with cut node x between i and j and consider the Kron reduction $G' := G/\{i,j,x\}^c$. By Proposition 3.20, we know that x is also a cut node in G'. Since G' is connected, it must thus be a path graph with end nodes i, j and cut node x. Since (i,x) and (x,j) are cut links, we have $\omega_{ix} = 1/c'_{ix}$ and $\omega_{jx} = 1/c'_{jx}$. Now consider the Kron reduction $G'' := G'/\{x\}$, which is a K_2 graph on i,j with $\omega_{ij} = 1/c''_{ij}$. By properties of the elementary Kron reduction $G' \to G''$, we find that

$$c''_{ij} = \frac{c'_{ix}c'_{xj}}{k'_x} = \frac{c'_{ix}c'_{xj}}{c'_{ix}+c'_{xj}} \quad \text{or} \quad \frac{1}{c''_{ij}} = \frac{1}{c'_{ix}} + \frac{1}{c'_{xj}} \Rightarrow \omega_{ij} = \omega_{ix} + \omega_{xj}$$

as required.

(converse direction) Let G be a graph and i, x, j three nodes for which the triangle equality holds with respect to x and consider the Kron reduction $G' := G/\{i,j,x\}^c$. As calculated above, the path graph with cut node x and link weights $c'_{ix} = \omega_{ix}^{-1}$ and $c'_{jx} = \omega_{jx}^{-1}$ is a graph realization that is consistent with the effective resistances between i, j, x that satisfy the triangle equality with respect to x. Since the effective resistances in G' fully determine the graph (e.g. by the Fiedler–Bapat identity) this means that G' must be the path graph with cut node x. By Proposition 3.20 on cut nodes in the Kron reduction, this confirms that x must be a cut node in G as well, which completes the proof. □

Remark 4.4. An alternative proof of the forward direction of Proposition 4.3 is as follows: let x be a cut node between i and j. Then $Q_{x^c x^c}$ is block diagonal with i and j in two separate blocks. Consequently, also $(Q_{x^c x^c})^{-1}$ is block diagonal with i, j in two different blocks and thus

$$0 = \left[(Q_{x^c x^c})^{-1}\right]_{ij} \overset{\text{(Prop. 3.28)}}{=} \frac{1}{2}(\omega_{ix} + \omega_{xj} - \omega_{ij}),$$

as required.

In particular, this result implies that all effective resistances in a graph are determined by the effective resistances within each of the biconnected components and

that the effective resistance between two nodes i, j only depends on those nodes and links that are contained in some path between i and j.

We now note a comparison result with the shortest-path distance on graphs. For weighted graphs, where the link weight c_{ij} is a measure of 'affinity' or 'closeness', the inverse weight c_{ij}^{-1} can be taken as the length of a link[1]. The weighted shortest-path distance is then equal to

$$d_{\mathrm{sp}}(i,j) := \min_P \sum_{(i,j)\in P} c_{ij}^{-1} = \min_P c(P)^{-1} \text{ over all } i-j \text{ paths } P,$$

which corresponds to the standard definition in terms of the number of links between two nodes in the case of unweighted graphs (with $c = 1$). The subscript 'sp' will sometimes be omitted. We find the following relation with effective resistances [11, 148]:

Proposition 4.5. *The effective resistance is smaller than or equal to the weighted shortest-path distance $\omega_{ij} \leq d_{sp}(i,j)$ with equality if and only if there is a unique path between i and j.*

Proof. Let P be a shortest weighted path between i and j with ordered nodes $i = 1 \sim 2 \sim \cdots \sim \ell = j$. By the triangle inequality for effective resistances and the resistance upper bound $\omega_{ij} \leq c_{ij}^{-1}$, we have

$$\omega_{ij} \leq \sum_{k=1}^{\ell-1} \omega_{k(k+1)} \leq \sum_{k=1}^{\ell-1} c_{k(k+1)}^{-1} = d_{\mathrm{sp}}(i,j),$$

with equality in the first (triangle) inequality if and only if all nodes in P except i and j are cut nodes, and equality in the second inequality ($\omega_{ij} \leq c_{ij}^{-1}$) if and only if all links in P are cut links; this corresponds to a unique path between i and j, which completes the proof. □

We note in particular that the effective resistance and the shortest-path distance are equal in tree graphs. The upper-bound in Proposition 2.10 generalizes Proposition 4.5, since it can take into account any independent set of paths and their weights instead of just the shortest path.

[1] While other monotonically decreasing function are possible, this particular choice of length leads to connections with the effective resistance; see for instance Proposition 4.5.

4.2 Classification and geometry of the resistance metric

We recall that a general metric space is denoted by (X, d), where X is a set of n elements and $d : X \times X \to [0, \infty)$ a metric between pairs of elements in X. This metric space can also be characterized by its $n \times n$ *distance matrix* D with entries given by the pairwise distances $(D)_{ij} = d(i, j)$. Since ω is a metric between the nodes of a graph, we can define the following metric space:

Definition 4.6 (resistance metric space)**.** A *resistance metric space* (X, d) is a finite metric space where $X = \mathcal{N}$ is the set of nodes of a graph and $d = \omega$ is the resistance distance.

Definition 4.6 is constructive: let G be any graph. Then we can calculate the resistance matrix, which is the distance matrix of the resistance metric space. It is less obvious, however, how to decide whether a given metric (space) is a resistance metric (space). One solution is given by our new characterization of resistance matrices in Theorem 3.9, by which we can recognize resistance metric spaces based on their distance matrices:

Proposition 4.7. *A metric space (X, d) is a resistance metric space if and only if its distance matrix D is invertible and satisfies $(\mathbf{u}^T D^{-1} \mathbf{u})(D^{-1})_{ij} \geq (D^{-1}\mathbf{u}\mathbf{u}^T D^{-1})_{ij}$ for all $i \neq j$.*

A second question following Definition 4.6 is whether resistance metric spaces can be classified in terms of known types of metric spaces. This question is answered starting in Section 4.2.1 below by showing that resistance metric spaces are a subclass of strict negative type metric spaces and that they can be characterized in terms of a particular class of geometric objects in Euclidean space.

4.2.1 Negative type metrics

Metric spaces of negative type first appeared in the work of Schoenberg in distance geometry [25, 210] and due to their nice properties (e.g. their relation to Euclidean spaces), they are used in various applications [5, 45, 168]. A negative type metric space is defined as follows:

Definition 4.8 ((strict) negative type)**.** A metric space (X, d) is/has *negative type* if its distance matrix satisfies

$$\mathbf{f}^T D \mathbf{f} \leq 0 \text{ for all } \mathbf{f} \in \operatorname{span}(\mathbf{u})^{\perp}. \tag{4.1}$$

If the inequality is strict for all nonzero $\mathbf{f} \in \text{span}(\mathbf{u})^\perp$, it is/has *strict negative type*.

We also say that (X, d) is a (strict) negative type metric space or that d is a (strict) negative type metric; if a metric space has strict negative type, it clearly also has negative type. While we only consider the finite case here, the theory of negative type metrics extends to infinite spaces. Meckes listed a number of examples in [173], which include some very common spaces: any subset of points X in Euclidean space, hyperbolic space or on the sphere, with their respective distances d is either a negative type or strict negative type metric space. Not all metrics are negative type however, and the shortest-path distance on the complete bipartite graph[2] $K_{2,3}$ is an example of a metric does not have negative type.

In practice, one can check whether a metric space with distance matrix D is negative type by calculating the *centered distance matrix* $D_c := (I - \frac{\mathbf{u}\mathbf{u}^T}{n})^T D (I - \frac{\mathbf{u}\mathbf{u}^T}{n})$. If this matrix is negative semidefinite, then $\mathbf{f}^T D \mathbf{f} = \mathbf{f}^T D_c \mathbf{f} \leq 0$ for all $\mathbf{f} \in \text{span}(\mathbf{u})^\perp$, which means that the metric space has negative type by Definition 4.8. If the centered distance matrix is negative semidefinite with kernel $\text{span}(\mathbf{u})$, then the metric space has strict negative type. For the resistance metric with resistance matrix as distance matrix, we thus find:

Proposition 4.9. *The resistance metric is a strict negative type metric.*

Proof. The resistance matrix satisfies $\mathbf{f}^T \Omega \mathbf{f} = -2\mathbf{f}^T Q^\dagger \mathbf{f} < 0$ for nonzero $\mathbf{f} \in \text{span}(\mathbf{u})^\perp$. □

This result is implicit in the work of Fiedler [96] and Subak-Sharpe & Moore [179].

The following example shows that Proposition 4.9 does not give a full classification of resistance metric spaces, i.e. there exist strict negative type metric spaces which are not resistance metric spaces.

Example 4.10 (resistance metric $\subsetneq$ strict negative type metric). Consider the matrix[3]

$$D = \begin{pmatrix} 0 & 20 & 11 & 11 \\ 20 & 0 & 11 & 11 \\ 11 & 11 & 0 & 8 \\ 11 & 11 & 8 & 0 \end{pmatrix}.$$

This matrix is symmetric, has a zero diagonal and the triangle inequality can be verified, so it is a distance matrix. Furthermore, the eigenvalues of the centered distance matrix are equal to $(0, -8, -8, -20)$ which means that it is a strict negative

[2] This is the 5-node graph with links $\{(i,j)\}_{1 \leq i \leq 2,\, 3 \leq j \leq 5}$.

[3] This example appeared in [70] with a different scaling.

type metric space. If we try to construct the corresponding Laplacian matrix following expression (3.9) however, we find that

$$-2D^{-1} + 2\frac{D^{-1}\mathbf{u}\mathbf{u}^T D^{-1}}{\mathbf{u}^T D^{-1}\mathbf{u}} = \frac{1}{80}\begin{pmatrix} 9 & 1 & -5 & -5 \\ 1 & 9 & -5 & -5 \\ -5 & -5 & 15 & -5 \\ -5 & -5 & -5 & 15 \end{pmatrix}.$$

Since the $(1, 2)$ entry is positive, we know from Proposition 4.7 that D is not a resistance matrix. This shows an example of a negative type metric space that is not a resistance metric space.

We thus know that the resistance metric space is a strict subclass of strict negative type metric spaces, which means that the former must carry some additional structure or constraints compared to the latter. This additional structure is most clearly seen by following the geometric perspective on these metric spaces.

One of the fundamental properties of negative type metric spaces is that they have a corresponding Euclidean embedding [25, 211]:

Property 4.11. *Let (X, d) be a metric space of negative type. Then there exists an embedding $\phi : X \to \mathbb{R}^{n-1}$ such that $\|\phi(i) - \phi(j)\|^2 = d(i, j)$.*

In other words, there exists an *isometric embedding* ϕ of $(X, \sqrt{d})$ into Euclidean space. Further on (below Proposition 4.13), we will show how this embedding follows from a simple transformation of the distance matrix into a Gram matrix. We remark that the embedding ϕ is not unique; if $\tau : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ is a *rigid transformation* of Euclidean space — a transformation that leaves distances invariant, which can be generated by combinations of rotations, translations and reflections — then the composition $\phi' = \tau \circ \phi$ is again a valid embedding. The dimension of the embedding space can be smaller than $n - 1$ if the embedded points are not affinely independent (see Definition 4.14).

Property 4.11 allows one to characterize (strict) negative type metric spaces in terms of the geometry of the embedded points $\phi(X)$. For negative type metric spaces, for instance, this already gives a complete characterization [25, 211]:

Proposition 4.12. *A metric space (X, d) has negative type if and only if there exists an isometric embedding of $(X, \sqrt{d})$ into $\mathbb{R}^{n-1}$.*

See Section 4.2.2 for a proof. Any negative type metric space can thus be characterized uniquely (up to rigid transformations) by some set of n points in which any

three points determine a nonobtuse triangle — i.e. a triangle with acute (smaller than $\pi/2$) or right (equal to $\pi/2$) angles. This angle requirement follows from the triangle inequality for d, which implies that

$$\|\phi(i) - \phi(j)\|^2 \leq \|\phi(i) - \phi(x)\|^2 + \|\phi(x) - \phi(j)\|^2 \text{ for all } i, x, j.$$

By trigonometry (e.g. the cosine rule), this inequality implies that $\phi(i), \phi(x), \phi(j)$ are the vertices of a triangle with a nonobtuse angle at x. As an example, the squared distance between the vertices of a hypercube or the vertices of a simplex (see Definition 4.14) forms a negative type metric space.

In Section 4.2.2, we show a similar characterization of strict negative type metric spaces and resistance metric spaces, based on further constraints on the geometric object determined by the points $\phi(X)$.

4.2.2 Hyperacute simplices

Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be a set of n points (vectors) in $\mathbb{R}^{n-1}$. The *Gram matrix* of these points is the $n \times n$ matrix with entries $\mathbf{b}_i^T \mathbf{b}_j$. To characterize these points up to rigid transformations, we define the *centered Gram matrix* M

$$M := \left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right) B^T B \left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right) \text{ where } B = [\mathbf{b}_1 \ \ldots \ \mathbf{b}_n],$$

which is invariant with respect to rigid transformations of the point set $B \to OB + \mathbf{x}\mathbf{u}^T$ for some orthogonal matrix O (which implies that $O^T O = I$) and translation $\mathbf{x} \in \mathbb{R}^{n-1}$. We remark that 'centered Gram matrix' is not standard terminology in the context of distance geometry.

A centered Gram matrix is characterized as follows:

Proposition 4.13. *A matrix M is a centered Gram matrix if and only if it is positive semidefinite with $M\mathbf{u} = 0$.*

Proof. By its definition as the product between a matrix with zero row sum and its transpose, a centered Gram matrix is positive semidefinite with zero row sum. Conversely, if M is a positive semidefinite matrix, then there exists a matrix B such that $M = B^T B$ (see Proposition A.4), and thus M is the Gram matrix of the n columns of B. Since $M\mathbf{u} = 0$, we know that $B\mathbf{u} = 0$ and thus that M can also be written as $M = (I - \mathbf{u}\mathbf{u}^T/n)B^T B(I - \mathbf{u}\mathbf{u}^T/n)$, the centered Gram matrix of the columns of B. □

For a set of points $\mathbf{b}_1, \ldots, \mathbf{b}_n$, the centered Gram matrix M (with entries $(M)_{ij} = \mathbf{b}_i^T \mathbf{b}_j$) and the squared Euclidean distance matrix E (with entries $(E)_{ij} := \|\mathbf{b}_i - \mathbf{b}_j\|^2$) are related as follows:

$$M = -\frac{1}{2}\left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right) E \left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right) \text{ and } E = \mathbf{u}\,\mathrm{diag}(M)^T + \mathrm{diag}(M)\mathbf{u}^T - 2M. \quad (4.2)$$

These relations show that the spectral properties of M and E are closely related, which will be crucial in describing the geometry of (strict) negative type metric spaces. For instance, this leads to a proof of Property 4.11.

Proof of Property 4.11 The negative type condition (4.1) on a distance matrix D implies that the centered distance matrix D_c is a negative semidefinite matrix with $D_c\mathbf{u} = 0$. By Proposition 4.13, this means that $-\frac{1}{2}D_c$ is a centered Gram matrix and thus that D is a squared Euclidean distance matrix as required. □

If condition (4.1) is strict, this further restricts D_c to have rank $n-1$, which in turn implies that the embedded points determine a 'simplex'.

A *simplex* is a generalization of triangles to higher dimensions, where 'three points not on a line' is generalized to the notion of *affine independence*. More precisely, a simplex is defined as follows:

Definition 4.14 (simplex and affine independence)**.** An *n-simplex* S is the convex hull of n points $\mathbf{b}_1, \ldots, \mathbf{b}_n \in \mathbb{R}^{n-1}$

$$S := \left\{ \sum_{i=1}^{n} \theta_i \mathbf{b}_i \text{ with } \theta_i \geq 0 \text{ and } \sum_{i=1}^{n} \theta_i = 1 \right\},$$

where the points are *affinely independent*: for every j with $1 \leq j \leq n$, the vectors $\{\mathbf{b}_i - \mathbf{b}_j\}_{i \neq j}$ are linearly independent.

Figure 4.1 shows four of simplices in low dimensions. A simplex is a convex polytope $S \subset \mathbb{R}^{n-1}$ and while S could live in (a subspace of) a higher-dimensional space, at least $n-1$ dimensions are required for affine independence to hold. The n points $\{\mathbf{b}_i\}_{i=1}^n$ that determine a simplex are called the *vertices* of the simplex.

Simplices are used for instance as the building blocks of *triangulations*, where a space is approximated by a collection of simplices glued together at their faces. Triangulations of manifolds appear for instance in discretizations of general relativity in the so-called Regge calculus [106, 205], in theories of discrete exterior calculus [69, 126] and in finite element calculations of partial differential equations [32, 33].

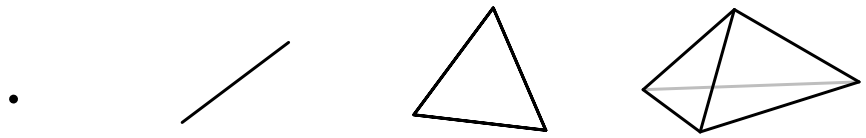


Figure 4.1: Example of simplices in $0, 1, 2$ and 3 dimensions.

Similarly, a topological space can be triangulated using abstract simplices[4] glued together at their faces; this gives rise to a *simplicial complex*, which is a central object in algebraic topology [124] and topological data analysis [44, 105].

How can we determine whether a given set of points $\mathbf{b}_1, \ldots, \mathbf{b}_n$ are the vertices of a simplex, i.e. whether they are affinely independent? The following result shows that affine independence translates to centered Gram matrices with maximal rank [96, Thm. 1.4.7]:

Proposition 4.15. *A matrix M is the centered Gram matrix of a simplex if and only if it is positive semidefinite with* $\ker(M) = \operatorname{span}(\mathbf{u})$.

Proof. *(forward direction)* Let $M = (I - \mathbf{u}\mathbf{u}^T/n)^T B^T B (I - \mathbf{u}\mathbf{u}^T/n)$ be the centered Gram matrix of a simplex, i.e. with vertices given by the columns $\mathbf{b}_1, \ldots, \mathbf{b}_n$ of B. Then M is positive semidefinite because it is the product of a matrix and its transpose; and from

$$\mathbf{x}^T M \mathbf{x} = 0 \Leftrightarrow \left\| B\left(I - \tfrac{\mathbf{u}\mathbf{u}^T}{n}\right)\mathbf{x} \right\|^2 = 0 \Leftrightarrow \mathbf{x} \in \ker(B_c),$$

where we define the *right-centered matrix* $B_c := B(I - \mathbf{u}\mathbf{u}^T/n)$, we know that $\ker(M) = \ker(B_c)$. By construction, we know that $\operatorname{span}(\mathbf{u}) \subseteq \ker(B_c)$. Suppose for contradiction that this subset is strict. Then there must exist a nonzero vector $\mathbf{f} \in \operatorname{span}(\mathbf{u})^\perp$ for which $B_c\mathbf{f} = 0$. But then for any j, we can write $f_j = -\sum_{i \neq j} f_i$ (since $\mathbf{f}$ is a zero-sum vector) and thus

$$0 = \sum_i f_i \mathbf{b}_i = \sum_{i \neq j} f_i (\mathbf{b}_i - \mathbf{b}_j).$$

This is in contradiction with affine independence of the vertices, which means that no such vector $\mathbf{f}$ can exist and that $\ker(M) = \ker(B_c) = \operatorname{span}(\mathbf{u})$, as required.

[4]While geometric simplices are defined up to rigid transformations, the simplices in topological applications are allowed to deform non-rigidly as long as the affine independence of its vertices is not violated. As a result, one of the main 'algebraic' properties that characterizes such abstract/topological simplices is that they are subset-closed (since each subset of an affinely independent set is again affinely independent).

(converse direction) Let M be an $n \times n$ positive semidefinite matrix with $\ker(M) = \operatorname{span}(\mathbf{u})$ and thus in particular $\operatorname{rank}(M) = n - 1$. Since M is positive semidefinite, it can be written as the product $M = B_c^T B_c$, where the right-centered matrix B_c has n columns. We now show that the columns of B_c are affinely independent and thus the vertices of an n-simplex. As in the forward direction, we find that $\ker(B_c) = \ker(M) = \operatorname{span}(\mathbf{u})$. Suppose for contradiction that the columns of B_c are not affinely independent. Then there would exist a column index j with $1 \leq j \leq n$ and some coefficients $\{f_i\}_{i \neq j}$ (not all zero) such that

$$\sum_{i \neq j} f_i(\mathbf{b}_i - \mathbf{b}_j) = 0 \Leftrightarrow \sum_{i=1}^{n} g_i \mathbf{b}_i = 0 \Leftrightarrow \mathbf{g} \in \ker(B_c) = \operatorname{span}(\mathbf{u}),$$

where we define $g_i = f_i$ for $i \neq j$ and $g_j = -\sum_{i \neq j} f_i$. But by its definition, $\mathbf{g}$ can only be constant if it is zero and thus if $\mathbf{f}$ is zero, which is a contradiction. Thus the columns of B_c are the affinely independent vertices of a simplex with centered Gram matrix M. This completes the proof. □

Proposition 4.15 leads to the following characterization of strict negative type metric spaces (see for instance [96, Thm. 1.2.4],[25]):

Proposition 4.16. *A metric space (X, d) has strict negative type if and only if there exists an isometric embedding of $(X, \sqrt{d})$ into $\mathbb{R}^{n-1}$ such that the embedded points are the vertices of a simplex.*

Proof. *(forward direction)* Let (X, d) be a strict negative type metric space. By Property 4.11, there exists an isometric embedding ϕ of $(X, \sqrt{d})$ into $\mathbb{R}^{n-1}$. The distance matrix D is the squared Euclidean distance matrix of the embedded points $\phi(X)$ and, following expression (4.2), we find that $\mathbf{f}^T D \mathbf{f} = -2\mathbf{f}^T M \mathbf{f}$ for all $\mathbf{f} \in \operatorname{span}(\mathbf{u})^\perp$, where M is the centered Gram matrix of the points $\phi(X)$. By strict negative type of (X, d), we then know that

$$\mathbf{f}^T D \mathbf{f} < 0 \text{ for all nonzero } \mathbf{f} \in \operatorname{span}(\mathbf{u})^\perp \Leftrightarrow \; M \text{ is PSD with } \ker(M) = \operatorname{span}(\mathbf{u}),$$

where 'PSD' abbreviates positive semidefinite. By Proposition 4.15, this means that M is the centered Gram matrix of the vertices of a simplex and thus that $\phi(X)$ are the vertices of a simplex as required.

(converse direction) Conversely, let (X, d) be a metric space which has an isometric embedding ϕ of $(X, \sqrt{d})$ into $\mathbb{R}^{n-1}$ such that the embedded points $\phi(X)$ are the vertices of a simplex. In particular, we then know that D is the squared Euclidean

distance matrix of a simplex. By Proposition 4.15, the centered Gram matrix M of the points $\phi(X)$ is PSD with kernel $\operatorname{span}(\mathbf{u})$ and thus

$$\mathbf{f}^T D\mathbf{f} = -2\mathbf{f}^T M\mathbf{f} < 0 \text{ for all nonzero } \mathbf{f} \in \operatorname{span}(\mathbf{u})^\perp,$$

which proves that (X, d) has strict negative type. This completes the proof. □

We again remark that implicitly all triangular faces need to be nonobtuse triangles due to the triangle inequality of d. Any strict negative type metric space can thus be characterized uniquely (up to rigid transformations) by a simplex with nonobtuse triangular faces. An example of a non-strict negative type metric is the squared distance between the vertices of a hypercube of dimension at least 2.

The distinction between strict and non-strict negative metric spaces originates from the strictness in condition (4.1), which translates to affine independence of the embedded points. It now remains to show which additional geometric constraints characterize resistance metric spaces. As shown in the next paragraph, this will be in terms of the angles of the simplex.

A *face* of a simplex is a simplex determined by a subset of the vertices[5]; for a specified set $\mathcal{V}$ of vertices, we will call this a $\mathcal{V}$-face. A face containing all but one vertex is also called a *facet* or an *elementary face*[6], and the interior angle between two facets is called the *dihedral angle* (see Figure 4.2). Fiedler studied the qualitative (smaller, equal or larger than $\pi/2$) patterns of dihedral simplex angles in [96]; here we consider the following class of simplices:

Definition 4.17. A *hyperacute simplex* is a simplex with nonobtuse dihedral angles.

Hyperacute simplices (also called "nonobtuse simplices") are important in triangulations, where they lead for instance to a 'discrete maximum principle' [32, Ch. 9],[33].

To calculate the angles in a simplex, we will consider the *centered pseudoinverse Gram matrix* $M^\dagger$ of a simplex, which is the Moore–Penrose pseudoinverse of the centered Gram matrix M. Repeating the arguments for the Laplacian and its pseudoinverse, we find that $\mathbf{f}^T M^\dagger \mathbf{f} \geq 0$, with equality if and only if $\mathbf{f} \in \operatorname{span}(\mathbf{u})$, and that $MM^\dagger = M^\dagger M = I - \mathbf{u}\mathbf{u}^T/n$. We find the following characterization of dihedral angles [96, Thm. 2.1.1]:

[5] If this subset were affinely dependent, then also the full set of vertices would be dependent; as a result, any face of a simplex is again a simplex.

[6] For example, the i^c-face for some $i \in \mathcal{N}$ is the face determined by all vertices but i.

Proposition 4.18 (dihedral angles)**.** *The dihedral angle θ_{ij} between the i^c-face and j^c-face of a simplex is given by the centered pseudoinverse Gram matrix $M^\dagger$ as*

$$\cos(\theta_{ij}) = \frac{-(M^\dagger)_{ij}}{\sqrt{(M^\dagger)_{ii}(M^\dagger)_{jj}}}.$$

Proof. Let S be a simplex and let $B = [\mathbf{b}_1 \ \dots \ \mathbf{b}_n]$ be the $(n-1) \times n$ matrix with the vertex vectors of S as its columns. Since angles are invariant under rigid transformations, we may assume that S is centered, with $\sum_{i=1}^n \mathbf{b}_i = 0$, and thus that B has zero row sums $B\mathbf{u} = 0$. As in the proof of Proposition 4.15, we know that affine independence implies that $\ker(B) = \operatorname{span}(\mathbf{u})$. Finally, we note that $M = B^T B$ is the centered Gram matrix of S.

The Moore–Penrose pseudoinverse $B^\dagger$ is an $n \times (n-1)$ matrix and we write $B^\dagger = [\tilde{\mathbf{b}}_1 \ \dots \ \tilde{\mathbf{b}}_n]^T$, which means that $B^\dagger$ has rows $\tilde{\mathbf{b}}_i^T$. Following Proposition A.5 in Appendix A.3, we know that $B^\dagger B = I - \mathbf{u}\mathbf{u}^T/n$, which can be used to show that $M^\dagger = B^\dagger B^{\dagger T}$ by checking that it agrees with the pseudoinverse definition. Finally, we have $(B^T B)^\dagger = B^\dagger B^{\dagger T}$ as a property of the Moore–Penrose pseudoinverse of the product between transposed matrices B^T and B; see [113] (in general, $(AB)^\dagger \neq B^\dagger A^\dagger$).

We now show the proposition. The i^c-face of the simplex S determines a hyperplane (affine space)

$$\mathcal{H}_i = \left\{B\mathbf{f} : \mathbf{f} \in \mathbb{R}^n \text{ s.t. } f_i = 0 \text{ and } \mathbf{u}^T\mathbf{f} = 0\right\}$$

of vectors parallel with the i^c-face. Indeed, any vector $\mathbf{f}$ with $f_i = 0$ and $\mathbf{f}^T\mathbf{u} = 0$ can be decomposed as $\mathbf{f} = \alpha(\mathbf{g} - \mathbf{g}')$, where $\alpha \in \mathbb{R}$ is some scalar and where $\mathbf{g}, \mathbf{g}' \in \mathbb{R}^n$ are two vectors with $g_i = g_i' = 0$ and $\mathbf{u}^T\mathbf{g} = \mathbf{u}^T\mathbf{g}' = 1$. By construction, we have that $B\mathbf{g}, B\mathbf{g}'$ are points in the i^c-face, which means that the vector $B\mathbf{f} = \alpha(B\mathbf{g} - B\mathbf{g}')$ is parallel to the facet and $\mathcal{H}_i$.

The rows of the pseudoinverse vertex matrix $B^\dagger$ satisfy

$$\tilde{\mathbf{b}}_i^T(B\mathbf{f}) = \mathbf{e}_i^T(B^\dagger B)\mathbf{f} = \mathbf{e}_i^T\left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right)\mathbf{f} = 0,$$

if $f_i = 0$ and $\mathbf{u}^T\mathbf{f} = 0$; in other words, $\tilde{\mathbf{b}}_i$ is a normal vector of the hyperplane $\mathcal{H}_i$. Since $\tilde{\mathbf{b}}_i^T\mathbf{b}_i = (I - \mathbf{u}\mathbf{u}^T/n)_{ii} = 1 - 1/n > 0$ (i.e. $\mathbf{b}_i$ and $\tilde{\mathbf{b}}_i$ point in the same direction) and the simplex S is centered, this shows that $\tilde{\mathbf{b}}_i$ is the inner-normal vector of the i^c-face of S. As illustrated in Figure 4.2, the dihedral (inner) angle θ_{ij} between a pair of facets can be calculated from the angle between these corresponding inner-normal vectors as follows:

$$\cos(\pi - \theta_{ij}) = -\cos(\theta_{ij}) = \frac{\tilde{\mathbf{b}}_i^T\tilde{\mathbf{b}}_j}{\|\tilde{\mathbf{b}}_i\|\|\tilde{\mathbf{b}}_j\|} = \frac{(M^\dagger)_{ij}}{\sqrt{(M^\dagger)_{ii}(M^\dagger)_{jj}}}.$$

This completes the proof. □

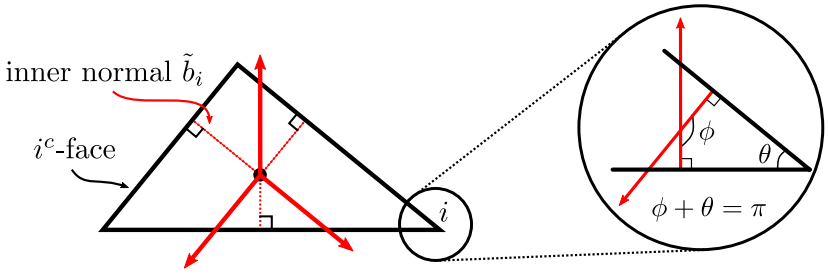


Figure 4.2: A 3-simplex with its inner normals (red) and a dihedral angle θ highlighted.

Thus if $(M^\dagger)_{ij}$ is larger/smaller/equal to zero then the dihedral angle θ_{ij} is obtuse/acute/right. This leads to the following characterization of resistance metric spaces [96, 179]:

Theorem 4.19. *A metric space (X, d) is a resistance metric space if and only if there exists an isometric embedding of $(X, \sqrt{d})$ into $\mathbb{R}^{n-1}$ such that the embedded points are the vertices of a hyperacute simplex.*

Proof. *(forward direction)* Let (X, d) be a resistance metric space. Since resistance metric spaces have strict negative type, there exists an isometric embedding ϕ of $(X, \sqrt{d})$ into $\mathbb{R}^{n-1}$ where $\phi(X)$ are the vertices of a simplex. The distance matrix D of these vertices is a resistance matrix and the corresponding centered Gram matrix $M = -\frac{1}{2}(I - \mathbf{u}\mathbf{u}^T/n)D(I - \mathbf{u}\mathbf{u}^T/n)$ is a pseudoinverse Laplacian matrix. This means that the pseudoinverse centered Gram matrix $M^\dagger$ is a Laplacian matrix with nonpositive off-diagonal entries, and thus that $\phi(X)$ are the vertices of a hyperacute simplex, as required.

(converse direction) Conversely, let (X, d) be a metric space with an isometric embedding ϕ of $(X, \sqrt{d})$ into $\mathbb{R}^{n-1}$ such that $\phi(X)$ are the vertices of a hyperacute simplex S. Following Proposition 4.15, we then know that the centered Gram matrix M of S is positive semidefinite with $\ker(M) = \operatorname{span}(\mathbf{u})$ and by extension the centered pseudoinverse Gram matrix $M^\dagger$ is also positive semidefinite with $\ker(M^\dagger) = \operatorname{span}(\mathbf{u})$. Furthermore, by hyperacuteness, the pseudoinverse centered Gram matrix has nonpositive off-diagonal entries $(M^\dagger)_{ij} \leq 0$, which means that $M^\dagger$ is a Laplacian matrix (Proposition 2.5) and thus that the distance matrix $D = \mathbf{u}\operatorname{diag}(M)^T + \operatorname{diag}(M)\mathbf{u}^T - 2M$ of the metric space (which is the squared Euclidean distance matrix of $\phi(X)$) is a resistance matrix. This completes the proof. □

We note once again that implicitly the embedded points must determine nonobtuse triangles. However, in contrast to the case of (strict) negative type metric spaces, this extra condition is automatically satisfied for hyperacute simplices due to the following result [96, Thm. 3.3.1]:

Theorem 4.20. *The face of a hyperacute simplex is a hyperacute simplex. In particular, all triangular faces are nonobtuse.*

Proof. Let S be a hyperacute simplex. Then its centered pseudoinverse Gram matrix has nonpositive off-diagonal entries and kernel $\operatorname{span}(\mathbf{u})$ and thus is a Laplacian matrix Q. The squared Euclidean distance matrix of S is then a resistance matrix Ω. This means that the $\mathcal{V}$-face of S has $(\Omega)_{\mathcal{V}\mathcal{V}}$ as its squared Euclidean distance matrix. Since this submatrix is a resistance matrix, it is again the squared Euclidean distance matrix of a hyperacute simplex by Theorem 4.19, which completes the proof. □

In addition to the constructive definition 4.6 and Proposition 4.7, which describes how to check whether a metric space is a resistance metric space, we thus have a second constructive definition of resistance metric spaces (see also [96, 179]):

Definition 4.21. A *resistance metric space* is a finite metric space (X, d), where X corresponds to the vertices of a hyperacute simplex and d to the squared distance between the vertices.

The equivalence with Definition 4.6 follows from Theorem 4.19 and the fact that triangular faces of a hyperacute simplex are hyperacute (Theorem 4.20), by which the squared edge lengths are a metric.

4.2.3 Closure and hierarchy of spaces

The classes of metric spaces discussed above all satisfy a specific property: they are *closed with respect to taking subspaces*. For negative type metric spaces, this follows directly from the definition in terms of distance matrices and geometrically from the fact that any subset of points with nonobtuse triples is again a set of point with nonobtuse triples. For strict negative type metric spaces, this follows from condition (4.1) on the distance matrix and the fact that a subset of affinely independent points (with nonobtuse triples) is again affinely independent. For resistance metric spaces, we know from Proposition 3.23 that the submatrix of a resistance matrix is again a resistance matrix and geometrically that faces of a hyperacute simplex are hyperacute (Theorem 4.20).

Taking the geometric 'closure of hyperacuteness' as a starting point leads to an alternative proof of the metric property of effective resistances. We note that this is not a circular argument, since the relation between effective resistances and hyperacute simplices is derived without reference to the fact that ω is a metric; it relies on properties of the Laplacian matrix (rank, kernel and off-diagonal signs) and how this is equivalent to the centered pseudoinverse Gram matrix of a hyperacute simplex.

Alternative proof of Theorem 4.2 The effective resistance is equal to the squared Euclidean distance between pairs of vertices of a hyperacute simplex S; thus it is symmetric and positive between distinct nodes. By closure of hyperacuteness (Theorem 4.20), every triangular face of S is a hyperacute triangle, which means that the squared edge lengths of every triangular face satisfy the triangle inequality (e.g. by the cosine rule), and thus that the effective resistances satisfy the triangle inequality. □

In [70], we remarked that this geometric perspective shows in a precise sense that '*the effective resistance is more than a distance*'. First of all, not all distances have an embedding — i.e. not all metrics are negative type. And, second, the effective resistance being a metric means that any three points in this embedding form a hyperacute triangle, which overlooks the much stronger constraint that any k points must form a hyperacute simplex and in particular that the n points form a hyperacute simplex; these are additional independence and angular constraints. This view was also considered in Klein [147] and by Subak-Sharpe & Moore [214]. The counterexample 4.10 illustrates an example of a simplex with triangular faces, which is not hyperacute. If the effective resistance were merely a metric, no such counterexample would exist.

The classification of resistance metric spaces can also be summarized in a hierarchy of metric spaces:

Proposition 4.22. *The following is a strict hierarchy of metric spaces: 'negative type ⊃ strict negative type ⊃ resistance metric space'. Geometrically, this corresponds to 'point sets with nonobtuse triangles ⊃ simplices with nonobtuse triangular faces ⊃ hyperacute simplices'. Each class is subset-closed.*

We remark that while these classes are closed under taking subsets, a subset can move 'downwards' in the hierarchy. Intuitively, this is because stricter classes in the hierarchy have a higher number of constraints, such that larger (smaller) spaces have more (less) 'degrees of freedom' to violate these constraints. For instance, in any

metric space, all three-point subspaces determine a nonobtuse triangle and are thus resistance metric spaces.

The utility of this hierarchy as a means to gradually generalize the effective resistance (or specialize to the effective resistance) is illustrated in Chapter 5, where we study the maximum-variance problem starting from general metrics over negative type metrics and finally to the resistance metric.

In Chapter 6, we will consider a further refinement of this hierarchy by defining positively curved graphs, in which the resistance curvature is positive $\mathbf{p} > 0$. Geometrically, these correspond to hyperacute simplices where the circumcenter lies inside the simplex (see Section 4.31), and the hierarchy is thus refined as

$$\begin{aligned} &\ldots \supset \text{resistance metric spaces} \supset \text{positively curved resistance metric spaces} \\ &\ldots \supset \text{hyperacute simplices} \supset \text{hyperacute simplices with interior circumcenters,} \end{aligned}$$

which are still subset-closed as shown in Proposition 6.29. Fiedler calls hyperacute simplices with interior circumcenters 'totally hyperacute simplices'.

4.3 The simplex geometry of graphs

4.3.1 graph–simplex correspondence

The previous section classified resistance metric spaces as strict negative type metric spaces and ended up with a geometric characterization in terms of hyperacute simplices. Since every resistance metric space is uniquely related to a weighted graph G, this geometric characterization extends to graphs [92]:

Theorem 4.23 (graph–simplex correspondence)**.** *There is a bijection between weighted graphs and hyperacute simplices, up to rigid transformations.*

Proof. Let G be a graph and define the mapping f : graphs $\to$ simplices, where $f(G)$ is the hyperacute simplex (seen as a representative of any of its rigid transformations) with squared Euclidean distance matrix equal to the resistance matrix Ω of G. We show that this mapping f is injective and surjective, and thus a bijection.

(injective) If two graphs G, G' map to the same simplex $f(G) = S = f(G')$, then this implies that $\Omega' = \Omega$ and thus that $G = G'$.

(surjective) Let S be a hyperacute simplex. Then there exists a graph G such that $f(G) = S$. Let D be the squared Euclidean distance matrix of S (which is invariant under rigid transformations); this is a resistance matrix Ω and the corresponding

graph G satisfies $f(G) = S$, as required. (We will not use the notation $f(G)$ any further) □

The bijection between graphs and hyperacute simplices is established by associating resistance matrices of graphs and squared Euclidean distance matrices of hyperacute simplices. Fiedler proved a broader version of the graph–simplex correspondence [92, 94], including additional equivalent objects, and Subak-Sharpe & Moore [179] found the same characterization but in terms of resistance matrices instead of graphs. The graph–simplex correspondence is constructive: for a given graph G with Laplacian matrix Q, let $\mathbf{z}_1, \dots, \mathbf{z}_{n-1}$ be an orthonormal basis of eigenvectors of the space $\ker(Q)^\perp = \operatorname{span}(\mathbf{u})^\perp$, corresponding to the positive eigenvalues $\mu_1, \dots, \mu_{n-1}$. Then the vectors $\mathbf{b}_1, \dots, \mathbf{b}_n$ defined as

$$(\mathbf{b}_i)_k := (\mathbf{z}_k)_i \mu_k^{-1/2} \text{ for all } i \in \mathcal{N} \text{ and } 1 \leq k \leq n-1$$

are the vertices of a hyperacute simplex S. The Gram matrix of these points is $Q^\dagger$ and their squared distances thus equal the effective resistance. Conversely, let S be a hyperacute simplex with vertices $\mathbf{b}_1, \dots, \mathbf{b}_n$. Then the centered Gram matrix with entries $\mathbf{b}_i^T \mathbf{b}_j$ is a pseudoinverse Laplacian and its Moore–Penrose pseudoinverse is the Laplacian matrix of the corresponding graph G. We will say that G, S is a corresponding *graph–simplex pair* if they are in bijection, or simply say that S is the simplex of (or associated to) G and vice versa.

As the proof of Theorem 4.20 already suggests, the graph–simplex correspondence also respects the Kron reduction and face relations between graphs and simplices (see [96, Thm. 3.3.1 & Thm. 3.2.3]):

Theorem 4.24. *There is a bijection between weighted graphs and hyperacute simplices. Furthermore if G, S is a graph–simplex pair, then the $\mathcal{V}$-face S' of S and the $\mathcal{V}$-Kron reduction G' of G is a graph–simplex pair G', S', for any non-empty subset $\mathcal{V}$ of nodes.*

Proof. The bijection was proven in Theorem 4.23. Let G, S be a graph–simplex pair, where Ω is the resistance matrix of G and Ω is the squared Euclidean distance matrix of S. Then the $\mathcal{V}$-face S' of S has squared Euclidean distance matrix $(\Omega)_{\mathcal{V}\mathcal{V}}$, equal to the resistance matrix of the $\mathcal{V}$-Kron reduction G', and thus G', S' is a graph–simplex pair. □

Schematically, this theorem can be summarized by the fact that the diagram

$$\begin{array}{ccc} G & \longleftrightarrow & S \\ {\scriptstyle \mathcal{V}\text{-Kron}}\downarrow & & \downarrow{\scriptstyle \mathcal{V}\text{-face}} \\ G' & \longleftrightarrow & S' \end{array}$$

commutes, i.e. going from G to S', we can first take the corresponding simplex S (via Ω) and then the $\mathcal{V}$-face S' or first take the $\mathcal{V}$-Kron reduction G' and then the corresponding simplex S' (via Ω'). Similarly, the two paths from S to G' commute. The right language to make this conceptual diagram (and the commuting relation) more precise is that of category theory; see for instance [98, 155, 170]. We may define a *category* **G** *of graphs* and a *category* **S** *of hyperacute simplices*. The objects in **G** are weighted graphs and the morphisms between graphs follow from the Kron reduction. The objects in **S** are (equivalence classes of) hyperacute simplices and a morphism between two simplices indicates that one is the face of the other[7]. In these categorical terms, Theorem 4.24 comes down to saying that *the graph-simplex bijection is a functor between* **G** *and* **S**. In other words, the graph–simplex correspondence is a so-called functorial relation that not only maps graphs to simplices but also maps '(Kron) relations between graphs' to '(face) relations between simplices'. Figure 4.3 below illustrates some of these relations for a given graph–simplex pair.

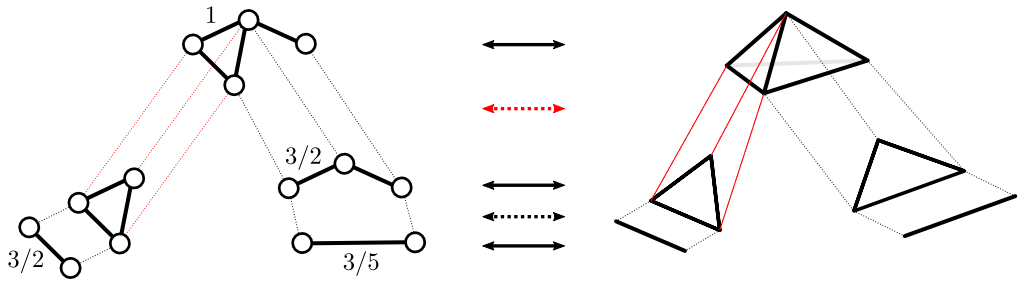


Figure 4.3: The graph–simplex correspondence relates weighted graphs (left) and hyperacute simplices (right), as well as Kron relations between graphs and face relations between simplices (red lines). Graph links without a number have weight $c = 1$.

Fiedler describes several applications of the graph–simplex correspondence in [96], and [239] studies low-dimensional graph embeddings based on the graph–simplex correspondence.

[7] A careful definition of these categories would be in terms of (node–) labeled graphs and (vertex–) labeled simplices.

4.3.2 The resistive embedding

Following the classification of resistance metric spaces, we know that there is an isometric Euclidean embedding of $(\mathcal{N}, \sqrt{\omega})$, which maps the nodes of a graph G onto the vertices of a hyperacute simplex S. This embedding can be extended to functions on the nodes as follows:

Definition 4.25 (resistive embedding)**.** Let G be a graph. A *resistive embedding* is a function $\varphi : \mathbb{R}^n \to \mathbb{R}^{n-1}$ defined by $\|\varphi(\mathbf{e}_i) - \varphi(\mathbf{e}_j)\|^2 = \omega_{ij}$ for all i, j and which is linearly extended to functions on the nodes as follows:

$$\varphi(\mathbf{f}) := \sum_{i \in \mathcal{N}} f_i \varphi(\mathbf{e}_i) \text{ for all } \mathbf{f} \in \mathbb{R}^n.$$

A resistive embedding is thus defined by mapping the basis vectors in $\mathbb{R}^n$ to the vertices of a hyperacute simplex in $\mathbb{R}^{n-1}$. These embeddings can be decomposed as

$$\varphi(\mathbf{f}) = \varphi'(\mathbf{f}) + (\mathbf{u}^T \mathbf{f})\overline{\varphi},$$

where φ' satisfies $\sum_{i \in \mathcal{N}} \varphi'(\mathbf{e}_i) = 0$ and with $\overline{\varphi} := \sum_{i \in \mathcal{N}} \varphi(\mathbf{e}_i)$. The embedding φ' is called a *centered embedding* and if $\overline{\varphi} = 0$, then $\varphi = \varphi'$ is a centered embedding. In terms of matrices, a resistive embedding can be written as

$$\varphi(\mathbf{f}) = B\mathbf{f} + (\mathbf{u}^T \mathbf{f})\overline{\varphi} \text{ with } B = [\varphi'(\mathbf{e}_i) \ \ldots \ \varphi'(\mathbf{e}_j)], \tag{4.3}$$

where B is the matrix with the centered embedding of the vertices as columns, such that $B\mathbf{u} = 0$. The matrix $B^T B = Q^\dagger$ is the centered Gram matrix of the hyperacute simplex S corresponding to G.

We now show a number of results that illustrate how resistive embeddings can be used to study *functional data* on a graph, i.e. node functions $f : \mathcal{N} \to \mathbb{R}$ with a corresponding vector representation $\mathbf{f} = (f(1), \ldots, f(n))^T$. A first result deals with the distance between two embedded points:

Proposition 4.26. *Let G be a graph with resistive embedding φ. Then*

$$\|\varphi(\mathbf{f}) - \varphi(\mathbf{g})\|^2 = (\mathbf{f} - \mathbf{g})^T Q^\dagger (\mathbf{f} - \mathbf{g}) \text{ for all } \mathbf{u}^T \mathbf{f} = \mathbf{u}^T \mathbf{g}.$$

Proof. Following decomposition (4.3) of the resistive embedding, we find

$$\|\varphi(\mathbf{f}) - \varphi(\mathbf{g})\|^2 = \|B(\mathbf{f} - \mathbf{g})\|^2 = (\mathbf{f} - \mathbf{g})^T Q^\dagger (\mathbf{f} - \mathbf{g}),$$

as required. This completes the proof. □

One consequence is the following property of resistive embeddings:

Property 4.27. *Let G be a graph with resistive embedding φ. Then φ is a bijection between constant-sum vectors in $\mathbb{R}^n$, and $\mathbb{R}^{n-1}$.*

Proof. Let $F := \{\mathbf{f} \in \mathbb{R}^n : \mathbf{u}^T\mathbf{f} = \alpha\}$ be vectors in $\mathbb{R}^n$ whose entries sum to α; we will also call these α-sum vectors. We prove that $\varphi : F \to \mathbb{R}^{n-1}$ is injective and surjective and thus bijective. *(injective)* If two vectors $\mathbf{f}, \mathbf{g} \in F$ are mapped onto the same point, then $\|\varphi(\mathbf{f}) - \varphi(\mathbf{g})\|^2 = 0$. By Proposition 4.26, we then know that $(\mathbf{f} - \mathbf{g})^T Q^\dagger (\mathbf{f} - \mathbf{g}) = 0$, which implies that $\mathbf{f} - \mathbf{g} \in \text{span}(\mathbf{u})$. But since $\mathbf{f}$ and $\mathbf{g}$ are α-sum vectors, they cannot differ by a constant and thus must be equal. *(surjective)* Let $\mathbf{x} \in \mathbb{R}^{n-1}$. Then we show that there is a function $\mathbf{f} \in F$ such that $\varphi(\mathbf{f}) = \mathbf{x}$. Decomposing the resistive embedding, we find $\mathbf{x} = \varphi(\mathbf{f}) = B\mathbf{f} + \alpha\overline{\varphi}$. Since the columns of B are the vertices of a hyperacute simplex, they are affinely independent with $\text{rank}(B) = n - 1$ and $\ker(B) = \text{span}(\mathbf{u})$. Therefore, the system $B\mathbf{f} = \mathbf{x} - \alpha\overline{\varphi}$ is underdetermined and has a family of solutions $\{\mathbf{f} + \lambda\mathbf{u}\}_{\lambda\in\mathbb{R}}$ for some $\mathbf{f}$, which intersects with F at $\lambda = (\alpha - \mathbf{u}^T\mathbf{f})/n$. This completes the proof. □

We now focus our attention on *unit-sum vectors*, which are vectors with $\mathbf{u}^T\mathbf{f} = 1$. By Property 4.27, these unit-sum vectors are in bijection with $\mathbb{R}^{n-1}$ under the resistive embedding φ. In other words, every point $\mathbf{x} \in \mathbb{R}^{n-1}$ has a unique associated unit-sum vector with resistive embedding $\mathbf{x}$; we call this the (unit-sum) *coordinate* of $\mathbf{x}$ and write $\varphi^{-1}(\mathbf{x})$. This is also called the 'barycentric coordinate'. Combining Proposition 4.26 and Property 4.27, we find that φ induces a metric between unit-sum vectors:

$$d(\mathbf{f}, \mathbf{g}) = \|\varphi(\mathbf{f}) - \varphi(\mathbf{g})\| = \sqrt{(\mathbf{f} - \mathbf{g})^T Q^\dagger (\mathbf{f} - \mathbf{g})} \text{ for all unit-sum } \mathbf{f}, \mathbf{g} \in \mathbb{R}^n.$$

This illustrates how the resistive embedding provides additional structure on functional data on graph.

We consider one further restriction; the nonnegative unit-sum vectors in $\mathbb{R}^n$ are given by

$$\Delta := \left\{\mathbf{f} \in \mathbb{R}^n : \mathbf{f} \geq 0 \text{ and } \mathbf{f}^T\mathbf{u} = 1\right\}.$$

The set Δ is called the *probability simplex* since it is an n-simplex in $\mathbb{R}^n$ with vertices $\mathbf{e}_1, \ldots, \mathbf{e}_n$; its elements $\mathbf{f} \in \Delta$ are called *distributions*. This terminology follows because nonnegative unit-sum vectors correspond to distributions on the nodes; see Chapter 5. The probability simplex and the simplex of a graph are related as follows:

Theorem 4.28. *A resistive embedding φ is a bijection between the probability simplex Δ and the hyperacute simplex S associated with φ.*

Proof. We prove that $\varphi : \Delta \to S$ is injective and surjective, and thus bijective. *(injective)* Since φ is injective between unit-sum vectors and $\mathbb{R}^{n-1}$, it is injective in particular between Δ and S. *(surjective)* Any point in the simplex $\mathbf{x} \in S$ can be expressed as a convex combination of the n vertices of S as $\sum_{i \in \mathcal{N}} f_i \varphi(\mathbf{e}_i)$. The coefficients $f_1, \ldots, f_n$ of this convex combination are nonnegative and sum to one, which implies that they determine a vector $\mathbf{f} \in \Delta$ in the probability simplex for which $\varphi(\mathbf{f}) = \mathbf{x}$. This completes the proof. □

The coordinate of a point in the simplex is thus a distribution on the graph and likewise, every distribution corresponds to a unique point on the simplex. We may also write

$$\varphi(\Delta) = S \text{ and } \varphi^{-1}(S) = \Delta$$

This bijection is a starting point and an important motivation for the results in Chapter 5, which looks in more detail at how to characterize and study distributions on a graph. To illustrate Theorem 4.23 and Proposition 4.26, we give some examples of points on the simplex S and their coordinates.

Example 4.29 (**vertices**)**.** Following Definition 4.25 of resistive embeddings, the basis vectors $\mathbf{e}_i$ are mapped to the vertices $\varphi(\mathbf{e}_i)$ of the simplex S.

Example 4.30 (**centroid**)**.** The point $\varphi(\mathbf{u}/n)$ in the simplex that corresponds to the uniform distribution $\mathbf{u}/n$ is called the *centroid* of S; this is also called the 'barycenter' or 'center of mass'. Since $Q^\dagger \mathbf{u} = 0$, we know that $\mathbf{f}^T Q^\dagger \mathbf{f} = \|\varphi(\mathbf{f}) - \varphi(\mathbf{u}/n)\|^2$ for any distribution $\mathbf{f} \in \Delta$. In other words, the quadratic form $\mathbf{f}^T Q^\dagger \mathbf{f}$ is equal to the squared distance from $\varphi(\mathbf{f})$ to the centroid of S. In particular, we find

$$(Q^\dagger)_{ii} = \|\varphi(\mathbf{e}_i) - \varphi(\mathbf{u}/n)\|^2 \text{ for each node } i,$$

which means that the pseudoinverse diagonal entries are equal to the squared distance from the corresponding vertices to the simplex centroid. This geometric interpretation of $(Q^\dagger)_{ii}$ was used in [203] to argue for the pseudoinverse diagonal as a centrality measure.

Example 4.31 (**circumcenter**)**.** Every simplex S has a unique sphere, the *circumsphere*, that passes through all of its vertices. The center and radius of this sphere are called the *circumcenter* and *circumradius*, respectively. The following result underlies the name 'resistance radius' for σ^2 (see [96, Thm. 2.1.1]):

Proposition 4.32. *Let G be a graph with associated simplex S. Then $\mathbf{p}$ is the unit-sum coordinate of the circumcenter of S and σ the circumradius of S.*

Proof. Let $\mathbf{f}$ be the unit-sum coordinate of the circumcenter. Then the circumsphere equations $r^2 = \|\varphi(\mathbf{f}) - \varphi(\mathbf{e}_i)\|^2$ for all i can be written as follows:

$$r^2 = (\mathbf{f} - \mathbf{e}_i)^T Q^\dagger (\mathbf{f} - \mathbf{e}_i) = -\frac{1}{2}(\mathbf{f} - \mathbf{e}_i)^T \Omega (\mathbf{f} - \mathbf{e}_i) = -\frac{1}{2}\mathbf{f}^T \Omega \mathbf{f} + \mathbf{e}_i^T \Omega \mathbf{f} \text{ for all } i.$$

This shows that $\Omega\mathbf{f}$ must be a constant vector $\Omega\mathbf{f} = \left(r^2 + \frac{1}{2}\mathbf{f}^T\Omega\mathbf{f}\right)\mathbf{u}$. Inverting the resistance matrix and using the unit-sum property of $\mathbf{f}$, this yields the solution $\mathbf{f} = \mathbf{p}$ and $r = \sigma$, as required. □

If the resistance curvature is nonnegative $\mathbf{p} \geq 0$, this implies that $\mathbf{p} \in \Delta$ and thus that $\varphi(\mathbf{p}) \in \Delta$; i.e. that the circumcenter is a point in the simplex (or one of its faces). This is the geometric meaning of positively curved graphs studied in Chapter 6.

The formula for the circumradius of a simplex in terms of its squared Euclidean distance matrix seems to have been rediscovered a number of times; for instance in [61, 90, 110, 131, 261]. And while some recursive relations between the circumcenter of different faces of a simplex are known; see for instance [126, §2.4], the concise expressions (3.15) that follow from the Schur complement of the Fiedler–Bapat identity do not seem to be widely known.

We remark that the circumradii and circumcenters of different faces in a simplex satisfy interesting orthogonality relations. Let S' be a face of S; since both $\varphi(\mathbf{p})$ and $\varphi(\mathbf{p}')$ are equidistant to all vertices of S', the circumcenter $\varphi(\mathbf{p}')$ of S' is the orthogonal projection of $\varphi(\mathbf{p})$ onto the affine space determined by S'. By Pythagoras' Theorem, the distance between $\varphi(\mathbf{p})$ and $\varphi(\mathbf{p}')$ is then given by $\|\varphi(\mathbf{p}) - \varphi(\mathbf{p}')\|^2 = \sigma^2 - \sigma'^2$ which in the case of elementary faces yields

$$\|\varphi(\mathbf{p}) - \varphi(\mathbf{p}')\|^2 = \sigma^2 - \sigma'^2 = \frac{p_i^2}{k_i} \text{ for the } i^c\text{-face } S' \text{ of } S, \tag{4.4}$$

where we recall that the i^c-face of S is the face determined by all vertices except i. Expression (4.4) yields a new interpretation of p_i in terms of the distance from the circumcenter of a simplex to the corresponding elementary face (facet); see also [96, Eq. 2.7]. The fact that $\varphi(\mathbf{p}) - \varphi(\mathbf{p}')$ is a normal to the face S' can be used for an alternative proof of Proposition 4.18 on the dihedral angles in a hyperacute simplex:

Alternative proof of Proposition 4.18 (restricted to hyperacute simplices) Let S be a hyperacute simplex and consider the i^c-face $S' = S/i$ and the j^c-face $S'' = S/j$. We consider the associated graph G and a resistive embedding φ. From the discussion above, we know that the vector $\varphi(\mathbf{p}) - \varphi(\mathbf{p}')$ is an inner-normal vector

of S' and, similarly, that $\varphi(\mathbf{p}) - \varphi(\mathbf{p}'')$ is an inner-normal vector of S''. The dihedral angle θ_{ij} between S' and S'' can be determined based on the angle between these two normal vectors as follows (see Figure 4.2):

$$\begin{aligned}\cos(\pi - \theta_{ij}) &= \frac{(\varphi(\mathbf{p}) - \varphi(\mathbf{p}'))^T(\varphi(\mathbf{p}) - \varphi(\mathbf{p}''))}{\|\varphi(\mathbf{p}) - \varphi(\mathbf{p}')\|.\|\varphi(\mathbf{p}) - \varphi(\mathbf{p}'')\|} \\ &= \frac{(\mathbf{p} - \mathbf{p}')^T B^T B(\mathbf{p} - \mathbf{p}'')}{\sqrt{(\mathbf{p} - \mathbf{p}')^T Q^\dagger(\mathbf{p} - \mathbf{p}')(\mathbf{p} - \mathbf{p}'')^T Q^\dagger(\mathbf{p} - \mathbf{p}'')}} \quad \text{(by Eq. (4.3))} \\ &= \frac{(\mathbf{p} - \mathbf{p}')^T Q^\dagger(\mathbf{p} - \mathbf{p}'')}{\sqrt{(\mathbf{p} - \mathbf{p}')^T Q^\dagger(\mathbf{p} - \mathbf{p}')(\mathbf{p} - \mathbf{p}'')^T Q^\dagger(\mathbf{p} - \mathbf{p}'')}}.\end{aligned}$$

By properties of the Kron reduction (3.16), we know that

$$\mathbf{p} - \mathbf{p}' = p_i \left(\mathbf{e}_i - \sum_{k:k\sim i} \frac{c_{ik}}{k_i} \mathbf{e}_k \right) = \frac{p_i}{k_i} Q\mathbf{e}_i$$

and similarly for $\mathbf{p} - \mathbf{p}''$, and thus

$$\begin{aligned}\cos(\pi - \theta_{ij}) &= \frac{\frac{p_i}{k_i}\mathbf{e}_i^T QQ^\dagger Q\mathbf{e}_j \frac{p_j}{k_j}}{\sqrt{\frac{p_i^2 p_j^2}{k_i^2 k_j^2}\mathbf{e}_i^T QQ^\dagger Q\mathbf{e}_i . \mathbf{e}_j^T QQ^\dagger Q\mathbf{e}_j}} \\ &= \frac{(Q)_{ij}}{\sqrt{(Q)_{ii}(Q)_{jj}}} \quad (\text{since } QQ^\dagger Q = Q).\end{aligned}$$

Since $\cos(\pi - \theta_{ij}) = -\cos(\theta_{ij})$ and the pseudoinverse centered Gram matrix $M^\dagger$ of S is equal to the Laplacian Q of the associated graph G, we thus retrieve $\cos(\theta_{ij}) = -(M^\dagger)_{ij}/\sqrt{(M^\dagger)_{ii}(M^\dagger)_{jj}}$, as required. □

The same proof works for simplices that are not hyperacute.

Chapter 5

Variance and covariance of distributions on graphs

In this chapter, we generalize the concepts of variance and covariance to distributions — i.e. nonnegative unit-sum node functions — on a graph. Roughly speaking, the graph variance measures the average distance between two randomly sampled nodes to characterize the spread of a distribution. This addresses a methodological gap in the analysis of data defined on the nodes of graphs and networks[1] and, being both simple in form and concept, we expect the graph variance to be widely applicable. We illustrate its use in a number of examples and experiments on a real-world network. The second part of this chapter deals with the 'maximum-variance problem'. We find several interesting results on this problem, such as an equivalent formulation, a complexity hierarchy that partly matches the metric hierarchy of the resistance distance, and a relation to graph boundaries.

Section 5.1 introduces the new variance and covariance measures with several examples for illustration.

In Section 5.2 we use variance and covariance to study a network of mathematical concepts (based on Wikipedia) and scientific articles that invoke these concepts. We use variance to characterize the diversity of articles and use covariance to quantify the degree to which related concepts are likely to appear together in an article (Figures 5.3–5.6 and Table 5.1). We propose some basic null models and statistical tests to compare the empirical values against.

Section 5.3 introduces the maximum-variance problem (5.6), which asks to find the largest possible variance for a given graph G and distance d. We formulate

[1] We will also call this *functional data*, as opposed to *structural data* which is data about the graph itself.

necessary conditions for the solution and as a first result find that the maximum-variance problem can be rephrased in terms of submatrices of the distance matrix (Theorem 5.15), in close analogy with the work of Leinster & Meckes on maximum diversity [158]. As the second result, we consider the maximum-variance problem for different classes of metrics; we find that by gradually adding constraints to the metric, the problem transitions from 'hard' for general distances to 'computable' for negative type metrics and finally to 'greedily computable' for resistance metrics (Example 5.16, Proposition 5.20 and Theorem 5.23), in close relation with the metric hierarchy in Proposition 4.22.

Section 5.4 describes a third feature of the maximum variance and its solutions: maximum variance distributions tend to concentrate on the 'boundary' of a graph. Finally, we briefly discuss the graph variance in the context of the associated simplex geometry.

5.1 Variance and covariance on graphs

Variance is a fundamental concept in probability theory and statistics and is routinely applied throughout the sciences and engineering. Intuitively speaking, the variance reflects how spread-out samples of a distribution are. In many practical cases however, distributions are defined on the nodes of a graph and the usual definition of variance does not apply. Here we address this methodological gap and propose a measure of variance and covariance for distributions defined on the nodes of a graph. The structure of the graph is taken into account by using distances between the nodes.

Our contribution relates to the pursuit of other research topics, such as graph signal processing and graph neural networks: to take tools that are traditionally designed for data on Euclidean or simple grid-like spaces, and generalize them to work for data defined on graphs and other irregular structures [35, 36, 215, 257, 268].

Our proposed variance is also closely related to the setting of 'diversity measures' [68, 157, 202, 204, 227] which are widely used, for instance in ecology [42, 209]. In particular, our variance is a special case of Rao's quadratic entropy [204], which measures diversity of distributions on sets with respect to 'differences' defined between elements of the set, and fits in the framework of Leinster & Cobbold's family of diversity measures [157] (see Section 5.3.1 for the latter).

5.1.1 Variance

We recall the notion of distributions introduced in Chapter 4. A *distribution on a graph* is a function on the nodes $f : \mathcal{N} \to [0, 1]$, which assigns a nonnegative number $f(i)$ to each node in the graph, such that all numbers add up to one: $\sum_{i \in \mathcal{N}} f(i) = 1$. This corresponds to a vector $\mathbf{f} \in \mathbb{R}^n$ with nonnegative entries $\mathbf{f} \geq 0$ and unit sum $\mathbf{u}^T\mathbf{f} = 1$; in other words, a distribution $\mathbf{f}$ is an thus an element of the *probability simplex* Δ. Distributions can be used to define a *random node* N, which is a random element of the node set, where the probability to sample any node i is given by the corresponding function value $f(i)$. In other words, we can 'sample' the random node N and get any of the nodes of G with probability

$$\Pr[N = i] = f(i) \text{ for all } i \in \mathcal{N}.$$

For this reason, the value $f(i)$ of a node is called the *probability* of i. To specify the underlying distribution of a random node N, we will write[2] $N \overset{d}{\sim} f$ and say that N is distributed according to f. For some function $\gamma : \mathcal{N} \to \mathbb{R}$, the expected value (average) of γ of a random node is defined as

$$\mathbb{E}(\gamma(N)) := \sum_{i \in \mathcal{N}} f(i)\gamma(i) \text{ with } N \overset{d}{\sim} f.$$

For instance, the average degree of a random node $N \overset{d}{\sim} f$ is given by $\mathbb{E}(k_N) = \mathbf{k}^T\mathbf{f}$.

As introduced, the interpretation of a graph distribution is twofold: it can represent a random node, or it can be a signal (function, vector) on the nodes of a graph. In both cases, it is natural to ask whether the probabilities of a distribution are concentrated on a small part of the graph — and thus that we might restrict our attention to this small part — or if instead the distribution is spread out over the graph. In the context of probability theory, the variance is a standard way to answer this question; it reflects how spread out a distribution is, as the average squared difference between a random outcome of the distribution and the mean. Of course, there exist many other characterizations of distributions which convey complementary information to the variance and their generalization to distributions on graphs could be an interesting line of future research.

To generalize the standard notion of variance to graphs, we propose to take into account distances between nodes in a graph, i.e. some function $d : \mathcal{N} \times \mathcal{N} \to [0, \infty)$,

[2] A more standard notation in probability theory would be $N \sim f$; we choose the notation '$\overset{d}{\sim}$' in this thesis to avoid confusion with adjacency of nodes, which is denoted by $i \sim j$ if $(i, j) \in \mathcal{L}$.

from which a notion of 'being spread out' on a graph can be derived. Formally we require d to be a *semimetric*, which satisfies all properties of a metric except for the triangle inequality[3] — we will use the terms "distance" and "semimetric" interchangeably in this section. The corresponding distance matrix D is symmetric, has a zero diagonal and positive off-diagonal entries. We propose the following variance:

Definition 5.1 (graph variance)**.** Let G be a graph with a distance d between its nodes. Then the *graph variance* of a distribution $f \in \Delta$ (with respect to d) is defined as

$$\mathrm{var}(f) := \frac{1}{2} \sum_{i,j \in \mathcal{N}} f(i) f(j) d^2(i,j). \tag{5.1}$$

The proposed variance is thus equal to the average squared distance between two random nodes sampled according to f. To make the dependence of variance on the distance explicit, we will sometimes include a subscript var_d.

The specific choice for definition (5.1) follows as a generalization of the classical notion of variance. For a random variable $N \overset{d}{\sim} f$ taking values in a finite set of real numbers $\mathcal{N} \subset \mathbb{R}$, the variance is defined as

$$\mathbb{E}\left([N - \mathbb{E}(N)]^2\right) = \frac{1}{2} \sum_{i,j \in \mathcal{N}} f(i) f(j) (i-j)^2.$$

The (squared) difference $(i-j)^2$ is not defined for general sets $\mathcal{N}$ such as the nodes of a graph, but if we interpret it as the squared distance between elements in $\mathbb{R}$, this formula retrieves the graph variance (5.1).

The graph variance can be expressed as a quadratic form determined by the matrix $D^{(2)}$ containing squared pairwise distances between the nodes:

$$\mathrm{var}(\mathbf{f}) = \frac{1}{2} \mathbf{f}^T D^{(2)} \mathbf{f} \text{ for some } \mathbf{f} \in \Delta.$$

This highlights that the graph variance is a very simple measure given by a quadratic form, which is readily calculated by standard linear algebra software.

As noted, the distance d need not be a metric, but symmetry and positivity for distinct points are sensible requirements for the formula to make intuitive sense. For a discussion on graph metrics, we refer to Section 2.3.3, and we recall that in many applied contexts, a distance or metric between the nodes of a graph is induced by some embedding $\phi : \mathcal{N} \to M$ of the graph into another metric space, as follows: $d = d_M \circ \phi$. The graph variance (and covariance, see Section 5.1.2) can thus be

[3] We note that many of the subsequent results still hold if we also relax the requirement that $d(x,y) = 0$ only if $x = y$.

applied off the shelf in these cases. Moreover, some standard choices for M such as Euclidean, hyperbolic and spherical spaces lead to negative type induced metrics for which var_d has additional nice properties, such as concavity (see Section 5.3.2).

More generally, Definition 5.1 for the variance (and later Definition 5.7 for the covariance) can be used on any metric space. While we focus on the case of graphs and their distances, the same questions and applications are clearly also interesting in the more general setting of metric spaces.

We conclude the introduction of the graph variance with a number of examples for illustration.

Example 5.2. Figure 5.1 illustrates how the graph variance captures the spread of distributions on a graph.

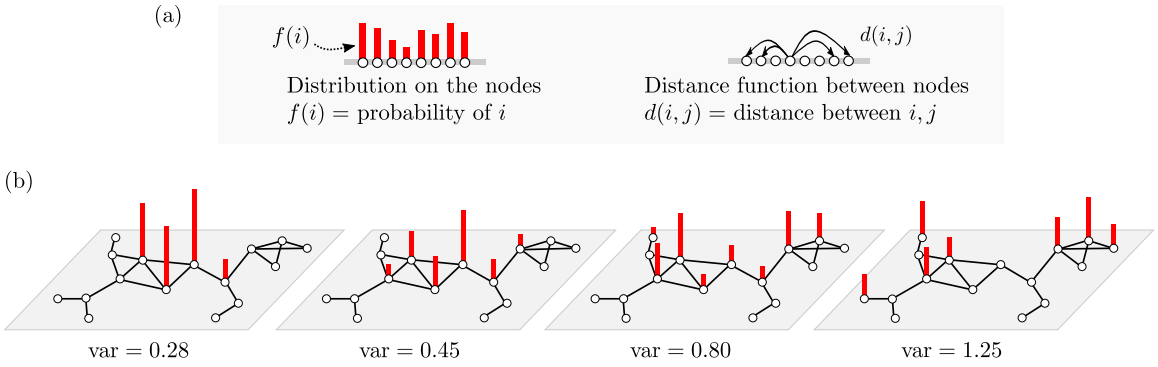


Figure 5.1: (a) Illustration of the data that goes into defining the graph variance and (b) a graph with four distributions with increasing variance, measured with respect to the square-root effective resistance as $\frac{1}{2}\mathbf{f}^T\Omega\mathbf{f}$.

Example 5.3 (square-root distance). If d is a (semi)metric, then $\sqrt{d}$ is also a (semi)metric[4]. This means that for any distance d on the graph, we can also define a variance as

$$\mathrm{var}(\mathbf{f}) = \frac{1}{2}\mathbf{f}^T D\mathbf{f} \text{ with distance matrix } D. \tag{5.2}$$

Since metrics d are generally better studied than squared metrics d^2, formula (5.2) is often more practical than (5.1), for instance to relate the graph variance to known

[4]Clearly, $\sqrt{d}$ is symmetric, equal to zero for identical elements and positive otherwise. Suppose for contradiction that the triangle inequality does not hold for a triple of elements i,j,k. Then we obtain the contradiction

$$d(i,k) > \left(\sqrt{d(i,j)} + \sqrt{d(j,k)}\right)^2 = d(i,j) + d(j,k) + 2\sqrt{d(i,j)d(j,k)} \geq d(i,k).$$

expressions. The square-root distance will be used in the next example and in Section 5.3.

Example 5.4 (**Wiener and Kirchhoff index**)**.** The *uniform distribution* $\mathbf{f} = \mathbf{u}/n$ is a canonical distribution. For the square root shortest-path distance $\sqrt{d_{sp}}$ as a metric (see Example 5.3), the variance is equal to

$$\mathrm{var}_{d_{sp}}(\mathbf{u}/n) = \frac{1}{2n^2}\sum_{i,j} d_{sp}(i,j).$$

The number $\frac{1}{2}\sum_{i,j} d_{sp}(i,j)$ is a widely-studied graph invariant known as the *Wiener index*, after Harry Wiener, who introduced it to characterize chemical compound graphs in mathematical chemistry [262].

Similarly, the variance of the uniform distribution with respect to the square root effective resistance $\sqrt{\omega}$ is equal to

$$\mathrm{var}_{\omega}(\mathbf{u}/n) = \frac{1}{2n^2}\sum_{i,j} \omega_{ij} = \frac{1}{n^2} Kf,$$

with Kirchhoff index Kf.

Example 5.5 (**random walks and Kemeny's constant**)**.** Random walks are well-studied (families of) distributions on a graph. A *random walk* is a process in which a 'random walker' makes its way across the nodes of a graph, following the links in the graph with certain probabilities [81, 159, 165]. More precisely, a random walk describes a sequence of random nodes $\{N_t\}_{t\in\mathbb{N}}$ where the consecutive random nodes N_t represent the position of the random walker at timestep t. The distributions of the random nodes are related via the transition probabilities (different probabilities give rise to different random walks)

$$\Pr[N_{t+1} = j | N_t = i] = c_{ij}/k_i.$$

In other words, the probability that the walker leaves a node via any given link is proportional to the weight of that link. On a connected graph, the distribution of N_t converges to a unique stationary distribution π with probabilities $\pi(i) = k_i/(n\langle k\rangle)$, where we recall that $\langle k\rangle$ is the average degree. The variance of the stationary distribution with respect to the square root effective resistance is then equal to

$$\mathrm{var}_{\omega}(\pi) = \frac{1}{2n^2\langle k\rangle^2}\sum_{i,j} k_i k_j \omega_{ij}.$$

The number $n\langle k\rangle \operatorname{var}_\omega(\pi)$ is equal to the *Kemeny constant* of a graph, as shown in [254]. This constant is the average time it takes for a random walker in the stationary distribution to reach a node i — as discovered by Kemeny [140], this average time is independent of i. Studying the variance of random walk distributions more generally (i.e. for all times t) is an interesting line of future research.

Example 5.6 (**diffusion**)**.** An important class of graph distributions follows from the discrete diffusion (heat) equation

$$\frac{d}{dt}\mathbf{f} = -Q\mathbf{f} \text{ with solution } \mathbf{f}(t) = \exp(-Qt)\mathbf{f}_0$$

for some initial vector $\mathbf{f}(0) = \mathbf{f}_0 \in \mathbb{R}^n$. Solutions to the diffusion equation can also be written as $\mathbf{f}(t) = P_t\mathbf{f}_0$ using the diffusion matrix[5] $P_t := \exp(-Qt)$. If the initial state is a distribution, then the solution will remain a distribution for all times[6]; in other words $P_t : \Delta \to \Delta$ for all $t \geq 0$, where we recall that Δ is the probability simplex. For any initial distribution, the solution converges to the uniform distribution $\mathbf{f}(t) \to \mathbf{u}/n$ for $t \to \infty$.

The diffusion distribution can also be interpreted as the occupation probability of a continuous-time random walker with initial distribution $\mathbf{f}_0$. In this context, the normalized diffusion process $\frac{d}{dt}\mathbf{f} = -Q\operatorname{diag}(\mathbf{k})^{-1}\mathbf{f}$ is more common since it is closely related to the classical discrete random-walk process. In Chapter 6, we discuss how the small-t variance of the diffusion distribution starting at a single node ($\mathbf{f}_0 = \mathbf{e}_i$) reflects the resistance curvature at that node (Eq. (6.9)).

5.1.2 Covariance

The study of random nodes is extended to pairs of random nodes by considering joint distributions. A *joint distribution on a graph* is a function $P : \mathcal{N} \times \mathcal{N} \to [0, 1]$, which assigns a nonnegative number $P(i, j)$ to each pair of nodes in the graph, such that all numbers add up to one. A joint distribution can be used to define a random pair of nodes (N, M), which are two random elements of the node set, with probability of being sampled given by the joint distribution as

$$\Pr[(N, M) = (i, j)] = P(i, j) \text{ for all } (i, j) \in \mathcal{N} \times \mathcal{N}.$$

[5]The family of matrices $\{P_t\}_{t>0}$ is also called the *heat semigroup* since it forms a semigroup with composition '$\circ$' as group operation: $P_s \circ P_t = P_{s+t}$.

[6]To see this, note (i) that $\frac{d}{dt}f_m = \sum_{j\sim m} c_{mj}(f_j - f_m) \geq 0$ if $f_m = f_{\min}$ and thus that nonnegativity is conserved by diffusion and (ii) that $\mathbf{u}^T\frac{d}{dt}\mathbf{f} = 0$, which means that the unit-sum property is conserved.

Any joint distribution P on node pairs also leads to two distributions $\tilde{f}$ and $\tilde{g}$ on the nodes, defined by $\tilde{f}(i) = \sum_{j \in \mathcal{N}} P(i,j)$ and $\tilde{g}(j) = \sum_{i \in \mathcal{N}} P(i,j)$, which are called the *marginal distributions* of P. We remark that P need not be symmetric in general, and thus that the ordering of the pair of nodes matters — i.e. (N, M) is in general different from (M, N).

For joint distributions, covariance is used much in the same way as the variance to quantify whether random pairs of nodes are on average close together, or far apart. Making use of a distance d (a semimetric) between the nodes of the graph again, we propose a generalization of covariance for joint distributions on a graph:

Definition 5.7 (graph covariance)**.** Let G be a graph with a distance d between its nodes. Then the *graph covariance* of a joint distribution P (with respect to d) is defined as

$$\operatorname{cov}(P) = \frac{1}{2} \sum_{i,j \in \mathcal{N}} [\tilde{p}(i)\tilde{q}(j) - P(i,j)]\, d^2(i,j). \tag{5.3}$$

The covariance thus measures the average distance between a random pair of nodes sampled according to the marginals $\tilde{f}, \tilde{g}$ minus the average distance between a pair sampled from P. We note that this expression is closely related to the 'distance covariance' [168, 233]. Our definition of covariance follows as a generalization of the classical notion of covariance: for a pair of random variables $(N, M) \overset{d}{\sim} P$ taking values in a set $\mathcal{N} \subset \mathbb{R}$ of numbers, the covariance is defined as (see Section 5.1.2.1)

$$\mathbb{E}\left([N - \mathbb{E}(N)][M - \mathbb{E}(M)]\right) = \frac{1}{2} \sum_{i,j \in \mathcal{N}} \Big(\tilde{f}(i)\tilde{g}(j) - P(i,j)\Big)\,(i-j)^2.$$

This corresponds to the graph covariance formula (5.3) since $(i-j)^2 = d^2(i,j)$ is equal to the Euclidean distance. The covariance can be expressed using the squared distance matrix $D^{(2)}$ and the $n \times n$ matrix P containing all pairwise probabilities as

$$\operatorname{cov}(P) = \frac{1}{2} \operatorname{tr}\left((P\mathbf{u}\mathbf{u}^T P - P)D^{(2)}\right) = \frac{1}{2}\tilde{\mathbf{f}}^T D^{(2)} \tilde{\mathbf{g}} - \frac{1}{2} \operatorname{tr}(PD^{(2)}).$$

In order to reduce the effect of the variance of the marginal distributions on the covariance, we propose a normalization, which we call the *graph correlation*:

$$\operatorname{corr}(P) := \frac{\operatorname{cov}(P)}{\sqrt{\operatorname{var}(\tilde{f})\operatorname{var}(\tilde{g})}}.$$

An alternative notion of correlation on graphs was defined in [60] by generalizing the Pearson correlation while taking into account the structure of a graph. We now show a number of examples for illustration.

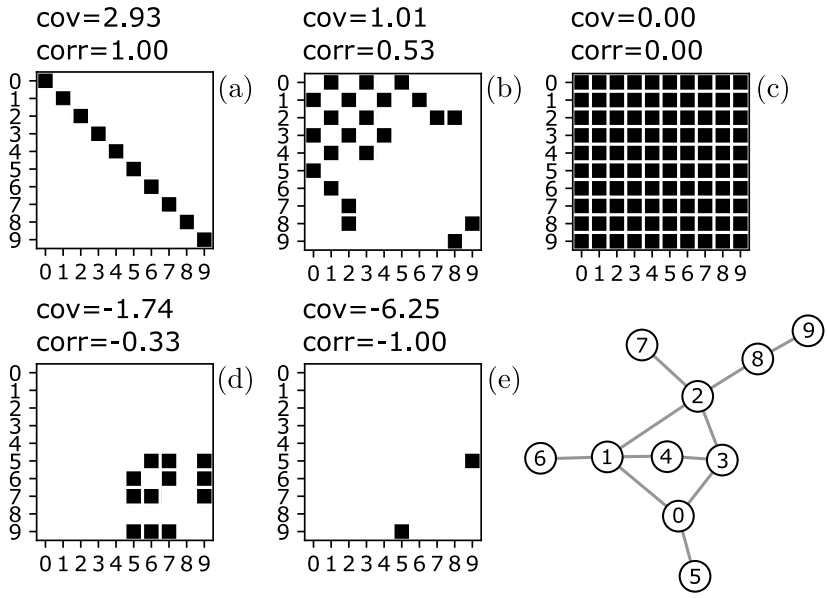


Figure 5.2: Five joint distributions on a 10-node graph. Each joint distribution is illustrated by a 10×10 grid with white squares indicating $P(i,j) = 0$ and black squares a constant such that P normalizes. The covariance and correlation are calculated with respect to the shortest-path distance for (a) the joint distribution for two identical nodes (N, N) with uniform distribution, (b) the joint distribution for the ends of a random link, i.e. (N, M) is a link sampled uniformly at random from the graph, (c) the uniform joint distribution, corresponding to two independent uniform random nodes, (d) the joint distribution for a pair of leaf nodes, i.e. with (N, M) a random distinct pair of nodes from $\{5, 6, 7, 9\}$ and (e) the joint distribution for a randomly permuted pair of nodes $(N, M) = (5, 9)$ or $(9, 5)$.

Example 5.8. Figure 5.2 shows a small graph with some examples of joint distributions and their covariance.

Example 5.9 (**independent/identical pairs**)**.** If two random nodes N, M are independent, then their joint distribution decouples as $\Pr[(N, M) = (i, j)] = \Pr[N = i] \Pr[N = j]$ and $P = \tilde{\mathbf{f}}\tilde{\mathbf{g}}^T$ and the covariance is $\mathrm{cov}(N, M) = 0$ with respect to any distance. For two copies of the same random node N, the joint distribution is given by $\Pr[(N, N) = (i, j)] = \Pr[N = i]$ if $i = j$ and zero otherwise. In other words, we have $P = \mathrm{diag}(\tilde{\mathbf{f}})$ and find that $\mathrm{cov}(P) = \mathrm{var}(\tilde{\mathbf{f}})$ and $\mathrm{corr}(P) = 1$ with respect to any distance.

Example 5.10 (**network modularity**)**.** There has been a large research effort in the network science community (and more generally) to develop methods that identify groups of nodes which are tightly connected within each of the groups but poorly

connected between different groups; these mesoscale structures are often called *communities*[7] [100, 184, 200]. One approach to community detection[8] is based on Newman & Girvan's measure of *network modularity* [185], which assigns a score $M(\kappa)$ to each partitioning $\kappa : \mathcal{N} \to \{1, \dots, \ell\}$ of the nodes into ℓ partitions (communities); this is defined as

$$M(\kappa) := \frac{1}{n\langle k \rangle} \sum_{i,j \in \mathcal{N}} \left(c_{ij} - \frac{k_i k_j}{n\langle k \rangle} \right) \delta_{\kappa(i)\kappa(j)},$$

with the Kronecker delta defined as $\delta_{\kappa(i)\kappa(j)} = 1$ if two nodes are in the same partition $\kappa(i) = \kappa(j)$ and zero otherwise. If we use the "distance"[9] $d_\kappa(i,j) := 1 - \delta_{\kappa(i)\kappa(j)}$ and define the joint distribution $P(i,j) = c_{ij}/(n\langle k \rangle)$, then we find that

$$\mathrm{cov}_{d_\kappa}(P) = \frac{1}{2} M(\kappa).$$

In other words, network modularity is equal to the covariance of the end points of a uniform random link in the graph, measured with respect to the distance d_κ. In [73], we show that the related measure of 'Markov stability' [67] can also be obtained as a covariance.

Example 5.11 (**distance between distributions**). Let $\mathbf{f}, \mathbf{g} \in \Delta$ be two distributions on a graph and define the joint distribution $P = \frac{1}{2}(\mathbf{f}\mathbf{g}^T + \mathbf{g}\mathbf{f}^T)$. This can be seen as a random pair of nodes (N, M) distributed conditionally on a third random variable: define a coin flip $C \in \{H, T\}$ with uniform probabilities; if the coin lands on heads 'H', then $N \overset{d}{\sim} f$ and $M \overset{d}{\sim} g$, and if the coin lands on tails 'T', the other way around $N \overset{d}{\sim} g$ and $M \overset{d}{\sim} f$. The random variables (N, M) are conditionally independent once the outcome of C is known. The covariance of this distribution is equal to

$$\mathrm{cov}(P) = \frac{1}{8}(\mathbf{f} - \mathbf{g})^T D^{(2)} (\mathbf{f} - \mathbf{g}).$$

From Chapter 4, we know that if $D^{(2)}$ is the squared Euclidean distance matrix of a set of points in $\mathbb{R}^k$ (for some k) and with $k \times n$ matrix B containing these points as columns, then we may write

$$\mathrm{cov}(P) = -\frac{1}{4} \|B\mathbf{f} - B\mathbf{g}\|^2,$$

[7] These structures are also called 'assortative communities' to distinguish them from more general mesoscale community structures.

[8] There exist many different approaches to community detection, for instance based on spectral clustering or Bayesian inference; see [2, 100, 153, 193, 200, 250] for some reviews on community detection methods and their applications.

[9] This is not a semimetric since d_κ is zero between distinct nodes in the same partition. We furthermore note that d_κ can be seen as an induced 'distance' $d_\kappa = d' \circ \kappa$ where d' is a metric on the partitions defined by $d'(a,b) = 1 - \delta_{ab}$ for $1 \leq a, b \leq \ell$.

i.e., equal to the distance between the two points $B\mathbf{f}, B\mathbf{g} \in \mathbb{R}^k$. In other words, the covariance of the conditionally independent distribution of $\mathbf{f}$ and $\mathbf{g}$ is related to the distance between their embeddings in the case of a squared Euclidean metric. For instance, if we let $D^{(2)} = \Omega$, this is the distance between the points $\varphi(\mathbf{f})$ and $\varphi(\mathbf{g})$ in the hyperacute simplex associated with Ω.

5.1.2.1 Classical covariance

We show that the proposed graph covariance is a generalization of the classical covariance. For a random pair of elements (X, Y) from a subset $\mathcal{X} \subset \mathbb{R}$ of the real numbers, the expectation operator is defined as

$$\mathbb{E}(\gamma(X, Y)) := \sum_{i,j \in \mathcal{X}} P(i, j)\gamma(i, j) \text{ with } (X, Y) \overset{d}{\sim} P,$$

where $\gamma : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is some function on pairs of reals (this could be the identity function). The (classical) *covariance* is then defined as

$$\operatorname{cov}(P) = \mathbb{E}\left([X - \mathbb{E}(X)][Y - \mathbb{E}(Y)]\right) \text{ with } (X, Y) \overset{d}{\sim} P. \tag{5.4}$$

To arrive at an expression in terms of distances, we start by combining two identities that follow from the definition of covariance and linearity of expectation:

$$\begin{cases} \operatorname{cov}(P) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \\ \mathbb{E}([X - Y]^2) = \mathbb{E}(X^2) - 2\mathbb{E}(XY) + \mathbb{E}(Y^2) \end{cases}$$
$$\Rightarrow \operatorname{cov}(P) = \frac{1}{2}\left[\mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y^2) - \mathbb{E}([X - Y]^2)\right]. \tag{5.5}$$

Expanding the first three terms in expression (5.5) for the covariance, we find

$$\begin{aligned} &\mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y^2) \\ &= \sum_{i \in \mathcal{X}} \tilde{f}(i)i^2 - 2\sum_{i \in \mathcal{X}} \tilde{f}(i)i \sum_{j \in \mathcal{X}} \tilde{g}(j)j + \sum_{j \in \mathcal{X}} \tilde{g}(j)j^2 \\ &= \sum_{i \in \mathcal{X}} \tilde{f}(i)i \left(i - \sum_{j \in \mathcal{X}} \tilde{g}(j)j\right) + \sum_{j \in \mathcal{X}} \tilde{g}(j)j \left(j - \sum_{i \in \mathcal{X}} \tilde{f}(i)i\right) \\ &= \sum_{i \in \mathcal{X}} \tilde{f}(i)i \left(\sum_{j \in \mathcal{X}} \tilde{g}(j)(i - j)\right) + \sum_{j \in \mathcal{X}} \tilde{g}(j)j \left(\sum_{i \in \mathcal{X}} \tilde{f}(i)(j - i)\right) \\ &= \sum_{i,j \in \mathcal{X}} \tilde{f}(i)\tilde{g}(j)(i - j)^2. \end{aligned}$$

Including the last term in (5.5), the covariance can thus be written as

$$\mathrm{cov}(P) = \frac{1}{2} \sum_{i,j \in \mathcal{X}} (\tilde{f}(i)\tilde{g}(j) - P(i,j))(i-j)^2$$

where $(i-j)^2 = d^2(i,j)$ is the squared Euclidean distance between elements of $\mathcal{X}$. The proposed graph covariance (5.3) is a generalization of this expression with $\mathcal{N}$ a set of nodes and d any semimetric between pairs of nodes.

5.2 Experiments

As an example application of the (co)variance measures, we study a 'network of knowledge' made up of mathematical ideas, results and their relations. We use a dataset from [206] that consists of a list of mathematical concepts (theorems, lemmas and equations) compiled from four general-topic Wikipedia pages, and with links between the concepts inferred from hyperlinks between their respective Wikipedia pages. More information about the data retrieval and filtering of the data set can be found in [206]. The resulting network of concepts consists of $n = 1150$ nodes and $m = 4109$ links and is shown in Figure 5.3.

Figure 5.3: Hyperlink network of Wikipedia pages of the considered mathematical concepts (see [206]). The size of the nodes is proportional to their PageRank [34] (calculated with standard parameters) and the color coding corresponds to communities found using the Louvain algorithm [24].

The structural information of this network captures the relations between different concepts, as highlighted in Figure 5.3 by emphasizing 'central' nodes and coloring different communities (groups of clustered concepts). To study the functional aspects

of the network of knowledge — in particular, how the mathematical concepts are used — we use a corpus of 142071 papers from the arXiv preprint repository. For each paper, we count which of the mathematical concepts appear, and represent this by a uniform distribution over the used concepts. Every paper x thus has a corresponding subset $\mathcal{V}_x$ of concepts and distribution $f^{(x)}$ uniform over this set of concepts, as illustrated in Figure 5.4.

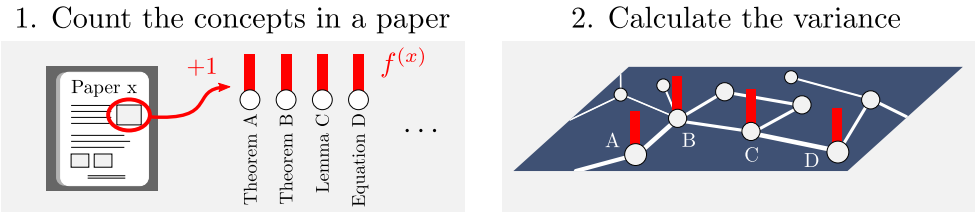


Figure 5.4: Illustration of the functional data.

A first question we address is whether mathematical papers contain 'coherent' sets of mathematical concepts. This is measured by the variance of the (empirical) concept distributions $f^{(x)}$ relative to the variance of 'virtual papers' generated by a null model that reproduces some of the observed statistics. More precisely, we consider null model distributions $\tilde{f}^{(x)}$ that are uniform over a number of randomly sampled concepts, according to their relative frequency over the full corpus. Figure 5.5 shows that the empirical paper distributions generally have a smaller variance than the virtual paper distributions. This is confirmed by performing the one-sided Mann–Whitney U test [128, Ch. 15] (see Section 5.2.0.1), from which we find with high significance (p-values $< 10^{-12}$) that the variance of a paper is typically smaller than what would be expected from the null model; more details on the test results are given in Section 5.2.0.1. Intuitively, this observation reflects the idea that a group of mathematical concepts is more likely to be considered in a paper if these concepts are related (in the constructed network). Furthermore, Figure 5.5 clearly indicates a 'typical' range of variance values which could be used to identify papers with exceptionally small (or large) variances. As a second application, the variance of paper distributions can be used for a (qualitative) comparison between different fields of study. In the bottom plots of Figure 5.5, different sub-fields are ranked based on their variance distributions, from more concentrated on small variances on the top left to more concentrated on large values in the bottom right. These comparisons are confirmed by one-sided Mann–Whitney U tests.

On an aggregate level, we study the corpus of papers by counting the co-occurrences

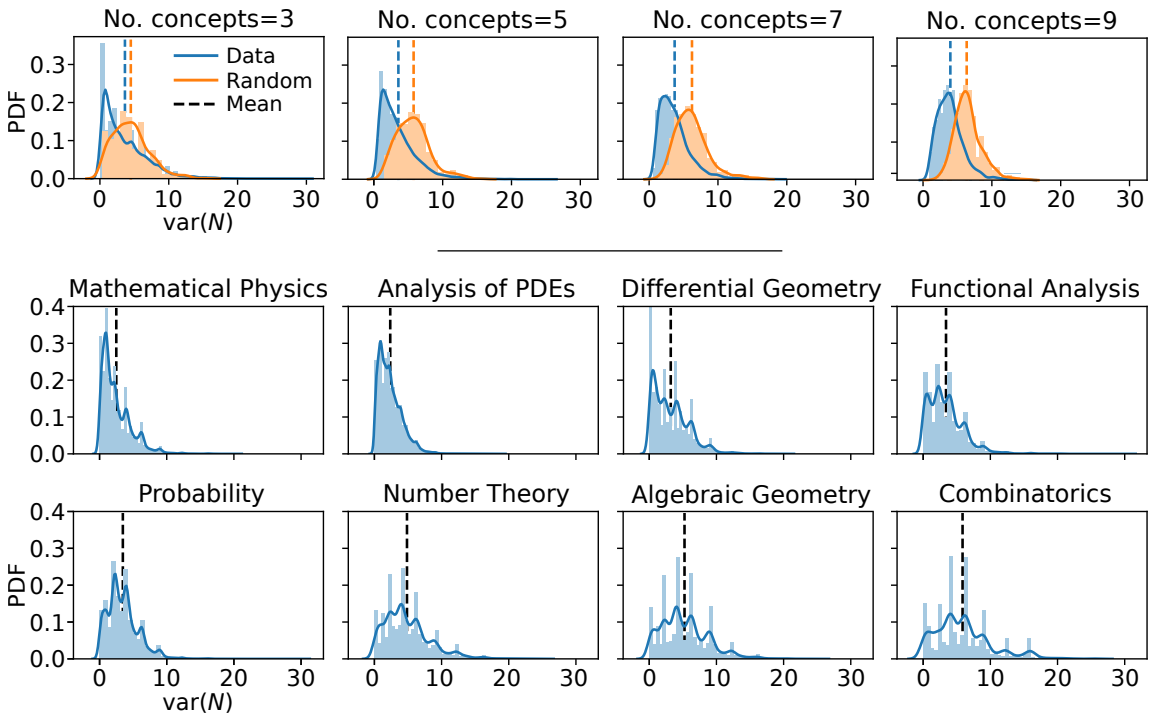


Figure 5.5: (Top panels) Distribution of network variances of concept distributions $f^{(x)}$ in papers, calculated with respect to the shortest-path distance in the Wikipedia network. In each panel, the probability density function (pdf) of variances is calculated for papers containing 3 (resp. 5, 7 and 9) concepts present in the Wikipedia network, and compared with the variance pdf for a collection of (null model) virtual papers with the same number of concepts. The empirical variance distributions are concentrated on smaller variances compared to the null model variances, as confirmed by the one-sided Mann–Whitney U test. (Bottom panels) Distribution of network variances of concept distributions $f^{(x)}$ for arXiv papers from different subfields. The fields are ordered from top left to bottom right according to concentration on increasing variances, again based on the Mann–Whitney U test.

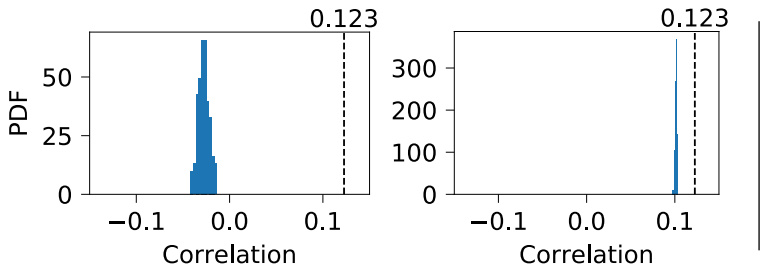
of pairs of concepts over all papers. This gives a (symmetric) joint distribution P, where $P(i,j)$ is proportional to the frequency of co-occurrence of concepts i and j. For instance, if concept i and concept j appear in 10% of the papers, then $P(i,j) = 0.1$ before P is normalized to have unit sum. To quantify how much the function of the network (i.e. concepts being used together) aligns with its structure (i.e. relations between the concepts), we again compare the covariance of the empirical distribution P to null model distributions. In the first model, the distribution P is left constant but the underlying graph is randomized, while in the second model, the distribution P is randomized with the graph left constant. Since we are comparing joint distributions with potentially different marginals, we use the correlation $\mathrm{corr}(P) = \mathrm{cov}(P)/\mathrm{var}(\tilde{f})$.

In the first null model, we perform a degree-preserving rewiring [176] of the Wikipedia network while keeping the joint distribution constant. As seen in the left panel of Figure 5.6, measuring the covariance of P with respect to these randomized graphs yields significantly smaller correlation values compared to the observed correlation; this suggests that the empirical correlation is not simply a consequence of the degree distribution of the network. The second null model consists of a marginal-preserving randomization of P while leaving the graph intact. We pick two pairs of nodes (i,j) and (i',j') with non-zero joint probabilities and reshuffle the joint probabilities as

$$P \to P - \alpha(v_{ij} + v_{i'j'} - v_{ij'} - v_{i'j})$$

where $v_{ij} = \mathbf{e}_i\mathbf{e}_j^T + \mathbf{e}_j\mathbf{e}_i^T$ and with a uniform random value $\alpha < \min(P(i,j), P(i',j'))$. In other words, there is a shift of probability mass α from $(i,j) \to (i,j')$ and from $(i',j') \to (i',j)$. We repeat this procedure until all pairs of nodes have been involved, which produces a randomized joint distribution P' with the same marginals as P. The middle panel in Figure 5.6 shows that these randomized joint distributions are concentrated on significantly lower correlation values. We remark that the empirical correlation is highly atypical for both null model distributions and it is thus very likely that these null models discard too much structure to give a reliable baseline for the empirical correlation. A further development of appropriate null models for (joint) distributions on a network is an interesting line of further research. The table on the right in Figure 5.6 reports the co-occurrence correlations for a number of sub-fields.

The above analysis illustrates a practical scenario where structural and functional data is available, for which our variance and covariance measures enable a unified treatment of the system data. Our (co)variance measures are of course not restricted to this specific example, but can be applied in the context of many other network problems in which a combination of structural and functional data is important.



Field	Correlation
Analysis of PDEs	0.021
Mathematical Physics	0.065
Probability	0.042
Algebraic Geometry	0.051
Differential Geometry	0.054
Number Theory	0.058
Dynamical Systems	0.081
Combinatorics	0.073
Functional Analysis	0.067

Figure 5.6: (Left and middle figures) Comparison between empirical correlation of the co-occurrence distribution P on the Wikipedia network of concepts and two null models, measured with respect to the shortest-path distance. The left panel shows the correlation of the empirical distribution P on the original Wikipedia networks (dashed line) and 100 realizations of a degree-preserving randomization of the network (blue bars). The middle panel shows the correlation of the empirical distribution P (dashed line) and 100 realizations of a marginal-preserving randomization of P (blue bars) on the Wikipedia network. (Right table) Correlation of the joint probability distribution of concepts for papers of different sub-fields.

5.2.0.1 Statistical tests

Figure 5.5 shows a comparison between the variances of concept distributions obtained from real arXiv papers and the variances of concept distributions of virtual papers constructed based on the overall frequencies of the concepts. Visually, the variance distribution of the real papers appears to be concentrated on smaller values than the variance distribution of the virtual papers.

To quantify this observation, we perform a one-sided Mann–Whitney U test[10] [128, Ch. 15] comparing the two distributions (for a given number of concepts in the papers), and the outcome of this test is recorded in Table 5.1. In all cases, the p-value ($< 10^{-12}$) rejects the Mann–Whitney null hypothesis with high significance, in favour of the alternative hypothesis that samples from the null model distribution are typically larger than samples from the empirical distribution — i.e. $\Pr[\mathrm{var}(f_R) < \mathrm{var}(f_V)] > 1/2$, where f_R is the concept distribution of a randomly selected real paper and f_V the distribution of a randomly selected virtual paper. In this sense, the variance of the papers is typically smaller than what would be expected from the null model. We note again that the extremely high significance values indicate that the null model distribution likely does not give a reliable baseline for comparison.

[10]The test statistic U counts the number of samples from one distribution that are larger than samples from the other distribution and works under the assumption that all samples are independent.

no. concepts	no. samples (real/virt.)	U value	U norm.	p-value
3 concepts	25716/400	6202641.5	0.60	6.52e-13
4 concepts	17578/400	4995856	0.71	1.71e-47
5 concepts	11742/400	3553001.5	0.76	1.13e-68
6 concepts	7449/400	2338463.5	0.78	1.19e-82
7 concepts	4545/400	1473777.5	0.81	7.06e-95
8 concepts	2629/400	875333.5	0.83	2.27e-102
9 concepts	1428/400	478299	0.84	4.65e-95
10 concepts	785/400	269039	0.86	2.91e-90

Table 5.1: Mann–Whitney U test statistics comparing the empirical distribution of paper variances to the distribution of (null model) virtual paper variances, for papers containing 3–10 concepts. Tabulated from left to right are the number of concepts, the number of real and virtual papers (no. samples), the U value of the Mann–Whitney test, the normalized U value $U/(n_{\text{real}}n_{\text{virt.}})$ and the p-value of the U test statistic.

5.3 The maximum-variance problem

This section deals with the maximum-variance problem: what is the largest possible variance on a given graph with distance (semimetric) d, and what are the distributions that achieve it? Starting from general distances, we characterize the solution(s) to the maximum-variance problem and show how adding further constraints to the distances makes the problem increasingly tractable. In the rest of this section, we consider variance with respect to square root distances such that the variance is a quadratic form determined by the distance matrix D.

Hjorth et al. [127] studied an equivalent optimization problem in the context of finding a generalized ('transfinite') diameter of metric spaces, and found that a unique solution exists for strict negative type metrics; this is our Proposition 5.20. Dankelmann [66] considered the quadratic form $\mathbf{f}^T D\mathbf{f}$ for the geodesic graph distance and solved the corresponding maximization problem for trees and cycle graphs. Bapat, Neogy et al. considered the quadratic form $\mathbf{f}^T\Omega\mathbf{f}$ in [12, 83] and found that the corresponding maximum-variance problem can be solved efficiently.

5.3.1 General distances

Let d be any distance (semimetric) and var_d the corresponding variance; the *maximum-variance problem* is defined as follows:

$$\begin{aligned} &\text{maximize } \frac{1}{2}\mathbf{f}^T D\mathbf{f} \\ &\text{subject to } \mathbf{f} \in \Delta. \end{aligned} \tag{5.6}$$

Intuitively, the maximum-variance problem asks to distribute a fixed probability mass over the nodes of a graph such that sampling two nodes according to the resulting distribution yields nodes that are as far apart on average as possible. While a first guess for such a maximizing distribution might be to put probability $1/2$ on the two nodes which are furthest apart, this means that in $1/4^{\text{th}}$ of the cases two identical nodes (at distance 0) are selected[11]. In addition to selecting far apart nodes, the maximum distribution thus also favours being spread out over a larger set of nodes and the problem is more complex than it might appear at first sight.

The solution to the maximum-variance problem is the *maximum variance* and is denoted by $\mathrm{var}_d^\star$ (or simply $\mathrm{var}^\star$). Since the variance is a continuous real function on a convex set Δ, its maximum is achieved by at least one distribution, which we call *a maximum variance distribution* and denote by $\mathbf{f}^\star$. The support of this distribution is denoted by $\mathcal{V}_d^\star := \{i : f_i^\star > 0\}$ (or $\mathcal{V}^\star$) and is called the *maximum variance support*.

To characterize the maximum variance, we start by deriving some necessary conditions on the maximum variance distribution:

Proposition 5.12. *A necessary condition for maximum variance distributions is*

$$\mathbf{f} \in \Delta \text{ and } (\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f} \geq 0 \text{ for all } i \in \mathrm{supp}(\mathbf{f}) \text{ and all } j. \tag{5.7}$$

Proof. Let d be a distance with distance matrix D and $\mathbf{f}^\star$ a maximum variance distribution with support $\mathcal{V}^\star$. We consider the vector $\mathbf{f}' = \mathbf{f}^\star - \epsilon(\mathbf{e}_i - \mathbf{e}_j)$ with $i \in \mathcal{V}^\star$ and any j; for small enough $\epsilon > 0$, this vector is a distribution since $f_i^\star > 0$. By maximality of $\mathbf{f}^\star$, we must have $\mathrm{var}_d(\mathbf{f}') \leq \mathrm{var}_d(\mathbf{f}^\star)$, which can be written as

$$\begin{aligned} &\mathbf{f}^{\star T} D\mathbf{f}^\star - 2\epsilon(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f}^\star + \epsilon^2(\mathbf{e}_i - \mathbf{e}_j)^T D(\mathbf{e}_i - \mathbf{e}_j) \leq \mathbf{f}^{\star T} D\mathbf{f}^\star \\ \Leftrightarrow &(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f}^\star \geq -\epsilon d(i,j) \text{ for all } i \in \mathcal{V}^\star \text{ and all } \epsilon > 0. \end{aligned} \tag{5.8}$$

[11] This first guess yields a variance of $\frac{1}{4}\,\mathrm{diam}(d)$. Proposition 5.17 says that this variance is achieved if and only if the graph is a path graph and that for graphs that are not paths, the maximum variance can be up to a factor 2 larger.

Now suppose for contradiction that there exists some $i \in \mathcal{V}^\star$ and j such that $(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f}^\star < 0$. Since $i \neq j$ and thus $d(i,j) > 0$, we can then choose an ϵ with $\epsilon < |(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f}^\star|/d(i,j)$ in equation (5.8); this produces a contradiction:

$$(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f}^\star \geq -\epsilon d(i,j) > (\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f}^\star.$$

This shows that no $i \in \mathcal{V}^\star$ and j with a negative value for $(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f}^\star$ can exist, which completes the proof. □

The necessary condition (5.7) encompasses three different aspects of maximum variance distributions: a basic feasibility condition $\mathbf{f} \in \Delta$, which says that $\mathbf{f}$ must be a distribution; a 'local' optimality condition, which reflects optimality with respect to perturbations of the distribution without changing its support — this corresponds to the inequalities with $j \in \operatorname{supp}(\mathbf{f})$ — and finally, a 'nonlocal' optimality condition, which reflects optimality with respect to perturbations that change the support — this corresponds to the inequalities with $j \notin \operatorname{supp}(\mathbf{f})$.

The local optimality condition can be translated to the following requirement on maximum variance distributions.

Proposition 5.13. *A necessary condition for maximum variance distributions is*

$$\mathbf{f} \in \Delta \text{ with support } \mathcal{V}, \text{ and } D_{\mathcal{V}\mathcal{V}}\mathbf{f}_\mathcal{V} = (\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u})^{-1}\mathbf{u}. \tag{5.9}$$

Proof. Let $\mathbf{f}$ be a maximum variance distribution with support $\mathcal{V}$. By the necessary conditions (5.7), we know that

$$(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f} \geq 0 \text{ and } (\mathbf{e}_j - \mathbf{e}_i)^T D\mathbf{f} \geq 0 \text{ for all } i,j \in \mathcal{V},$$

which implies that $(\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f} = 0$ for all $i,j \in \mathcal{V}$, and thus that $D\mathbf{f}$ is constant on the support of $\mathbf{f}$. In other words,

$$D_{\mathcal{V}\mathcal{V}}\mathbf{f}_\mathcal{V} = \alpha\mathbf{u} \text{ for some constant } \alpha.$$

Since $D_{\mathcal{V}\mathcal{V}}$ is a distance matrix with positive off-diagonals (and $|\mathcal{V}| \geq 2$; see Proposition 5.30) and $\mathbf{f}_\mathcal{V}$ is a positive vector, we know that $\alpha > 0$. This implies that $\mathbf{u} \in \ker(D_{\mathcal{V}\mathcal{V}})^\perp$ since no solution could exist otherwise, and thus by properties of the Moore–Penrose pseudoinverse (Proposition A.5) that $D_{\mathcal{V}\mathcal{V}}[D_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u} = \mathbf{u}$. We then find

$$\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u} = \alpha^{-1}\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger D_{\mathcal{V}\mathcal{V}}\mathbf{f}_\mathcal{V} = \alpha^{-1}\mathbf{u}^T\mathbf{f}_\mathcal{V} = \alpha^{-1},$$

which completes the proof. □

Proposition 5.13 implies the following result for the maximum variance:

Corollary 5.14. *The maximum variance is equal to* $\text{var}_d^\star = \frac{1}{2}(\mathbf{u}^T[D_{\mathcal{V}^\star\mathcal{V}^\star}]^\dagger\mathbf{u})^{-1}$.

Proof. Let $\mathbf{f}^\star$ be a maximum variance distribution with support $\mathcal{V}^\star$. The maximum variance is then equal to

$$\text{var}_d^\star = \frac{1}{2}\mathbf{f}^{\star T} D\mathbf{f}^\star = \frac{1}{2}\mathbf{f}_{\mathcal{V}^\star}^{\star T} D_{\mathcal{V}^\star\mathcal{V}^\star}\mathbf{f}_{\mathcal{V}^\star}^\star \overset{\text{(Prop. 5.13)}}{=} \frac{1}{2}(\mathbf{u}^T[D_{\mathcal{V}^\star\mathcal{V}^\star}]^\dagger\mathbf{u})^{-1},$$

as required. □

In other words, the maximum variance can be expressed in terms of a submatrix of the distance matrix, without explicit reference to the maximum variance distribution. Of course, the question remains how to find the maximum variance support $\mathcal{V}^\star$. Combining Proposition 5.13 and Corollary 5.14, we find the following reformulation of the maximum-variance problem in terms of submatrices of the distance matrix:

Theorem 5.15. *The maximum-variance problem is equivalent to*

$$\begin{aligned}&\text{maximize } \frac{1}{2}(\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u})^{-1}\\ &\text{subject to } \mathcal{V} \subseteq \mathcal{N} \text{ for which } D_{\mathcal{V}\mathcal{V}}\mathbf{x} = \mathbf{u} \text{ has a nonnegative solution}\end{aligned}$$

Proof. We construct an equivalent optimization problem to the maximum-variance problem (5.6) by adding additional constraints to the optimization domain $\Delta \to \widetilde{\Delta}$, such that the restricted domain $\widetilde{\Delta}$ still contains the solution. We define

$$\widetilde{\Delta} := \{\mathbf{f} \in \Delta \text{ with support } \mathcal{V} \text{ such that } D_{\mathcal{V}\mathcal{V}}\mathbf{f}_\mathcal{V} = (\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u})^{-1}\mathbf{u}\}.$$

A solution with support $\mathcal{V}$ exists if and only if $D_{\mathcal{V}\mathcal{V}}\mathbf{x} = \mathbf{u}$ has a nonnegative and nonzero solution for $\mathbf{x}$. For any candidate distribution $\mathbf{f} \in \widetilde{\Delta}$ in the support, the maximum variance is given by

$$\text{var}_d(\mathbf{f}) = \frac{1}{2}\mathbf{f}^T D\mathbf{f} = \frac{1}{2}\mathbf{f}_\mathcal{V}^T D_{\mathcal{V}\mathcal{V}}\mathbf{f}_\mathcal{V} = \tfrac{1}{2}(\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u})^{-1},$$

only dependent on the support $\mathcal{V}$. Thus all possible variances that correspond to distributions in $\widetilde{\Delta}$ are covered if we consider $\frac{1}{2}(\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u})^{-1}$ over all subsets $\mathcal{V}$ for which $D_{\mathcal{V}\mathcal{V}}\mathbf{x} = \mathbf{u}$ has a nonnegative nonzero solution. This completes the proof. □

The maximum-variance problem can thus be solved by iterating over all subsets, checking whether a nonzero nonnegative solution exists — this amounts to determining whether a certain system of linear equations $A\mathbf{x} \geq \mathbf{b}$ has a solution, which is equivalent to solving a linear program in general — and finally comparing which of these subsets with solutions produces the largest value for $\frac{1}{2}(\mathbf{u}^\mathbf{T}[\mathbf{D}_{\mathcal{V}\mathcal{V}}]^\dagger\mathbf{u})^{-1}$. In general, this procedure requires to check all 2^n subsets of nodes, which is not practical. The following example based on [158] shows that the maximum-variance problem is in fact NP-hard.

Example 5.16 (**max variance and max clique**). Let G be an unweighted graph and let $\tilde{d}(i,j) = 1$ if $i \sim j$ and zero otherwise; in other words, the distance matrix $\widetilde{D}$ is the adjacency matrix of G; see e.g. [28]. As discussed in [158, Cor. 3], the maximum-variance problem $\max_{\mathbf{f}\in\Delta} \frac{1}{2}\mathbf{f}^T\widetilde{D}\mathbf{f}$ has as solution $\mathrm{var}_{\tilde{d}}^\star = \frac{1}{2}(1-\mathrm{cl}(G)^{-1})$, where $\mathrm{cl}(G)$ is the clique number of G, the size of the largest clique in the graph. Since calculating the clique number is NP-hard in general [139], so must be the maximum-variance problem with respect to $\tilde{d}$. While $\tilde{d}$ is not a semimetric because it can be zero between distinct nodes, we may define the semimetric $d_\epsilon(i,j) = \tilde{d}(i,j) + \epsilon$ if $i \neq j$ and $d_\epsilon(i,i) = \tilde{d}(i,i)$ otherwise. By letting $\epsilon > 0$ be small enough, the maximum variance with respect to d_ϵ resolves to the rational solution[12] found by $\tilde{d}$, and thus the maximum-variance problem for general semimetrics must be NP-hard as well.

While calculating the maximum variance is thus not practical in general, we have the following estimate in terms of the diameter $\mathrm{diam}(d) := \max_{i,j} d(i,j)$:

Proposition 5.17. *The maximum variance of a graph with metric d is bounded by its diameter as*

$$\frac{1}{4}\mathrm{diam}(d) \leq \mathrm{var}_d^\star \leq \frac{n-1}{2n}\mathrm{diam}(d).$$

Equality in the lower bound (resp. upper bound) holds if and only if d is the resistance distance on a path graph (resp. complete graph).

Proof. *(lower bound)* The maximum variance is at least as large as the variance of any particular distribution $\mathbf{f}$. By choosing $\mathbf{f} = \frac{1}{2}(\mathbf{e}_i + \mathbf{e}_j)$ for some pair of nodes at distance $d(i,j) = \mathrm{diam}(d)$, we then find that

$$\frac{1}{4}\mathrm{diam}(d) = \mathrm{var}_d\left(\tfrac{1}{2}(\mathbf{e}_i + \mathbf{e}_j)\right) \leq \mathrm{var}_d^\star.$$

Next, we will derive the conditions for equality in the lower bound.

(lower bound equality, forward) Let G be a path graph with resistance distance $d = \omega$. Then Example 6.7 shows that the resistance curvature is $p = 1/2$ on the end nodes of the path and $p = 0$ for all other nodes. In particular, the resistance curvature is nonnegative. By Corollary 5.24 for the maximum variance distribution with respect to the resistance distance, we then know that the maximum variance distribution is equal to the resistance curvature $\mathbf{f}^\star = \mathbf{p} = \frac{1}{2}(\mathbf{e}_i + \mathbf{e}_j)$ where i and j are the end nodes of the path. This yields $\mathrm{var}_d^\star = \frac{1}{4}\mathrm{diam}(d)$ as required.

[12] For an n-node graph, the clique number $\mathrm{cl}(G)$ is an integer between 2 and n and thus $\mathrm{var}_{\tilde{d}}^\star = \frac{1}{2}(1-\mathrm{cl}(G)^{-1})$ must be a positive rational number.

(lower bound equality, converse) Conversely, if $\operatorname{var}_d^\star = \frac{1}{4}\operatorname{diam}(d)$, then a possible maximum variance distribution is $\mathbf{f}^\star = \frac{1}{2}(\mathbf{e}_i + \mathbf{e}_j)$ for some pair of nodes at distance $d(i,j) = \operatorname{diam}(d)$. From the necessary conditions (5.7), we then know that

$$(\mathbf{e}_i - \mathbf{e}_x)^T D(\mathbf{e}_i + \mathbf{e}_j) \geq 0 \text{ for all } x.$$

For $x = i, j$, this yields zero and for $x \neq i, j$, this translates to

$$0 \leq d(i,j) - d(x,j) - d(x,i) \leq 0,$$

where the second inequality is the triangle inequality for d. This implies that $d(i,j) = d(i,x) + d(x,j)$ for all $x \neq i, j$; in other words, there are two nodes i, j such that all other nodes x are a cut node between i and j. This implies that d is the resistance metric on a path with end nodes i, j.

(upper bound) The upper bound follows from

$$\operatorname{var}_d^\star = \frac{1}{2}\sum_{i,j} f_i^\star f_j^\star d(i,j) \leq \frac{1}{2}\operatorname{diam}(d)\sum_{i \neq j} f_i^\star f_j^\star = \frac{1}{2}\operatorname{diam}(d)\left(1 - \|\mathbf{f}^\star\|^2\right),$$

and the fact that $\|\mathbf{f}\|^2 \geq \frac{1}{n}$ for all $\mathbf{f} \in \Delta$ with equality if and only if $\mathbf{f} = \mathbf{u}/n$. Equality in the upper bound is thus achieved if and only if $\mathbf{f}^\star = \mathbf{u}/n$ and $d(i,j) = \operatorname{diam}(d)$ for all pairs of nodes $i \neq j$, which corresponds to the resistance metric on a complete graph. This completes the proof. □

We note that the lower-bound equality result requires d to be a metric while the inequalities work for general distances. Furthermore, we recall that the resistance distance is equal to the shortest-path distance for path graphs. The diameter of a metric space thus determines its maximum variance up to a factor of at most 2. These bounds were inspired by the results of Steinerberger [226] in a different context and proven using different techniques.

As noted before, the maximum-variance problem is closely related to the work of Leinster et al. on diversity measures [157, 158]. We briefly discuss their main results as they pertain to the maximum-variance problem.

A *similarity matrix* Z is a symmetric $n \times n$ matrix with unit diagonal $(Z)_{ii} = 1$ and bounded off-diagonal $0 \leq (Z)_{ij} \leq 1$ for all $i \neq j$. The *diversity of order* q of a distribution $\mathbf{f}$ with respect to a similarity matrix Z is defined as

$$\operatorname{div}_q^Z(\mathbf{f}) := \left(\mathbf{f}^T(Z\mathbf{f})^{q-1}\right)^{\frac{1}{1-q}} \text{ for some } \mathbf{f} \in \Delta,$$

where the powers are entrywise and $q \in [0, \infty]$. The maximum diversity problem asks to maximize div_q^Z over the probability simplex Δ, and the following result was shown by Leinster & Meckes [158, Thm. 1 & 2]:

Theorem 5.18. *(i) There exists a distribution* $\mathbf{f}^\star$ *that maximizes* div_q^Z *for all* $q \in [0, \infty]$. *Moreover, the maximum diversity* $\operatorname{div}_q^Z(\mathbf{f}^\star)$ *is independent of* q. *(ii) The maximum diversity problem is equivalent to* $\max_{\mathcal{V}} \mathbf{u}^T (Z_{\mathcal{V}\mathcal{V}})^\dagger \mathbf{u}$ *over subsets* $\mathcal{V}$, *for which* $Z_{\mathcal{V}\mathcal{V}}\mathbf{x} = \mathbf{u}$ *has a nonnegative solution.*

Result (i) says that the maximum diversity and the maximum diversity distribution are generic for a given similarity matrix, as they do not depend on the specific order of the measure. Result (ii) is called the 'computation theorem' and shows how the maximum diversity is determined by submatrices of the similarity matrix, in close analogy with Theorem 5.15 for the maximum-variance problem.

The relation between the variance with respect to a distance matrix D and the q-diversity with respect to a similarity matrix Z is found by defining the similarity matrix[13] $Z_\delta := \mathbf{u}\mathbf{u}^T - \delta D$ with $\delta \leq \operatorname{diam}(d)^{-1}$ and distance matrix D, for which we find

$$\operatorname{div}_2^{Z_\delta}(\mathbf{f}) = \frac{1}{1 - 2\delta \operatorname{var}_d(\mathbf{f})}.$$

As a consequence, finding the maximum diversity with respect to Z_δ corresponds to finding the maximum variance with respect to D. By Theorem 5.18, this implies that the maximum-variance problem and its solutions are relevant for a broad class of diversity measures ($0 \leq q \leq \infty$), and not just for the 'simple' quadratic variance ($q = 2$).

5.3.2 Negative type metrics

Since the general maximum-variance problem is NP-hard, we further restrict the class of distances under consideration. As a first result, we find that the variance has additional desirable properties for negative type metrics:

Proposition 5.19. *The variance* var_d *is (strictly) concave on* Δ *if and only if the distance matrix* D *is a squared Euclidean distance matrix (of a simplex). In particular,* var_d *is (strictly) concave for (strict) negative type metrics.*

Proof. Let G be a graph with metric d and corresponding variance var_d. By definition of concavity, the variance is concave on Δ if and only if

$$\theta \operatorname{var}_d(\mathbf{f}) + (1-\theta)\operatorname{var}_d(\mathbf{g}) \leq \operatorname{var}_d(\theta\mathbf{f} + (1-\theta)\mathbf{g}) \text{ for all } \mathbf{f}, \mathbf{g} \in \Delta \text{ and } \theta \in [0,1].$$

[13] We note that this is asymptotically equivalent to $(Z)_{ij} = \exp(-d(i,j)t)$ as $t \to \infty$. This general mapping from distances to similarity matrices is central to the theory of magnitude of metric spaces [154].

Introducing the expression for variance, this can be written as

$$\begin{aligned}(\text{var}_d \text{ concave}) &\Leftrightarrow \theta\mathbf{f}^T D\mathbf{f} + (1-\theta)\mathbf{g}^T D\mathbf{g} \leq (\theta\mathbf{f} + (1-\theta)\mathbf{g})^T D(\theta\mathbf{f} + (1-\theta)\mathbf{g}) \\ &\Leftrightarrow \theta(1-\theta)(\mathbf{f}-\mathbf{g})^T D(\mathbf{f}-\mathbf{g}) \leq 0 \text{ for all } \mathbf{f},\mathbf{g} \in \Delta \text{ and } \theta \in [0,1].\end{aligned} \tag{5.10}$$

Since $\theta(1-\theta) \geq 0$ and any vector $\mathbf{x} \in \text{span}(\mathbf{u})^\perp$ can be written as $\mathbf{x} = \alpha(\mathbf{f}-\mathbf{g})$ for some $\mathbf{f},\mathbf{g} \in \Delta$ and scalar α, we find that concavity is equivalent to

$$(\text{var}_d \text{ concave}) \Leftrightarrow \mathbf{x}^T D\mathbf{x} \leq 0 \text{ for all } \mathbf{x} \in \text{span}(\mathbf{u})^\perp \Leftrightarrow (D \text{ is an SED matrix}),$$

where 'SED' abbreviates squared Euclidean distance matrix (see Chapter 4). Similarly, for strict concavity, the inequality in (5.10) is strict and $\theta \in (0,1)$ in combination with $\mathbf{f} \neq \mathbf{g}$ is equivalent to $\mathbf{x}^T D\mathbf{x} < 0$ for all nonzero $\mathbf{x} \in \text{span}(\mathbf{u})^\perp$, which is equivalent to D being the squared Euclidean distance matrix of a simplex.

By Proposition 4.12, the distance matrix of a negative type metric is a squared Euclidean distance matrix and by Proposition 4.16, the distance matrix of a strict negative type metric is the squared Euclidean distance matrix of the vertices of a simplex. This completes the proof. □

It was observed in [158] that the maximum diversity problem simplifies significantly in the case of positive definite similarity matrices, which relates to Proposition 5.19. Concavity is a key property in the context of optimization, since it guarantees that local maxima are also global maxima [31]. For the maximum-variance problem, we find the following result:

Proposition 5.20. *A necessary and sufficient condition for a maximum variance distribution of a graph with negative type metric d is*

$$\mathbf{f} \in \Delta \text{ and } (\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f} \geq 0 \text{ for all } i \in \text{supp}(\mathbf{f}) \text{ and all } j,$$

and a solution can be calculated in polynomial time in n. For a strict negative type metric, these conditions are satisfied by a unique distribution.

Proof. For a negative type metric, the variance is a concave function on a convex set Δ. The maximum-variance problem is thus a *convex quadratic program* which can be solved in polynomial time in n; see [31].

(necessary) By Proposition 5.12, condition (5.7) is a necessary condition for the maximum variance distribution with respect to any distance, and thus in particular with respect to negative type metrics. We now show that the conditions are also sufficient for negative type metrics.

(sufficient) Let $\mathbf{f}$ and $\mathbf{g}$ be any two distributions and $\theta \in [0,1]$. By concavity of the variance for negative type metrics, we then have

$$\begin{aligned}
&\operatorname{var}_d(\theta\mathbf{f} + (1-\theta)\mathbf{g}) \geq \theta \operatorname{var}_d(\mathbf{f}) + (1-\theta)\operatorname{var}_d(\mathbf{g}) \\
\Leftrightarrow\ &\theta(\theta-1)\mathbf{g}^T D\mathbf{g} \geq \theta(\theta-1)\mathbf{f}^T D\mathbf{f} - 2\theta(\theta-1)(\mathbf{f}-\mathbf{g})^T D\mathbf{f} \\
\Leftrightarrow\ &\operatorname{var}_d(\mathbf{g}) \leq \operatorname{var}_d(\mathbf{f}) - 2\sum_{i\in\operatorname{supp}(\mathbf{f}),j} f_i g_j (\mathbf{e}_i - \mathbf{e}_j)^T D\mathbf{f} \quad \text{(since } \theta \leq 1\text{)}.
\end{aligned}$$

If $\mathbf{f}$ satisfies conditions (5.7), then the second term on the right-hand side is non-positive, which implies that $\operatorname{var}_d(\mathbf{g}) \leq \operatorname{var}_d(\mathbf{f})$ for all $\mathbf{g} \in \Delta$ and thus that $\mathbf{f}$ is the maximum variance distribution as required.

(uniqueness) Next, suppose for contradiction that two distinct distributions $\mathbf{f}$ and $\mathbf{f}'$ achieve the maximum variance $\operatorname{var}_d^\star$ for a strict negative type metric d. Then by strict concavity, we find that $\operatorname{var}_d\left(\frac{1}{2}(\mathbf{f}+\mathbf{f}')\right) > \operatorname{var}_d^\star$ which is a contradiction against maximality and thus proves that there is a unique maximum variance distribution for strict negative type metrics. □

In practice, the maximum variance can be calculated using standard convex programming software such as the CVX package [112]. We recall that (strict) negative type metrics appear commonly, for instance as point sets in $\mathbb{R}^d$, in hyperbolic space or on the sphere. Consequently, if a graph (or any dataset) is embedded into one of these spaces as $\phi : \mathcal{N} \to M$, then the maximum-variance problem can be solved efficiently for the graph with the induced metric $d_M \circ \phi$.

5.3.3 Resistance metrics

Translating the maximum variance results for general and negative type metrics to the special case of effective resistances, we find the following characterization:

Proposition 5.21. *The maximum variance of a graph with respect to the resistance distance $d = \omega$ is determined by a unique subset of nodes $\mathcal{V}^\star$ and the corresponding Kron reduction $G^\star = G/\mathcal{V}_\omega^{\star c}$: the maximum variance is equal to the resistance radius $\operatorname{var}_\omega^\star = \sigma_\star^2$ and the maximum variance distribution is equal to the resistance curvature $\mathbf{f}^\star = \mathbf{p}^\star$ of $G^\star$.*

Proof. Since the resistance metric has strict negative type, there is a unique maximum variance distribution $\mathbf{f}^\star$. Following Proposition 5.13, this distribution satisfies

$$\Omega_{\mathcal{V}^\star\mathcal{V}^\star}\mathbf{f}_{\mathcal{V}^\star}^\star = 2\operatorname{var}_\omega^\star \mathbf{u},$$

which by inverting the resistance submatrix, yields $\mathbf{f}^\star_{\mathcal{V}} = \mathbf{p}_{\mathcal{V}}$ and thus $\mathrm{var}^\star_\omega = \sigma^2_\star$, as required. □

In context of the resistance metric, we will thus use $\sigma^2_\star$ to refer to the maximum variance and $\mathbf{p}^\star$ for the maximum variance distribution[14], both of which are determined by the maximum variance support $\mathcal{V}^\star$. According to Proposition 5.20, this maximum variance support $\mathcal{V}^\star$ can be calculated in polynomial time using quadratic programming. However, making use of the specific properties of the effective resistance — in particular, positivity of the link weights of the associated graph — we obtain a simple 'greedy' method to solve the maximum-variance problem. We start from the following observation:

Lemma 5.22. *$p_i \geq p^\star_i > 0$ for all $i \in \mathcal{V}^\star$.*

Proof. Let G be a graph with its resistance metric ω and maximum variance distribution $\mathbf{p}^\star$. Condition (5.7) for the maximum variance distribution says that

$$(\mathbf{e}_i - \mathbf{e}_j)^T \Omega \mathbf{p}^\star \geq 0 \text{ for all } i \in \mathcal{V}^\star \text{ and all } j.$$

Summing $c_{ij}(\mathbf{e}_i - \mathbf{e}_j)^T \Omega \mathbf{p}^\star$ over the neighbours of a node $i \in \mathcal{V}^\star$, we then find the inequality

$$\sum_{j:j\sim i} c_{ij}(\mathbf{e}_i - \mathbf{e}_j)^T \Omega \mathbf{p}^\star = \mathbf{e}_i^T Q \Omega \mathbf{p}^\star \geq 0 \text{ for all } i \in \mathcal{V}^\star.$$

From the $Q\Omega$ identity (3.7) we then find that

$$\mathbf{e}_i^T(-2I + 2\mathbf{p}\mathbf{u}^T)\mathbf{p}^\star \geq 0 \Rightarrow p_i \geq p^\star_i > 0 \text{ for all } i \in \mathcal{V}^\star$$

which completes the proof. □

In other words, nodes in the maximum variance support $\mathcal{V}^\star$ have positive resistance curvature both in $G^\star$ (by definition) but also in the graph G. Conversely, the maximum variance support $\mathcal{V}^\star$ must be a subset of the nodes with positive resistance curvature in G. This observation leads to the following result:

Theorem 5.23. *Let G be a graph and repeatedly take the Kron reduction with respect to all nodes with positive resistance curvature $p_i > 0$ until all nodes have positive resistance curvature. The remaining graph is $G^\star = G/\mathcal{V}^{\star c}$.*

[14] This is a slight abuse of notation, since $\mathbf{p}^\star$ may also be the $|\mathcal{V}^\star| \times 1$ resistance curvature vector of the graph $G^\star$.

Proof. We first note that for all sets $\mathcal{W}$ with $\mathcal{V}^\star \subseteq \mathcal{W} \subseteq \mathcal{N}$, the maximum variance support $\mathcal{V}^\star$ of G is also the maximum variance support of the Kron reduction $G/\mathcal{W}^c$. This holds because the optimization domain of the maximum-variance problem for $G/\mathcal{W}^c$ is a subset of the optimization domain of the maximization problem for G, with the optimum contained in the subset. This implies that the nodes in $\mathcal{V}^\star$ have positive resistance curvature in the Kron reduction $G/\mathcal{W}^c$ for all $\mathcal{V}^\star \subseteq \mathcal{W} \subseteq \mathcal{N}$.

Now consider the sequence of graphs $G_0, G_1, \ldots$ defined by $G_0 := G$ and $G_{k+1} = G_k/\mathcal{V}_k$ for all $k \geq 0$, where $\mathcal{V}_k$ is the set of nodes with nonpositive resistance curvature $p_i \leq 0$ in G_k. By Lemma 5.22 and the discussion above, we know that $\mathcal{V}_k \cap \mathcal{V}^\star = \emptyset$ and thus that the nodes in $\mathcal{V}^\star$ are contained in every graph in the sequence. For every step $k \geq 0$ in the sequence, either the set $\mathcal{V}_k$ of nonpositively curved nodes is nonempty and the size of the graph decreases or $\mathcal{V}_k$ is empty and the sequence terminates. Let $G' = G_\ell$ be the graph where such an empty set occurs (i.e. with $\mathbf{p}' > 0$) and suppose for contradiction that $\mathcal{V}^\star \neq \mathcal{N}_\ell$. Then we know that $\mathcal{V}^\star \subset \mathcal{N}_\ell$ (this is because $\mathcal{V}^\star \subseteq \mathcal{N}_k$ is true for every graph in the sequence and equality is excluded) and that the resistance curvature $\mathbf{p}'$ satisfies (i) $p'_i > 0 = p^\star_i$ for all $i \in \mathcal{V}^{\star c}$ and (ii) $p'_i \geq p^\star_i$ for all $i \in \mathcal{V}^\star$ by Lemma 5.22 and by properties of the sequence. Together, we thus have that $\mathbf{p}' \geq \mathbf{p}^\star$ and in particular that $\mathbf{p}' - \mathbf{p}^\star$ is a nonnegative nonzero vector. But this produces the contradiction

$$0 \leq (\mathbf{p}' - \mathbf{p}^\star)^T \Omega (\mathbf{p}' - \mathbf{p}^\star) < 0 \text{ if } \mathbf{p}' \neq \mathbf{p}^\star$$

where the lower bound follows because the quadratic product of a nonnegative vector and a nonnegative matrix is nonnegative, and the upper bound follows because $(\mathbf{p}' - \mathbf{p}^\star) \in \text{span}(\mathbf{u})^\perp$. This contradiction implies that $\mathcal{N}_\ell = \mathcal{V}^\star$ and thus that the graph sequence strictly decreases in size every step and terminates if and only if the graph is equal to $G^\star$ as required. □

This characterization suggests a simple greedy algorithm to calculate an exact solution to the maximum-variance problem with respect to the resistance distance: (step 1) identify any node x with nonpositive resistance curvature, (step 2) if such a node exists, take the $\{x\}^c$-Kron reduction and return to step 1 or if no such node exists then the resulting graph is $G^\star$. Each step in this algorithm is local (once Q and Ω are known) and can be constant time if the largest degree is constant in n. Furthermore, the procedure is more transparent than using general convex program solvers and we found that it runs faster in practice.

As a consequence of Theorem 5.23, the maximum-variance problem is trivial for graphs with nonnegative resistance curvature.

Corollary 5.24. *If a graph has nonnegative resistance curvature* $\mathbf{p} \geq 0$*, then the maximum-variance problem with respect to the effective resistance is solved by* $\mathbf{f}^\star = \mathbf{p}$ *and* $\mathrm{var}_\omega^\star = \sigma^2$*.*

Proof. Let G be a graph with nonnegative resistance curvature $\mathbf{p} \geq 0$ with support $\mathcal{V}$ — i.e. nodes in $\mathcal{V}^c$ have zero resistance curvature. Applying Theorem 5.23, the first two graphs in the sequence are $G_1 = G$ and $G_2 = G/\mathcal{V}^c =: G'$. Following Proposition 3.25 on the resistance curvature for Kron reduction graphs and the fact that $p_i = 0$ for all $i \in \mathcal{V}^c$, we find that $\mathbf{p}'_i = \mathbf{p}_i > 0$ for all $i \in \mathcal{V}$ and thus that the sequence terminates at $G_2 = G^\star$. This proves that $\mathbf{f}^\star = \mathbf{p}' = \mathbf{p}$. Similarly, by Proposition 3.25, the resistance radius satisfies $\sigma'^2 = \sigma^2$ since $p_i = 0$ for all $i \in \mathcal{V}^c$, and thus $\mathrm{var}_\omega^\star = \sigma'^2 = \sigma^2$, as required. □

We note in particular that Corollary 5.24 in combination with Proposition 5.17 implies that $\frac{1}{4}\operatorname{diam}(\omega) \leq \sigma^2 \leq \frac{n-1}{2n}\operatorname{diam}(\omega)$ for graphs with $\mathbf{p} \geq 0$. In other words, the resistance diameter and twice the resistance radius are equal, up to a factor of at most two.

As a direction for future work we believe it would be interesting to study the set function

$$\eta(\mathcal{V}) := \frac{1}{2}\left(\mathbf{u}^T[D_{\mathcal{V}\mathcal{V}}]^\dagger \mathbf{u}\right)^{-1} \text{ for all } \mathcal{V} \subseteq \mathcal{N},$$

for some distance matrix D (or general matrix A) whenever this is well-defined. This definition is inspired by the resistance radius set function — which has many interesting properties such as an inclusion–exclusion property for cut sets in Proposition 3.40 and submodularity (see Eq. (6.29)) for nonnegatively curved graphs in Theorem 6.33 — and the maximum variance formulation in terms of distance submatrices (Theorem 5.15).

5.4 The geometry of variance

5.4.1 The maximum variance support

The results in the previous section are valid for any semimetric space, without explicit reference to an underlying graph. In the case of graph distances, which reflect the structure of the graph, the maximum-variance problem and its solutions can furthermore be interpreted in terms of the graph structure and geometry. More precisely, we propose that maximum variance distributions and their supports $\mathcal{V}_d^\star$ determine a notion of *graph boundary*. We support this intuition by showing that the maximum

variance support is contained in an existing notion of boundary, and then by a number of examples.

Chartrand et al. defined a notion of graph boundary in [48], which has been widely studied. Steinerberger considered a relaxation of this definition and proposed the following graph boundary [225]:

$$\partial_d G := \Big\{ i \in \mathcal{N} : \exists x \in \mathcal{N} \text{ s.t. } d(i,x) \geq \sum_{j \sim i} \tfrac{1}{d_i} d(j,x) \Big\}. \tag{5.11}$$

For the shortest-path distance, Steinerberger showed that the boundary $\partial_{d_{sp}} G$ satisfies a number of natural geometric properties, such as isoperimetric inequalities. We note that a node function f is called *superharmonic* at node i if $f(i) \geq \sum_{j:j\sim i} f(j)/d_i$ (see [164, Ex. 4.4]) and that the graph boundary thus contains nodes i for which there exists a node x such that the distance to x, i.e. $f_x : i \mapsto d(i,x)$, is superharmonic at i. We find the following relation between maximum variance supports and the graph boundary:

Proposition 5.25. *The support of a maximum variance distribution is a subset of Steinerberger's graph boundary for any distance d on a graph; i.e.*

$$\mathcal{V}_d^\star \subseteq \partial_d G$$

Proof. Let G be a graph with a maximum variance distribution $\mathbf{f}^\star$. Following the necessary conditions in Proposition 5.12, we know that $(\mathbf{e}_i - \mathbf{e}_j)^T D \mathbf{f}^\star \geq 0$ for all $i \in \mathcal{V}_d^\star$ and all j. Let $i \in \mathcal{V}_d^\star$ be some node in the maximum variance support; summing over its neighbours $j \sim i$, we find

$$\sum_{k \in \mathcal{V}_d^\star} f_k^\star \sum_{j:j\sim i} (d(i,k) - d(j,k)) \geq 0.$$

Since the maximum variance probabilities $f_k^\star$ are positive, there must be at least one $x \in \mathcal{V}_d^\star$ for which $\sum_{j:j\sim i}(d(i,x) - d(j,x)) \geq 0$, which means that $i \in \partial_d G$. As i was an arbitrary node in the maximum variance support. This completes the proof. □

Next, we discuss a number of examples in which the maximum variance support $\mathcal{V}_d^\star$ retrieves an intuitive set of boundary nodes of a graph.

Example 5.26 (**cut nodes & tree graphs**)**.** Chartrand et al. [48] found the intuitive result that their boundary (a subset of Steinerberger's boundary) cannot contain cut nodes. This result is reproduced by the maximum variance support with respect to the effective resistance:

Proposition 5.27. *The maximum variance support $\mathcal{V}_\omega^\star$ does not contain any cut nodes.*

Proof. Let G be a graph with cut node x. By Property 6.2, the resistance curvature of the cut node x is nonpositive in G, i.e. $p_x \leq 0$. By Lemma 5.22, we furthermore know that all nodes in $\mathcal{V}_\omega^\star$ have nonnegative resistance curvature in G, as $p_i \geq p_i^\star > 0$, which excludes the cut node x from the maximum variance support, as required. □

As a corollary, the maximum variance support of a tree graph must be a subset of its leaf nodes, which is the intuitive boundary of a tree. This result for tree graphs was obtained in [66, 73, 127, 225] for related problems and using a variety of approaches.

Example 5.28 (**node transitive graphs**)**.** Proposition 6.11 in Chapter 6 says that the resistance curvature in node transitive graphs is constant, positive and given by $\mathbf{p} = \mathbf{u}/n$. By Corollary 5.24, this implies that $G^\star = G$ and thus that the maximum variance support is equal to the full node set $\mathcal{V}_\omega^\star = \mathcal{N}$. This result is expected since nodes are indistinguishable in a node transitive graph and thus no set of nodes should be distinguished from the others as a boundary.

Example 5.29 (**size of the maximum variance support**)**.** As discussed in the introduction to this section, we not only expect the maximum variance distribution to be supported on nodes which are far apart, but also to be spread out over many nodes. The following result gives some further intuition:

Proposition 5.30. *The maximum variance support contains at least two nodes and exactly two nodes if and only if these nodes are the ends of a path graph with resistance (equivalently, shortest-path) metric.*

Proof. A distribution supported on a single node has zero variance, which is in contradiction with the positive lower bound for the maximum variance in Proposition 5.17. As a result, any maximum variance distribution must be supported on at least two nodes. We now consider when the case of two nodes occurs.

(forward direction) Proposition 5.17 states that the maximum variance of a path graph with resistance (or shortest-path) distance d is $\mathrm{var}_d^\star = \frac{1}{4}\operatorname{diam}(d)$, which is achieved by the distribution $\mathbf{f}^\star = \frac{1}{2}(\mathbf{e}_i + \mathbf{e}_j)$ supported on the two end nodes, as required.

(converse direction) Let G be a graph in which a maximum variance distribution $\mathbf{f}^\star$ is supported on two nodes i, j. Then we find

$$\mathrm{var}(\mathbf{f}^\star) = f_i^\star f_j^\star d(i,j) \leq f_i^\star(1 - f_i^\star)\operatorname{diam}(i,j) \leq \tfrac{1}{4}\operatorname{diam}(d)$$

with equality if and only if $d(i,j) = \operatorname{diam}(d)$ (the first inequality) and $f_i^\star = f_j^\star = 1/2$ (the second inequality). By Proposition 5.17, we also know that $\operatorname{var}(\mathbf{f}^\star) \geq \frac{1}{4}\operatorname{diam}(d)$, which means that equality must hold throughout, and by the equality condition in 5.17, we then know that this implies that d is the resistance metric of a path graph, which completes the proof. □

Example 5.31 (**Random geometric graphs**)**.** Random geometric graphs (RGGs) are random graphs with nodes sampled from some metric space $\mathcal{N} \subseteq M$ and with nearby nodes connected $i \sim j$ if $d_M(i,j) \leq r$ for some connection radius r. If the sampled domain has a clear boundary in the ambient space M, then we would hope that the nodes close to this boundary play the role of a boundary in the graph. Figure 5.7 shows that this is indeed the case for some RGGs constructed on a domain in Euclidean space and considering the boundary with respect to the effective resistance; Section 6.3.3 discusses RGGs in some more detail.

Importantly, as observed in the letter 'A' in Figure 5.7, the boundary $\mathcal{V}_\omega^\star$ does not seem to contain nodes near points on the boundary with negative curvature (in this case, locally concave). The precise interplay between boundary curvature and the set $\mathcal{V}_\omega^\star$ is unclear, but in Chapter 6 we discuss the interpretation of $\mathbf{p}$ as a notion of discrete curvature; in fact, this observation was the initial clue that motivated our investigation of $\mathbf{p}$ in the context of discrete curvature.

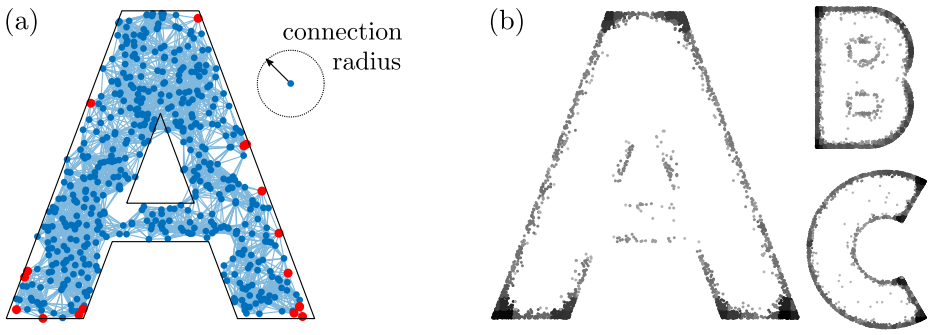


Figure 5.7: The maximum variance support $\mathcal{V}_\omega^\star$ of RGGs in a bounded domain of $\mathbb{R}^2$ is more likely to contain nodes that are located near the boundaries. Panel (a) shows one realization of an RGG on $n = 500$ nodes in a domain shaped as the letter 'A', with connection radius as indicated. The nodes in the maximum variance support are colored red, and are all located near the boundary of the domain. Panel (b) summarizes the maximum variance support node locations calculated in 250 independent RGG realizations on the given domains. A darker (respectively lighter) shading indicates a higher (respectively lower) density of maximum variance nodes. This figure supports that the maximum variance support nodes are more likely to be located near the domain boundaries and, in particular, near corners with high curvature.

5.4.2 Variance and the simplex

We briefly discuss how the variance measured with respect to $\sqrt{\omega}$ can be interpreted geometrically in terms of the resistive embedding φ and the associated simplex geometry. Recall that the resistive embedding determines a bijection between distributions on a graph (the probability simplex Δ) and the hyperacute simplex S associated with the graph, and that the circumcenter and circumradius are given by $\varphi(\mathbf{p})$ and σ.

Variance with respect to the effective resistance can be written in terms of distances related to the simplex as follows:

Proposition 5.32. *Let G be a weighted graph with resistive embedding φ. The variance of a distribution $\mathbf{f}$ with respect to the effective resistance is determined by the distance from its resistive embedding $\varphi(\mathbf{f})$ to the circumcenter $\varphi(\mathbf{p})$ of the simplex:*

$$\operatorname{var}_\omega(\mathbf{f}) = \sigma^2 - \|\varphi(\mathbf{f}) - \varphi(\mathbf{p})\|^2 \text{ for any resistive embedding } \varphi.$$

Proof. Expressing the squared distance between points on the simplex in terms of the resistance matrix and invoking the equilibrium identity $\Omega\mathbf{p} = 2\sigma^2\mathbf{u}$, we find

$$\|\varphi(\mathbf{f}) - \varphi(\mathbf{p})\|^2 = -\frac{1}{2}(\mathbf{f}-\mathbf{p})^T\Omega(\mathbf{f}-\mathbf{p}) = -\frac{1}{2}\mathbf{f}^T\Omega\mathbf{f} + \mathbf{p}^T\Omega\mathbf{f} - \frac{1}{2}\mathbf{p}^T\Omega\mathbf{p} = \sigma^2 - \operatorname{var}_\omega(\mathbf{f}),$$

as required. □

Following Proposition 5.32, the variance of a distribution becomes larger as its resistive embedding lies closer to the circumcenter $\varphi(\mathbf{p})$. This suggests a geometric reformulation of the maximum-variance problem:

$$\sigma_\star^2 = \sigma^2 - \min_{\mathbf{x}\in S} \|\mathbf{x} - \varphi(\mathbf{p})\|^2 = \sigma^2 - d^2(\varphi(\mathbf{p}), S)$$

with optimal solution $\mathbf{p}^\star$, where the distance d between a point and a set in $\mathbb{R}^{n-1}$ is defined as the smallest distance to any point in the set. If the graph has nonnegative resistance curvature $\mathbf{p} \geq 0$, the circumcenter is a point in the simplex and the maximum variance is achieved for $\mathbf{p}^\star = \mathbf{p}$. Otherwise, the maximum variance distribution is the projection of the circumcenter $\varphi(\mathbf{p})$ onto S, which is the closest point on the simplex to $\varphi(\mathbf{p})$. Following Theorem 5.23, this projection can thus be computed by a repeated Kron reduction of the associated graph until $\mathbf{p}' \geq 0$.

Chapter 6

Discrete curvature from effective resistances

This chapter deals with the interpretation of the resistance curvature $\mathbf{p}$ as a notion of discrete curvature. The contribution of this new perspective is twofold: first, we add to the conceptual understanding of the vector $\mathbf{p}$, with the new results in Section 6.4 on discrete Ricci flow and Section 6.5 on positively curved graphs as important consequences. Second, we add one more approach to the active and broad discrete curvature research landscape, with both clear connections to existing approaches as well as apparent unique features. While the contribution to the theory of $\mathbf{p}$ and the effective resistance more generally is already clear, the precise implications of this new approach for the study of discrete curvatures remains to be seen.

Section 6.2 introduces a definition for the resistance curvature in terms of relative resistances. This leads to bounds on $\mathbf{p}$ and examples of graphs with positive, negative and zero resistance curvature. We introduce a related notion of link resistance curvature κ (Definition 6.4) and derive similar bounds. Finally, we summarize the many equivalent definitions for $\mathbf{p}$ and σ^2 encountered in this thesis in Section 6.2.3, including new ones such as Proposition 6.13.

Next, Section 6.3 details our arguments for an interpretation of $\mathbf{p}$ as a notion of discrete curvature. We consider graphs with constant resistance curvature of the correct sign, we describe the relation to other notions of discrete curvature and present numerical and theoretical evidence for convergence to zero curvature in Euclidean RGGs (expression (6.18)).

Section 6.4 introduces a discrete Ricci flow of graphs $G(t)$ based on the resistance curvature. We show that this flow is a gradient flow and is described as a quadratic flow of Laplacian matrices. To develop a further understanding, we define and study

a simplified local flow (6.27). This flow is well-defined and can be interpreted as a continuous interpolation between a graph and an elementary Kron reduction (Proposition 6.21) with simple solutions for the effective resistances (Property 6.23) and resistance curvatures (Property 6.24) as a result.

Finally, Section 6.5 considers the class of graphs with nonnegative curvature $\mathbf{p} \geq 0$. We describe some examples and show that these graphs are closed under Kron reduction. Most importantly, we find the following properties: positively curved graphs are strongly connected in the technical sense of 'toughness' (Theorem 6.31) and as a second result, we show that the resistance radius σ^2 is a 'submodular* set function' in nonnegatively curved graphs (Theorem 6.33).

Most of this chapter appeared in [72], with the exception of Sections 6.4.2 and 6.5, and references to the relevant literature are given in the introduction section below.

6.1 Introduction

6.1.0.1 From continuous ...

The idea of curvature has a long and rich history in geometry and provides the mathematical language for one of the most important modern-day physical theories, Einstein's theory of general relativity, which describes the interplay between mass and energy and the (Ricci) curvature of spacetime. As illustrated in Figure 6.1, curvature is a geometric property of smooth spaces such as lines and surfaces, and quantifies how much and in which ways a smooth space differs from being flat around a point.

Several complementary and related notions of curvature exist. *Scalar curvature*, for instance, associates a single number to each point on a manifold and allows one to distinguish the qualitatively different cases of positive curvature (like the surface of a sphere), zero curvature (like the flat plane) and negative curvature (like a saddle point), as illustrated in Figure 6.1. The scalar curvature is a measure of how much the volume of (small) ϵ-balls around a point differ from the volume of ϵ-balls in Euclidean space of the same dimension. *Ricci curvature* associates a tensor to each point on a manifold, and reflects the difference of volume growth between geodesics emanating from the point in two tangential directions compared to Euclidean growth. While the Ricci curvature includes 'directional' information not present in the scalar curvature, the latter can be retrieved as the trace of the former. We refer the readers to [20, 135] for the background on curvature and differential geometry.

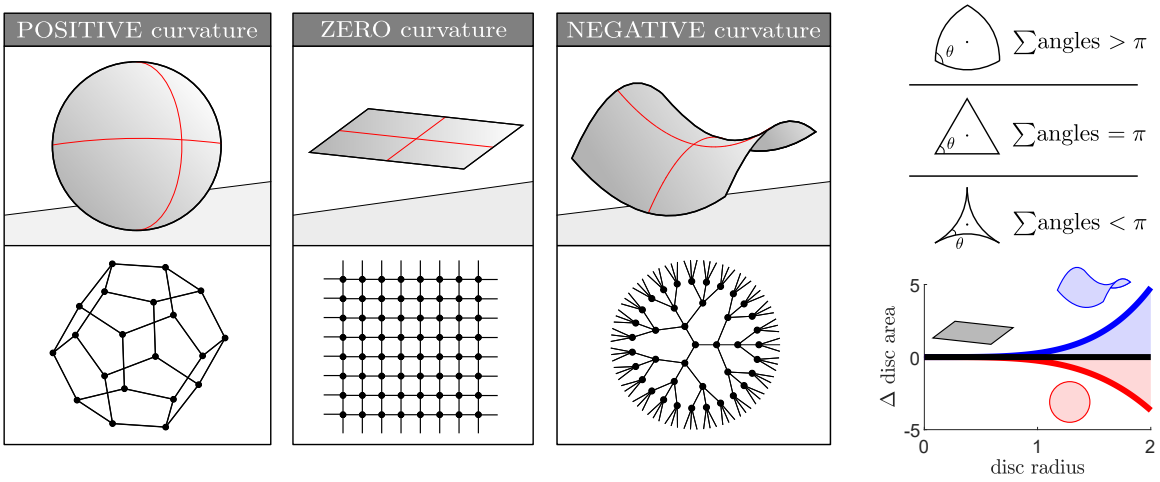


Figure 6.1: Scalar curvature distinguishes points with different local geometries. On smooth surfaces, such as the images shown on the left, points with positive/zero/negative curvature are locally like the surface of a sphere/flat plane/saddle point. The figures on the right illustrate how the curvature sign influences the local geometry around a point, in terms of triangles around the point (with different angle sums) and in terms of volume growth (where geodesic discs can grow faster/slower than in Euclidean space). The second row on the left shows some graphs, which are naturally associated to spaces of constant curvature; see Section 6.3.1.

6.1.0.2 ... to discrete

Discrete settings such as graphs lack the differential structure required for classical definitions of curvature. However, it is still widely acknowledged that these spaces exhibit relevant geometric features and many notions of *discrete curvature* have been proposed that work in the discrete setting [9, 30, 56, 87, 99, 115, 189, 205, 208]; for a survey, see for instance [26, 138, 183, 199]. In Section 6.3.2, we briefly discuss a number of these curvatures. Broadly speaking, the vision is that discrete curvature can mimic the successful classical differential theory and play a role in representing and understanding 'emergent' geometric properties in discrete structures — which at the local level appear purely combinatorial, but at a larger scale, feature much richer properties. Importantly, the subject of discrete curvature is thus not just about finding generalized notions of curvature, but also about trying to translate the multitude of curvature-related results (Bonnet–Myers, Lichnerowicz, etc.) to the discrete setting; this has been done successfully for instance in the case of Forman–Ricci curvature [99], Ollivier–Ricci curvature [161, 162, 181, 189], Bakry–Émery curvature [16], entropic curvature [87] and Steinerberger's curvature [226]. This aspect is less pronounced in this thesis (except perhaps for Section 6.5), but the logical next step

is to try and translate the classical curvature results to our proposed curvatures.

The idea of discrete curvature is perhaps most easily understood when the discrete space under consideration is closely related to some smooth space. For instance, a graph which can be isometrically embedded into a manifold (see Section 6.3.1), or random geometric graphs, which are natural 'coarse' representations of some smooth metric space (see Section 6.3.3). However, most notions of discrete curvature still apply even when no smooth ambient space is involved. In other words they are defined intrinsically.

Discrete curvature is not only studied out of mathematical curiosity, but also as a tool in applications, for instance in the context of various digital technologies. As noted in the preface of Najman & Romon's review on modern approaches to discrete curvature [183, p. v]: "*[discrete geometry] has nevertheless taken a new twist at the end of the twentieth century, first with the onset of computer science and engineering, requiring a science based on discrete objects, bits and integers, molecules or bricks, graphs, and networks.*". This includes applications that rely on finite approximations of continuous or smooth structures, as in computer graphics and image processing [133, 183]. More recently, notions of discrete curvature have also appeared in the design and analysis of algorithms [76, 132, 217, 238] and applications in network science, including community detection [109, 188, 216], methods for graph comparison [186] and alternative ways to capture connectivity in interconnected systems [134, 182, 187, 256]. Discrete curvature also plays a role in some approaches to quantum gravity, which attempt to unify the continuous theory of gravity with the discrete theory of quantum mechanics; see for instance[1] [108, 241].

6.2 Resistance curvature

We consider two types of discrete curvature: a curvature defined on the nodes of a graph as $p : \mathcal{N} \to \mathbb{R}$ and a curvature defined on the links of a graph as $\kappa : \mathcal{L} \to \mathbb{R}$. As introduced in Chapter 3, we refer to the nodal curvature as 'resistance curvature' and will use the slightly longer 'link resistance curvature' for the latter[2]. The relation of p and κ to existing notions of curvature suggests that p might be a *scalar curvature* and κ a *Ricci curvature*. However, these two types of curvature satisfy specific relations

[1] There are many different approaches to quantum gravity with discreteness arising in a variety of ways; in particular, discrete curvature does not always play as important of a role as it does in [241] or [108].

[2] The name "effective resistance curvature" was introduced in [114] for a different notion of curvature derived from effective resistances.

in the continuous case which are not reproduced by p and κ; for instance, scalar curvature is the trace of the Ricci curvature. This suggests that p or κ might instead be approximate notions of scalar and Ricci curvature.

6.2.1 Node resistance curvature

Chapter 3 introduced the resistance curvature $\mathbf{p}$ in context of the Fiedler–Bapat identity and we already encountered a number of alternative definitions throughout the thesis; Section 6.2.3 provides a summary of these alternative definitions. The discussion on the curvature properties of $\mathbf{p}$ will focus on the following definition in terms of relative resistances:

Definition 6.1 (resistance curvature)**.** The *(node) resistance curvature* is defined as

$$p_i := 1 - \frac{1}{2} \sum_{j:j\sim i} c_{ij}\omega_{ij} \text{ for any node } i \in \mathcal{N}. \tag{6.1}$$

This expression follows from the diagonal of the $Q\Omega$ identity as $\sum_{j:j\sim i} c_{ij}\omega_{ij} = -(Q\Omega)_{ii} = 2 - 2p_i$. While Definition 6.1 for the resistance curvature appeared before (under a different name) in [10, 70, 96, 228, 269], there has not been a systematic study of its properties, and its relation to discrete curvature in particular seems to be new. Initially, this curvature interpretation is suggested by the following qualitative observation: if the neighbourhood of a node is "tree-like" with few short cycles, then the local relative resistances will be large and thus p_i will be small, as expected for a notion of curvature [136, Thm. 1]. If the neighbourhood is "clique-like" with many short cycles, then the local relative resistances will be small with a large p_i as a result.

Expression (6.1) defines the resistance curvature based on the *relative resistance* of incident links on a node. We recall Theorem 2.13, which states that the relative resistance $c_{ij}\omega_{ij}$ of a link is the probability that this link is contained in a random spanning tree. This result bounds the relative resistance in $(0, 1]$ and leads to Foster's Theorem and the cut bounds (2.9). With these results, we find the following bounds on the node resistance curvature:

Property 6.2. *The resistance curvature of a node i is bounded as*

$$1 - \frac{d_i}{2} \leq p_i \leq 1 - \frac{\beta(G_i\backslash\{i\})}{2}$$

where $G_i\backslash\{i\}$ is the component G_i with node i removed. Equality is achieved if and only if all links connected to i are cut links, in which case both bounds are equal.

Proof. *(lower bound)* The lower bound follows from the relative resistance bound $c_{ij}\omega_{ij} \leq 1$ for each term in expression (6.1), with equality if and only if each link is a cut link. *(upper bound)* The upper bound follows by summing the relative resistance over cuts: if removing i disconnects G_i into $\beta' := \beta(G_i \backslash \{i\})$ components, this means that the links incident to node i can be partitioned into β' sets of links $\{C_k\}_{k=1}^{\beta'}$ that are cuts of G_i. By the cut bound (2.9), we then find

$$p_i = 1 - \frac{1}{2} \sum_{k=1}^{\beta'} \sum_{(i,j) \in C_k} c_{ij}\omega_{ij} \leq 1 - \frac{\beta(G_i \backslash \{i\})}{2}.$$

Furthermore, equality occurs in the cut bound if and only if the cut consists of a single link, which means that equality in the upper bound for p_i occurs if and only if all incident links are cut links. This completes the proof. □

We note that the bounds correctly assign $p_i = 1$ for a disconnected node. In tree graphs, every link is a cut link and thus Property 6.2 implies that the non-leaf nodes in a tree have nonpositive curvature while leaf nodes have resistance curvature equal to $1/2$. For path graphs, the two end nodes have resistance curvature $1/2$ while all other nodes have zero curvature. The bounds for $\mathbf{p}$ thus produce some examples where negative and zero curvature happen generically (independent of link weights), in tree and path graphs.

A second property of the node resistance curvature follows from Foster's Theorem:

Property 6.3. *The sum of node resistance curvatures in a graph is equal to the number of connected components:* $\sum_{i \in \mathcal{N}} p_i = \beta(G)$.

Proof. Let G be a connected graph. By Foster's Theorem, the sum over all relative resistances in the graph is equal to $n - 1$ and thus

$$\sum_{i \in \mathcal{N}} p_i = n - \sum_{j \sim i} c_{ij}\omega_{ij} = n - (n-1) = 1.$$

For a graph with more than one connected component, this result can be applied to each component separately to yield Property 6.3. We note that $\mathbf{u}^T\mathbf{p} = 1$ for a connected graph also follows immediately from the Fiedler–Bapat identity, as described in Corollary 3.7. □

In geometric terms, Property 6.3 can be thought of as a Gauss–Bonnet-like result, which relates the sum over all curvatures to topological properties of the graph. Furthermore, Property 6.3 suggests the cycle graph as an example of positive curvature: by virtue of the rotational symmetry of the cycle graph, all nodes are indistinguishable and thus should have the same resistance curvature. By $\sum p_i = 1$, this then

implies that the resistance curvature is positive and equal to $p_i = 1/n$ for all nodes in the n-cycle graph; this derivation is confirmed in Example 6.10. The cycle graph is thus a first example of a graph with *positive resistance curvature*. These examples are further discussed (and illustrated) at the end of the next section.

6.2.2 Link resistance curvature

The link resistance curvature is defined as follows:

Definition 6.4 (link resistance curvature)**.** The *link resistance curvature* is defined as

$$\kappa_{ij} := \frac{2(p_i + p_j)}{\omega_{ij}} \text{ for all } (i,j) \in \mathcal{L}. \tag{6.2}$$

The link resistance curvature of a link is equal to the sum of the (node) resistance curvature of its end nodes, divided by the effective resistance between them. Consequently, calculating the link resistance curvature is straightforward once p is known, and when p_i and p_j have the same sign, this immediately determines the sign of κ_{ij} as well. Tree, path and cycle graphs can thus serve as first examples of graphs with links of negative/zero/positive link curvature.

Similar to the node resistance curvature bounds, we find:

Property 6.5. *The link resistance curvature is bounded as*

$$\tfrac{1}{\omega_{ij}}(4 - d_i - d_j) \leq \kappa_{ij} \leq \tfrac{1}{\omega_{ij}}\left[6 - 2\beta(G_i \backslash (i,j)) - \beta(G_i \backslash \{i,j\})\right],$$

where $G_i\backslash(i,j)$ is the graph with link (i,j) removed and $G_i\backslash\{i,j\}$ the graph with nodes i and j removed. Equality is achieved if and only if all links incident to i and j are cut links, in which case both bounds are equal.

Proof. *(lower bound)* Both the lower bound and the conditions for equality follow directly from the lower bound $p_i \geq 1 - d_i/2$ on the node resistance curvature in Property 6.2. *(upper bound)* For simplicity we further assume that $G = G_i$ is a connected graph; this is without loss of generality since any link is contained in a connected component. We can write the link resistance curvature as

$$\frac{2(p_i + p_j)}{\omega_{ij}} = \frac{1}{\omega_{ij}}\left(4 - \sum_{\ell \in \mathcal{L}_{ij}} c_\ell \omega_\ell - 2c_{ij}\omega_{ij}\right), \tag{6.3}$$

where $\mathcal{L}_{ij} := \{(x,k) \in \mathcal{L}\backslash(i,j) \mid x \in \{i,j\}\}$ are the links incident on (i,j). We note that removing nodes i and j from G is equal to removing the links $\mathcal{L}_{ij}$ from G

and then removing the nodes i, j from the resulting graph. Hence, every connected component in $G\backslash\{i,j\}$ is disconnected from G by some cut $C_k \subseteq \mathcal{L}_{ij}$, which gives a partition of $\mathcal{L}_{ij}$ into cuts. For each cut, we may then invoke the cut bound (2.9) and find that

$$\sum_{\ell\in\mathcal{L}_{ij}} c_\ell\omega_\ell = \sum_{k=1}^{\beta(G\backslash\{i,j\})} \sum_{\ell\in C_k} c_\ell\omega_\ell \geq \beta(G\backslash\{i,j\}),$$

with equality if and only if all cuts consist of single links and thus if all links in $\mathcal{L}_{ij}$ (and thus also (i,j)) are cut links. Second, we know that the relative resistance satisfies $c_{ij}\omega_{ij} > 0$ and that $c_{ij}\omega_{ij} = 1$ if and only if (i,j) is a cut link. Hence, in general, we have $c_{ij}\omega_{ij} \geq [\beta(G\backslash(i,j)) - 1]$, where $G\backslash(i,j)$ has the link (i,j) removed; again, equality only holds if this is a cut link. Introducing the bounds for these link terms into (6.3) yields the proposed upper bound and conditions for equality, and thus completes the proof. □

A simpler but less tight version of the upper bound is $\kappa_{ij} \leq 2/\omega_{ij}$ which follows from the upper bound on p_i with $\beta(G_i\backslash\{i\}) \geq 1$ when i is not a isolated node.

Following the discussion in Section 2.3.2, the resistance curvatures can be approximated efficiently using Laplacian solvers and related tools. In particular, we note that the work in [194] suggests that the resistance curvatures can be approximated using local methods (i.e. only depending on the structure of the graph within a certain number of steps from the node or link) and the methods in [3] to efficiently approximate the pseudoinverse Laplacian diagonal $\boldsymbol{\zeta}$ can be used to calculate the node resistance curvature, based on expression $\mathbf{p} = \frac{1}{2}Q\boldsymbol{\zeta} + \mathbf{u}/n$.

We now discuss a number of examples to support the introduced definitions and to illustrate how negative, zero and positive curvature occur generically in certain classes of graphs.

Example 6.6 (**tree graphs**)**.** Since every link in a tree graph is a cut link, the resistance curvature bounds in Properties 6.2 and 6.5 hold with equality, and we find

$$c_{ij}\omega_{ij} = 1, \quad p_i = 1 - \frac{d_i}{2} \quad \text{and} \quad \kappa_{ij} = c_{ij}(4 - d_i - d_j)$$

for all nodes i and links (i,j). In other words, except for the leaf nodes and the links connected to them, all nodes and links have nonpositive node/link resistance curvature.

Example 6.7 (**path graphs**)**.** We recall that a path is a tree graph with two end nodes of combinatorial degree $d = 1$, and $d = 2$ for all other nodes. Following the result for tree graphs, the resistance curvatures in a path with $n > 2$ are given by

$$\begin{cases} p_i = \frac{1}{2} \text{ at the end nodes} \\ p_i = 0 \text{ otherwise} \end{cases} \qquad \begin{cases} \kappa_{ij} = c_{ij} \text{ if } i \text{ or } j \text{ is an end node} \\ \kappa_{ij} = 0 \text{ otherwise.} \end{cases}$$

In other words, except for the end nodes and the links connected to them, all nodes and links have zero curvature. The case of a 2-node path K_2 is different, as we get $\kappa_{ij} = 2c_{ij}$.

Example 6.8 (**cycle graphs**)**.** As explained in Example 6.10 below, the symmetries of a graph have implications for the resistance curvature. One example of a node and link transitive graph is the cycle graph, for which the node and link resistance curvatures are all equal, and thus

$$c\omega_{ij} = \frac{n-1}{n}, \quad p_i = \frac{1}{n} \quad \text{and} \quad \kappa_{ij} = \frac{4c}{n-1}$$

for all nodes i and links (i, j) and constant link weight c. In other words, all nodes and links are positively curved. Some more examples of positively curved graphs are discussed in Example 6.10 and Section 6.3.1. Figure 6.2 shows examples of tree, path and cycle graphs and their curvature signs.

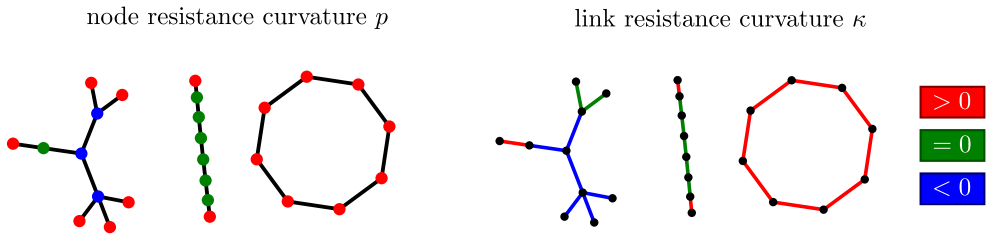


Figure 6.2: The node and link resistance curvatures in tree, path and cycle graphs.

Example 6.9 (**Erdős–Rényi random graph**)**.** An *Erdős–Rényi* (ER) graph $G(n, \rho)$ is a random graph where every pair of nodes is connected with probability ρ, see [29, 245]. Figure 6.3 shows how the mean link curvature $\frac{1}{m}\sum_{i\sim j}\kappa_{ij}$ and resistance curvature distributions evolve in these random graphs as a function of the connection probability (density) ρ. An intuitive explanation for the observed evolution of the mean link curvature is as follows: ER graphs with very small ρ typically consist of a collection of 2-paths and disconnected nodes (graph A in Figure 6.3) with a mean link curvature of $+2$ since $\kappa_{ij} = 2$ for a 2-path. As ρ increases, the typical ER graph

will transition from 2-paths to a collection of trees or locally tree-like components (graphs B, C) with decreasing κ as the density increases. As ρ increases further, the typical ER graph will be a dense, connected graph (graph D) in which κ increases along with the density, until it reaches mean link curvature $+2$ for the complete graph at $\rho = 1$. For the node resistance curvature, we find a similar explanation: small and medium-density ER graphs consist mainly of disconnected nodes and 2-paths with $p = 1$ and $p = 1/2$ respectively, or a collection of trees with $p_i = 1 - d_i/2$, which results in the 'discrete distribution' of p on half integers, as observed in Figure 6.3. For large density ρ, we observe that the locally tree-like graphs have a "continuous distribution" of possible node curvatures. This distribution narrows and converges to $p \approx 1/n$ when the density increases to 1; this can be explained by the high uniformity of dense ER graphs in combination with Foster's Theorem for $\beta = 1$. The description above suggests a possible explanation for the observed mean link resistance curvature in Figure 6.3. The Erdős–Rényi model is very well-studied and so we may hope that a detailed and precise explanation can be provided.

Example 6.10 (**symmetry and curvature**)**.** Certain graph symmetries have strong implications for the resistance curvatures in a graph. For resistance curvatures, the relevant notion is node (link) *transitivity*, which says that every node (link) can be mapped onto any other node (link) by some structure-preserving map. More precisely, a permutation $\pi : \mathcal{N} \to \mathcal{N}$ of the nodes of a graph is called an *automorphism* if it preserves the graph structure, i.e. $\pi(i) \sim \pi(j) \Leftrightarrow i \sim j$. A graph is called *node transitive*[3] if for every two nodes i, j there exists an automorphism with $\pi(i) = j$; intuitively, this means that all nodes are indistinguishable in the graph since any two nodes may be interchanged without changing the graph structure. Similarly, a graph is called *link transitive* if for every pair of links (i, j) and (x, y), there exists an automorphism with $(\pi(i), \pi(j)) = (x, y)$ (as unordered tuples). See for instance [22] for properties of such graphs. Some common examples of node and link transitive graphs are cycle graphs, the complete graph, the hypercube graph and the Platonic graphs (graph skeletons of the Platonic solids); furthermore, all 'Cayley graphs' are node transitive [22, 107]. For graphs with node and/or link transitivity, we can find the resistance curvatures exactly as follows:

Proposition 6.11 (transitivity and curvature)**.** *Let G be a finite connected graph on n nodes, m links and with constant link weights c:*

[3]This name refers to the fact that the automorphism group, with automorphisms as group elements and composition as group operation, acts transitively on the node set. Usually this is called vertex transitive when nodes are called vertices.

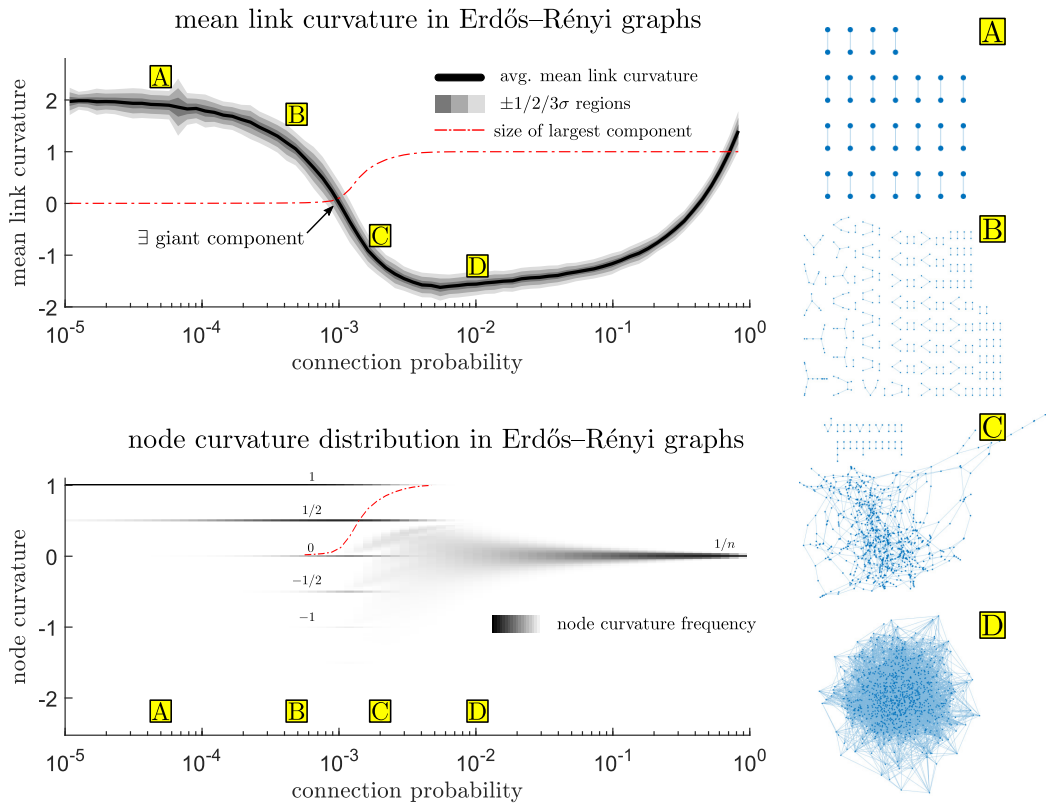


Figure 6.3: The mean link resistance curvature and node resistance curvature distributions in Erdős–Rényi (ER) graphs for different connection probabilities (densities) ρ. We select $N_\rho = 10^4$ connection probabilities $\log_{10}(\rho) \in [-5, -0.05]$ uniformly at random, construct an ER random graph on $n = 10^3$ nodes for each probability and calculate the mean link resistance curvature $\frac{1}{m}\sum_{j\sim i}\kappa_{ij}$ and node resistance curvatures p in each graph. (Top left figure) Plot of the sample mean of the mean link resistance curvature, with graph samples binned together according to their connection probabilities ($N_{\text{bins}} = 75$), and $\pm 1/2/3$ standard deviations around the average. The size of the largest connected component (as a fraction of the total number of nodes) is shown in red, again averaged according to the bins; we remark that the threshold for the appearance of a giant component at $(n-1)\rho = 1$ (see [88]) coincides with the mean link curvature crossing zero. (Bottom left figure) Plot of the aggregate distribution of node curvatures for each of the 75 connection probability bins. The shaded regions are proportional to the observed frequency of nodes with a given curvature p in ER graphs in a given density bin ρ; white–black corresponds to 0–100% observed frequency and to enhance the contrast, we plot the fourth power f^4 of the frequency $f \in [0, 1]$. (Right figure) Four ER graphs (A–D) with different connection probabilities to illustrate the evolution from a collection of many small tree-like components (A,B) for small ρ to the emergence of a larger locally tree-like component (C) and finally, a dense connected graph (D) as ρ further increases. Components consisting of single nodes are omitted for clarity.

(i) if G is node transitive, then it has constant resistance curvature $p = 1/n$,

(ii) if G is link transitive with nodes of (possibly equal) degree r_1 and r_2, then it has constant link resistance curvature $\kappa = \frac{4cm}{n-1} - c(r_1 + r_2)$, and

(iii) if G is node and link transitive, then it has constant node and link resistance curvatures $p = 1/n$ and $\kappa = 2c\rho$, where $\rho = m/\binom{n}{2}$ is the link density.

Proof. *(i)* By [22, Proposition 15.2], a permutation π of the nodes of a graph is an automorphism if and only if its corresponding permutation matrix P leaves the Laplacian matrix invariant as $PQP^T = Q$. Consequently, the structure-preserving row and column permutations of the Laplacian of a node transitive graph will act transitively on the row and column set and we say that the Laplacian is row and column transitive. Since G is connected, the pseudoinverse Laplacian $Q^\dagger$ is the inverse of the Laplacian in the space $\mathrm{span}(\mathbf{u})^\perp$ with the permutation-invariant constant vector $\mathbf{u}$ and we find that $PQP^T = Q \Rightarrow PQ^\dagger P^T = Q^\dagger$ from

$$\begin{aligned}
Q^\dagger &= \left(Q + \frac{\mathbf{u}\mathbf{u}^T}{n}\right)^{-1} - \frac{\mathbf{u}\mathbf{u}^T}{n} \qquad \text{(see Eq. (2.6))} \\
&= \left(PQP^T + \frac{\mathbf{u}\mathbf{u}^T}{n}\right)^{-1} - \frac{\mathbf{u}\mathbf{u}^T}{n} \quad \text{for all } P \text{ with } PQP^T = Q \\
&= P^T\left[\left(Q + \frac{\mathbf{u}\mathbf{u}^T}{n}\right)^{-1} - \frac{\mathbf{u}\mathbf{u}^T}{n}\right]P \qquad (\text{since } P\mathbf{u} = \mathbf{u} \text{ and } P^{-1} = P^T) \\
&= P^T Q^\dagger P \text{ for all } P \text{ with } PQP^T = Q.
\end{aligned}$$

This implies that the pseudoinverse Laplacian $Q^\dagger$ is also row and column transitive and consequently that $Q^\dagger$ has a constant diagonal: for every two nodes i, j, there exists a permutation matrix P such that $PQ^\dagger P^T = Q^\dagger$ and $P\mathbf{e}_i = \mathbf{e}_j$, and thus

$$(Q^\dagger)_{ii} = \mathbf{e}_i^T Q^\dagger \mathbf{e}_i = \mathbf{e}_i^T PQ^\dagger P^T \mathbf{e}_i = \mathbf{e}_j^T Q^\dagger \mathbf{e}_j = (Q^\dagger)_{jj} \text{ for all } i, j.$$

By definition of the resistance curvature in the Fiedler–Bapat identity (Theorem 3.6) we then find that

$$\mathbf{p} = \frac{1}{2}Q\boldsymbol{\zeta} + \frac{\mathbf{u}}{n} \overset{(\boldsymbol{\zeta} \text{ constant})}{=} \frac{\mathbf{u}}{n},$$

which proves that $p_i = 1/n$ for node transitive graphs, as required.

(ii) If a graph is link transitive and not node transitive, then it must be bipartite on $\mathcal{V}_1, \mathcal{V}_2$ and the automorphism group acts transitively on these partitions [107, Lemma 3.2.1]; i.e. the nodes are indistinguishable inside the partitions. In particular, this implies that all n_1 nodes in $\mathcal{V}_1$ have the same degree r_1 and all n_2 nodes in $\mathcal{V}_2$ have

degree r_2. The number of links is then equal to $m = n_1 r_1 = n_2 r_2$. By Foster's theorem and link transitivity — such that $c_{ij}\omega_{ij}$ must be equal for all links — we find that the effective resistance of every link is equal and given by $\omega = \frac{n-1}{n_1 r_1 c} = \frac{n-1}{n_2 r_2 c}$. The node curvatures in the two sets then follow as $p_1 = 1 - \frac{n-1}{2n_1}$ and $p_2 = 1 - \frac{n-1}{2n_2}$ and the link resistance curvature is calculated from their sum, resulting in the proposed formula for κ.

(iii) If a graph is link transitive and node transitive, then $r_1 = r_2$ in the previous derivation, and from $c(r_1+r_2) = 4mc/n$, we find the proposed link resistance curvature $2rc/(n-1)$. This completes the proof. □

Remark 6.12. An alternative proof of statement *(iii)* in Proposition 6.11 uses the fact that if the graph is node transitive. Then $p_i = 1/n$ as before and Foster's Theorem together with link transitivity implies that $\omega = \frac{2(n-1)}{nrc}$, such that $\kappa = 2rc/(n-1)$, as required.

6.2.3 Alternative definitions

There exist many alternative definitions for the node resistance curvature $\mathbf{p}$. This section summarizes the expressions that appeared in this thesis and introduces a few new ones. The alternative definitions not only highlight different aspects of the resistance curvatures and their relation to discrete curvature — in particular, the more 'geometric' definitions in terms of distances — but also illustrate the rich theory of the node resistance curvature, which originates in its definition in terms of the effective resistance and its role in the Fiedler–Bapat identity. In tandem, we summarize some related results for the resistance radius σ^2.

6.2.3.1 Distance–difference characterization

A first (family of) definition(s) for the resistance curvature follows from the inverse relation between the Laplacian and resistance matrix, as captured by the Fiedler–Bapat identity. In particular, from the $Q\Omega$ identity (3.7), we find that

$$\mathbf{p} = \frac{1}{2} Q\Omega\mathbf{f} + \mathbf{f} \text{ for any } \mathbf{u}^T\mathbf{f} = 1.$$

By expanding the matrix product, this can be written as follows:

$$p_i = \frac{k_i}{2} \sum_{k\in\mathcal{N}} f_k \left(\omega_{ik} - \sum_{j:j\sim i} \frac{c_{ij}}{k_i} \omega_{jk} \right) + f_i \text{ for any } \mathbf{u}^T\mathbf{f} = 1. \tag{6.4}$$

This expression in terms of distance differences between a node and its neighbours leads to a number of interesting expressions for the node resistance curvature. With $\mathbf{f} = \mathbf{e}_i$, expression (6.4) retrieves the definition in terms of relative resistances: $p_i = 1 - \frac{1}{2}\sum_{j\sim i} c_{ij}\omega_{ij}$. With $\mathbf{f} = \mathbf{e}_x$ for some $x \neq i$, expression (6.4) yields

$$p_i = \frac{k_i}{2}\left(\omega_{ix} - \sum_{j:j\sim i}\frac{c_{ij}}{k_i}\omega_{jx}\right) \text{ for any } x \neq i. \tag{6.5}$$

From Definition (5.11) of the graph boundary in Chapter 5, this implies that the graph boundary with respect to ω consists of nodes with nonnegative curvature $\partial_\omega G = \{i : p_i \geq 0\}$. For $\mathbf{f} = \mathbf{u}/n$ and using relation (3.4) between the pseudoinverse Laplacian diagonal and average resistance distances, expression (6.4) yields

$$\mathbf{p} = \frac{1}{2}Q\boldsymbol{\zeta} + \frac{\mathbf{u}}{n} \quad \text{and} \quad p_i = \frac{1}{2}\sum_{j:j\sim i} c_{ij}(\langle\omega_i\rangle - \langle\omega_j\rangle) + \frac{1}{n}.$$

In other words, the resistance curvature is determined by the Laplacian of the average resistance distances.

The $Q\Omega$ identity also leads to a family of distance-based definitions for the resistance radius; invoking the $Q\Omega$ identity (3.7) and the equilibrium characterization, equation (6.10) in Section 6.2.3.3, we find that $\Omega Q\Omega = -2\Omega + 4\sigma^2\mathbf{u}\mathbf{u}^T$, from which it follows that

$$\sigma^2 = \frac{1}{4}(\Omega Q\Omega)_{xx} = \frac{1}{4}\sum_{j\sim i} c_{ij}(\omega_{ix} - \omega_{jx})^2 \text{ for any } x. \tag{6.6}$$

Since this holds for any x, we also find $4n\sigma^2 = \operatorname{tr}(\Omega Q\Omega)$.

6.2.3.2 Dynamical characterization

The diffusion of a unit mass at a node i results in a distribution $\boldsymbol{\rho}_i(t) := \exp(-Qt)\mathbf{e}_i$ for $t \geq 0$. This distribution can be interpreted as a growing 'ball' around the node (see also Section 6.3.2.2) or as the occupation probability at time t of a continuous-time random walker starting from node i. Introducing this diffusion distribution as $\mathbf{f} = \boldsymbol{\rho}_i(t)$ in expression (6.4) leads to the following dynamical characterization for the resistance curvatures:

Proposition 6.13. *The node and link resistance curvature are equal to*

$$p_i = \lim_{t\to 0}\left(1 - \frac{1}{4t}\mathbb{E}(\omega_{N_tM_t})\right) \text{ with } N_t, M_t \overset{d}{\sim} \boldsymbol{\rho}_i(t) \text{ independently} \tag{6.7}$$

$$\kappa_{ij} = \lim_{t\to 0}\frac{1}{t}\left(1 - \frac{\mathbb{E}(\omega_{N_tM_t})}{\omega_{ij}}\right) \text{ with } N_t \overset{d}{\sim} \boldsymbol{\rho}_i(t), M_t \overset{d}{\sim} \boldsymbol{\rho}_j(t). \tag{6.8}$$

Proof. The average resistance distance between a pair of random nodes $N \overset{d}{\sim} \mathbf{f}$ and $M \overset{d}{\sim} \mathbf{g}$ is equal to $\mathbb{E}(\omega_{NM}) = \sum_{i,j} f(i)g(j)\omega_{ij} = \mathbf{f}^T\Omega\mathbf{g}$. Together with the definition of the diffusion distribution, the right-hand side of expression (6.7) can then be written as

$$\lim_{t\to 0}\left(1 - \frac{1}{4t}\mathbb{E}(\omega_{N_tM_t})\right) = \lim_{t\to 0}\left(1 - \frac{1}{4t}\mathbf{e}_i^T \exp(-Qt)\Omega\exp(-Qt)\mathbf{e}_i\right) \text{ for } N_t, M_t \overset{d}{\sim} \boldsymbol{\rho}_i(t).$$

By definition of the matrix exponential as the power series $\exp(A) = \sum_{k=0}^{\infty} A^k/(k!)$, the leading order term of the diffusion distribution around i in the $t \to 0$ limit is $(I - Qt)\mathbf{e}_i$. Making use of the $Q\Omega$ identity (3.7), we find

$$(I - Qt)\Omega(I - Qt) = \Omega + 4I - (2\mathbf{p}\mathbf{u}^T + 2\mathbf{u}\mathbf{p}^T)t - 2Qt^2,$$

and thus

$$\lim_{t\to 0}\left(1 - \frac{1}{4t}\mathbb{E}(\omega_{N_tM_t})\right) = \lim_{t\to 0}\left(p_i + \frac{1}{2}k_it\right) = p_i \text{ for } N_t, M_t \overset{d}{\sim} \boldsymbol{\rho}_i(t),$$

as required. Similarly, for the link resistance curvature, we find that the right-hand side in expression (6.8) is equal to

$$\lim_{t\to 0}\frac{1}{t}\left(1 - \frac{\mathbb{E}(\omega_{N_tM_t})}{\omega_{ij}}\right) = \lim_{t\to 0}\left(\frac{2(p_i + p_j)}{\omega_{ij}} - \frac{2c_{ij}t}{\omega_{ij}}\right) = \kappa_{ij} \text{ for } N_t \overset{d}{\sim} \boldsymbol{\rho}_i(t) \text{ and } M_t \overset{d}{\sim} \boldsymbol{\rho}_j(t)$$

as required. This completes the proof. □

In other words, the node resistance curvature is related to the *average distance between nodes in its neighbourhood*, as defined by the diffusion equation, and the link resistance curvature is related to *the average distance between the neighbourhoods of the end nodes of the link*. Equivalently, $\mathbb{E}(\omega_{N_tM_t})$ corresponds to the average distance between two independent continuous-time random walkers (as defined by the diffusion equation) starting at a node or at the end nodes of a link. Changing the diffusion equation to a normalized variant with $\tilde{\boldsymbol{\rho}}_i(t) = \exp(-Q\,\mathrm{diag}(\mathbf{k})^{-1}t)\mathbf{e}_i$, we arrive at normalized variants of the node and link resistance curvature:

$$\tilde{p}_i = \frac{p_i}{k_i} \quad \text{and} \quad \tilde{\kappa}_{ij} = \frac{2(\tilde{p}_i + \tilde{p}_j)}{\omega_{ij}}.$$

In Section 6.3.2.2, we discuss how the dynamical characterization for the link resistance curvature relates to the definition of Ollivier–Ricci curvature.

The average distance between two identically and independently distributed random nodes may also be interpreted as a variance since $\mathbb{E}(\omega_{NM}) = \mathbf{f}^T\Omega\mathbf{f} = 2\,\mathrm{var}(\mathbf{f})$

with $N, M \overset{d}{\sim} f$ independently. This allows one to define the node resistance curvature as

$$p_i = \lim_{t \to 0} \left(1 - \frac{1}{2t} \operatorname{var}(\boldsymbol{\rho}_i(t)) \right). \tag{6.9}$$

The diffusion variance thus grows as $2t$ around nodes with zero node resistance curvature. This is equal to the variance growth for diffusion or, equivalently, the mean square displacement of Brownian motion on the real line [84].

6.2.3.3 Equilibrium characterization

We recall from Corollary 3.7 the following expressions for the resistance curvature and radius:

$$\mathbf{p} = \frac{\Omega^{-1}\mathbf{u}}{\mathbf{u}^T\Omega^{-1}\mathbf{u}} \quad \text{and} \quad \sigma^2 = \frac{1}{2}(\mathbf{u}^T\Omega^{-1}\mathbf{u})^{-1}. \tag{6.10}$$

As discussed in Section 3.4, this definition relates to Leinster's magnitude and in Section 6.3.2.4 we discuss the relation between definition (6.10) and Steinerberger's discrete curvature [226]. The 'equilibrium' terminology is inspired by the theory of electrostatics and potential theory (see for instance [17]).

6.2.3.4 Variational characterization

In the context of maximum variance distributions, we found that $\sigma_\star^2$, the resistance radius of a particular Kron reduction $G^\star$ of a graph (see Theorem 5.23), maximizes the variance over all distributions Δ with the maximum achieved at the maximum variance distribution $\mathbf{p}^\star$. Similarly, we find the following variational characterization for the resistance curvature and radius as the solution to an optimization problem:

Proposition 6.14. *The resistance curvature and radius of a graph satisfy*

$$\sigma^2 = \max_{\mathbf{f}:\mathbf{u}^T\mathbf{f}=1} \tfrac{1}{2}\mathbf{f}^T\Omega\mathbf{f}, \tag{6.11}$$

with the maximum achieved uniquely at $\mathbf{f}^\star = \mathbf{p}$.

Proof. Let $\mathbf{f} \in \mathbb{R}^n$ be any unit-sum vector distinct from $\mathbf{p}$. Then we find

$$0 \overset{(a)}{>} (\mathbf{f}-\mathbf{p})^T\Omega(\mathbf{f}-\mathbf{p}) = \mathbf{f}^T\Omega\mathbf{f} - \mathbf{p}^T\Omega\mathbf{p} - 2(\mathbf{f}-\mathbf{p})^T\Omega\mathbf{p} \overset{(b)}{=} \mathbf{f}^T\Omega\mathbf{f} - \mathbf{p}^T\Omega\mathbf{p},$$

where step (a) follows from expression (3.2) or equivalently from the fact that the resistance distance has strict negative type, and step (b) uses the equilibrium expression $\Omega\mathbf{p} = 2\sigma^2\mathbf{u}$ in combination with the fact that $(\mathbf{f}-\mathbf{p})^T\mathbf{u} = 0$. This implies that $2\sigma^2 = \mathbf{p}^T\Omega\mathbf{p} > \mathbf{f}^T\Omega\mathbf{f}$ for all unit-sum vectors $\mathbf{f} \neq \mathbf{p}$, and completes the proof. □

6.2.3.5 Geometric characterization

Proposition 4.32 relates the resistance curvature and radius to the circumscribed sphere of the hyperacute simplex S associated to a graph. More precisely, $\mathbf{p}$ is the unit-sum coordinate of the circumcenter of S and σ is the circumradius. The circumsphere is thus defined as

$$\{\mathbf{x} \in \mathbb{R}^{n-1} : \|\varphi(\mathbf{p}) - \mathbf{x}\|^2 = \sigma^2\}$$

which contains the vertices $\varphi(\mathbf{e}_i)$ of S. Similarly, for each $\mathcal{V}$-face of the simplex S, the circumcenter and radius are defined by the resistance curvature and radius of the corresponding $\mathcal{V}$-Kron reduction. The relations between circumcenters and circumradii of different faces are determined by the Kron reduction formula (3.15); see also Section 4.3.2.

6.3 Curvature properties

So far, the only hints towards an interpretation of p and κ as notions of discrete curvature have been some examples of graphs with negative/zero/positive curvature. This section presents three arguments that further evidence this interpretation. A brief summary of the arguments is given below and each argument is then presented in more detail in the following subsections. These arguments do not constitute a 'proof' that p is a discrete scalar curvature and κ a discrete Ricci curvature, but instead highlight some relevant connections and provide a starting point for further research.

(1) The resistance curvatures are constant and of the correct sign for a number of graphs associated to constant curvature spaces. We show this for infinite regular lattices, infinite regular trees and regular tilings of the sphere.

(2) The resistance curvatures are related to established notions of discrete curvature on graphs. We show that the node resistance curvature is related to combinatorial curvature and Steinerberger's curvature and that the link resistance curvature satisfies 'Forman curvature $\leq$ link resistance curvature $\leq$ Ollivier curvature'.

(3) We present numerical and theoretical evidence that the node resistance curvature is zero in random geometric graphs in $\mathbb{R}^2$.

We note again that aside from being a potential contribution to the theory of discrete curvature, the proposed interpretation of p and κ as discrete curvatures is also important for the theory of effective resistances. As the results in Section 6.4 and 6.5 illustrate, thinking of p and κ as curvatures can guide their study by suggesting specific problems such as the discrete Ricci flow and graphs with positive resistance curvature.

6.3.1 Constant curvature graphs

Positive curvature (platonic graphs) Platonic graphs are the graph skeletons of the platonic solids (tetrahedron, cube, octahedron, dodecahedron, icosahedron), which are regular tilings of the 2-sphere $\mathbb{S}^2$. These graphs are node and link transitive and thus have constant and positive resistance curvatures $p_i = 1/n$ and $\kappa_{ij} = 2\rho$, where $\rho = m/\binom{n}{2}$ is the link density. This is in correspondence with the constant positive curvature of $\mathbb{S}^2$.

Zero curvature (infinite lattice) The rectangular/triangular/hexagonal lattices are infinite graphs that correspond to (the graph skeleton of) the regular tilings of the Euclidean plane. The effective resistance in these lattices has been studied extensively, and it was shown that the relative resistance of all links is equal to $2/d$ with d the combinatorial degree of the respective lattice [80, 97, 237]. As a result, lattice graphs have zero resistance curvatures $p_i = 0$ and $\kappa_{ij} = 0$, in correspondence with the zero curvature of $\mathbb{R}^2$.

Negative curvature (infinite regular tree) The infinite regular tree or Bethe lattice [180] corresponds to a regular tiling of the hyperbolic plane $\mathbb{H}^2$. Like all trees, the relative resistance of every link is equal to $c_{ij}\omega_{ij} = 1$, which implies that if $d > 2$ then these graphs have negative resistance curvatures $p_i = 1 - d/2$ and $\kappa_{ij} = 2c(2-d)$ in correspondence with the constant negative curvature of $\mathbb{H}^2$.

An example of these constant curvature graphs is shown in Figure 6.1. We remark that the correct curvature in regular tilings is not always reproduced correctly by other notions of discrete curvature. As noted in [141] for instance, the hexagonal lattice is often assigned a negative curvature because it is locally tree-like.

6.3.2 Relation to other discrete curvatures

As a second argument, we show that the resistance curvatures are related to existing notions of discrete curvature. In particular, we show that p is related to combinatorial

curvature (through random spanning trees) and Steinerberger's curvature (through the equilibrium definition) and that the link resistance curvature is related to Ollivier–Ricci and Forman–Ricci curvature (through tight two-sided bounds).

6.3.2.1 Combinatorial curvature

Combinatorial curvature [115, 138] measures discrete curvature for graphs embedded in the plane (or other surfaces) — i.e. with the nodes and links drawn in the plane such that the links only intersect at their endpoints. These drawings partition the plane into *faces* which are connected regions of the plane bordered by links that form a cycle in the graph. The combinatorial curvature is then defined as [138]

$$p_i^{(\mathrm{co})} := 1 - \frac{d_i}{2} + \sum_{\text{face } f \ni i} \frac{1}{d_f},$$

where $f \ni i$ are the faces that contain node i in their boundary. In tree graphs, there is only a single unbounded face, which does not contribute to the combinatorial curvature. As a result, every node in a tree graph has combinatorial curvature $1-d_i/2$ equal to the node resistance curvature in the tree. Following the connection between relative resistances and random spanning trees in Theorem 2.13, we find

$$\begin{aligned} p_i &= 1 - \frac{1}{2}\sum_{j\sim i} c_{ij}\omega_{ij} = 1 - \frac{1}{2}\sum_{j\sim i}\sum_{T\in\mathcal{T}} \Pr[\mathbf{T}=T]\mathbf{1}_{\{(i,j)\in T\}} \quad \text{(by Theorem 2.13)} \\ &= \sum_{T\in\mathcal{T}} \Pr[\mathbf{T}=T]\left(1 - \frac{1}{2}\sum_{j\sim i}\mathbf{1}_{\{(i,j)\in T\}}\right) = \sum_{T\in\mathcal{T}} \Pr[\mathbf{T}=T]\left(1 - \frac{d_i^{(T)}}{2}\right). \end{aligned}$$

In other words, the resistance curvature of a node i in a graph is equal to the expected combinatorial curvature of i in a random spanning tree:

$$p_i = \mathbb{E}[p_i^{(co)}(\mathbf{T})] \text{ with random spanning tree } \mathbf{T}. \tag{6.12}$$

6.3.2.2 Ollivier–Ricci curvature

Yann Ollivier [189] introduced a notion of curvature for metric spaces with an associated Markov chain. This notion of discrete curvature can be applied to graphs, for instance with the shortest-path distance as metric and with random walks as Markov chain.

For every node i, the Markov chain determines a distribution[4] $\boldsymbol{\mu}_{t,i} \in \Delta$ on the

[4]The subscript 't' of the Markov chain is the *laziness parameter* of the random walk and should not be confused with the 'Markov time' of a continuous-time Markov chain, which is often denoted by t as well. For a given Markov chain with transition probabilities $T_{ij} = \Pr[j \text{ at step } k+1 | i \text{ at step } k]$, the lazy Markov chain has transition probabilities $T'_{ij} = (1-t)I + tT_{ij}$; i.e. with probability $(1-t)$ to stay at the same node each step.

nodes which is concentrated around i and can thus be thought of as a 'ball' around i in the graph; think for instance of the occupation probability of a random walker a few steps after starting at node i. The (continuous) laziness parameter $t > 0$ captures how tightly the balls are concentrated around the nodes. The metric d can be used to provide a notion of distance between these balls by making use of the 1-Wasserstein distance W_1:

$$W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j}) := \min_P \operatorname{tr}(PD),$$

with distance matrix D and where the minimum is taken over all nonnegative $n \times n$ matrices P with $P\mathbf{u} = \boldsymbol{\mu}_{t,i}$ and $P^T\mathbf{u} = \boldsymbol{\mu}_{t,j}$; these matrices determine joint distributions with the balls $\boldsymbol{\mu}_{t,i}$ and $\boldsymbol{\mu}_{t,j}$ as marginals. *Ollivier–Ricci* (OR) curvature measures how much the *direct distance* $d(i,j)$ between two nodes differs from the *distance between balls* around these nodes $W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j})$ as a generalization of continuous Ricci curvature to the data $(\mathcal{N}, d, \boldsymbol{\mu}_t)$. Lin, Lu & Yau [161] further modified Ollivier's definition to a limit for shrinking balls as follows:

$$\kappa_{ij}^{(\mathrm{OR})} := \lim_{t \to 0} \frac{1}{t} \left(1 - \frac{W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j})}{d(i,j)} \right).$$

The two distances W_1 and d converge for $t \to 0$, and OR curvature thus measures how much they differ in the first order in t. If the distance between the balls is larger than the direct distance between the points, then $\kappa^{(\mathrm{OR})}$ will be negative, and the other way around for positive curvature.

The Ollivier–Ricci curvature on graphs is usually studied with respect to the shortest-path distance as metric and lazy random walks as Markov chain, defined by $\boldsymbol{\mu}_{t,i} = (I - Qt)\mathbf{e}_i$; see for instance [161, 181]. Since OR curvature is defined in the $t \to 0$ limit, the lazy random walk can also be approximated by the diffusion distribution $\boldsymbol{\rho}_i(t) = \exp(-Qt)\mathbf{e}_i$. This highlights the similarity between $\kappa_{ij}^{(\mathrm{OR})}$ and the link resistance curvature κ_{ij} in expression (6.8); the former is defined based on the Wasserstein distance between two distributions $\boldsymbol{\mu}_{t,i}$ and $\boldsymbol{\mu}_{t,j}$ whereas the latter is defined based on the mean resistance distance between[5] distributions $\boldsymbol{\rho}_i(t)$ and $\boldsymbol{\rho}_j(t)$; i.e. between pairs of nodes sampled from these distributions. As a consequence, we find the following relation when measuring OR curvature with respect to the effective resistance:

[5]One way to think of this difference, perhaps, is that there is a trade-off between simplicity in the node distance (d_{sp} is simple while ω_{ij} is complex) and simplicity in the ball distance (expected distance is simple while Wasserstein distance is complex).

Proposition 6.15. *The Ollivier–Ricci curvature with respect to the resistance distance and the lazy random walk is lower-bounded by the resistance curvature:* $\kappa_{ij}^{(OR)} \geq \kappa_{ij}$*, with equality if* (i,j) *is a cut link.*

Proof. We start by bounding the Wasserstein distance between the balls $\boldsymbol{\mu}_{t,i} = (I - Qt)\mathbf{e}_i$ and $\boldsymbol{\mu}_{t,j} = (I - Qt)\mathbf{e}_j$:

$$W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j}) = \min_P \operatorname{tr}(P\Omega) \overset{(a)}{\leq} \operatorname{tr}(\boldsymbol{\mu}_{t,i}\boldsymbol{\mu}_{t,j}^T\Omega) \overset{(b)}{=} \boldsymbol{\mu}_{t,j}^T\Omega\boldsymbol{\mu}_{t,i}, \tag{6.13}$$

where inequality (a) follows from the fact that $P = \boldsymbol{\mu}_{t,i}\boldsymbol{\mu}_{t,j}^T$ is a valid matrix for the Wasserstein distance — i.e. with nonnegative entries and the correct marginals — and where equality (b) invokes properties of the trace operator. By definition of $\boldsymbol{\mu}_{t,i}$ and $\boldsymbol{\mu}_{t,j}$, we obtain

$$W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j}) \leq \mathbf{e}_j^T(I - Qt)\Omega(I - Qt)\mathbf{e}_i \overset{(a)}{=} \omega_{ij} - 2t(p_i + p_j) + 2t^2 c_{ij},$$

where (a) follows from the $Q\Omega$ identity. With this bound for the Wasserstein distance, we find for the Ollivier–Ricci curvature with ω as distance:

$$\kappa_{ij}^{(\mathrm{OR})} = \lim_{t\to 0} \frac{1}{t}\left(1 - \frac{W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j})}{\omega_{ij}}\right) \geq \frac{2(p_i + p_j)}{\omega_{ij}}, \tag{6.14}$$

establishing the bound in Proposition 6.15.

Next, we show that equality is achieved in the case of cut links. If (i,j) is a cut link, then i is a cut node and thus for any two nodes x, y, which are disconnected by removal of i, we have the triangle equality $\omega_{xy} = \omega_{xi} + \omega_{iy}$. Similarly, j is a cut node. In particular, the triangle equality holds for effective resistances between the neighbours of i and the neighbours of j; in other words between the supports $\mathcal{I} := \operatorname{supp}(\boldsymbol{\mu}_{t,i})$ and $\mathcal{J} := \operatorname{supp}(\boldsymbol{\mu}_{t,j})$ of the balls around i and j respectively. Note that $i, j \in \mathcal{I}$ and $i, j \in \mathcal{J}$ by definition of the balls and by $i \sim j$. The effective resistances satisfy $\omega_{xy} = \omega_{xi} + \omega_{ij} + \omega_{jy}$ for all $x \in \mathcal{I}\backslash\{j\}$ and $y \in \mathcal{J}\backslash\{i\}$, and consequently the block matrix $\Omega_{\mathcal{J}\mathcal{I}}$ has the following decomposition:

$$\Omega_{\mathcal{J}\mathcal{I}} = \boldsymbol{\omega}_j\mathbf{u}^T + \mathbf{u}\boldsymbol{\omega}_i^T + \omega_{ij}(\mathbf{u}\mathbf{u}^T - 2\mathbf{e}_i\mathbf{u}^T - 2\mathbf{u}\mathbf{e}_j^T + 2\mathbf{e}_i\mathbf{e}_j^T)$$

where $\boldsymbol{\omega}_i$ is the $|\mathcal{I}| \times 1$ vector with effective resistances to the neighbours of i, as $(\boldsymbol{\omega}_i)_x = \omega_{ix}$ for $x \in \mathcal{I}$ and similarly for $\boldsymbol{\omega}_j$, and with $\mathbf{u}$ and $\mathbf{u}^T$ the all-one vectors of

appropriate size. The Wasserstein distance thus satisfies:

$$
\begin{aligned}
&W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j}) = \min_P \operatorname{tr}(P\Omega) \overset{\text{(a)}}{=} \min_P \operatorname{tr}(P_{\mathcal{I}\mathcal{J}}\Omega_{\mathcal{J}\mathcal{I}}) \\
&= \min_P \left(\mathbf{u}^T P_{\mathcal{I}\mathcal{J}}\boldsymbol{\omega}_j + \boldsymbol{\omega}_i^T P_{\mathcal{I}\mathcal{J}}\mathbf{u} + \omega_{ij}\left[\mathbf{u}^T P_{\mathcal{I}\mathcal{J}}\mathbf{u} - 2\mathbf{u}^T P_{\mathcal{I}\mathcal{J}}\mathbf{e}_i - 2\mathbf{e}_j^T P_{\mathcal{I}\mathcal{J}}\mathbf{u} + 2\mathbf{e}_j^T P_{\mathcal{I}\mathcal{J}}\mathbf{e}_i\right]\right) \\
&\overset{\text{(b)}}{=} \min_P \left(\boldsymbol{\mu}_{t,j}^T(\boldsymbol{\omega}_j - 2\omega_{ij}\mathbf{e}_i) + (\boldsymbol{\omega}_i^T - 2\omega_{ij}\mathbf{e}_j^T)\boldsymbol{\mu}_{t,i} + \omega_{ij}(1 + 2P_{ji})\right) \\
&= \boldsymbol{\mu}_{t,j}^T\Omega\boldsymbol{\mu}_{t,i} - 2\omega_{ij}\boldsymbol{\mu}_{t,j}^T\mathbf{e}_i\mathbf{e}_j^T\boldsymbol{\mu}_{t,i} + 2\omega_{ij}\min_P P_{ji} \\
&\overset{\text{(c)}}{\geq} \boldsymbol{\mu}_{t,j}^T\Omega\boldsymbol{\mu}_{t,i} - 2c_{ij}t^2
\end{aligned}
$$

where in step (a) we use that a nonnegative matrix P with $\operatorname{supp}(P\mathbf{u}) = \mathcal{I}$ and $\operatorname{supp}(P^T\mathbf{u}) = \mathcal{J}$ can only have nonzero entries in $\mathcal{I} \times \mathcal{J}$, in step (b) we introduce the marginals of P and in step (c) we use nonnegativity of P_{ji} for the minimization and the fact that $(\boldsymbol{\mu}_{t,i})_j = (\boldsymbol{\mu}_{t,j})_i = c_{ij}t$ and $c_{ij}\omega_{ij} = 1$ since the link is a cut link. With the definitions of $\boldsymbol{\mu}_{t,i}$ and $\boldsymbol{\mu}_{t,j}$, we obtain

$$
W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j}) \geq \mathbf{e}_j^T(I - Qt)\Omega(I - Qt)\mathbf{e}_i - 2c_{ij}t^2 = \omega_{ij} - 2t(p_i + p_j)
$$

such that

$$
\kappa_{ij}^{(\mathrm{OR})} = \lim_{t\to 0}\frac{1}{t}\left(1 - \frac{W_1(\boldsymbol{\mu}_{t,i}, \boldsymbol{\mu}_{t,j})}{\omega_{ij}}\right) \leq \frac{2(p_i + p_j)}{\omega_{ij}} \qquad \text{(if } (i,j) \text{ is a cut link).}
$$

Combined with the general lower-bound (6.14) for the OR curvature, this proves that equality must hold between $\kappa_{ij}^{(\mathrm{OR})}$ and κ_{ij} when (i,j) is a cut link. □

In tree graphs, all links are cut links and the link resistance curvature thus corresponds to the OR curvature throughout the graph. We recall that the resistance distance is equal to the shortest-path distance in trees.

The normalized lazy random walk, defined by $\boldsymbol{\mu}_{t,i} = (I - Q\operatorname{diag}(\mathbf{k})^{-1}t)\mathbf{e}_i$, where $\operatorname{diag}(\mathbf{k})$ is the diagonal matrix with the weighted degrees on its diagonal, is an alternative choice for the Ollivier–Ricci curvature Markov chain [161, 181]. Similar to the standard lazy random walk, this random walk corresponds to the $t \to 0$ limit of the continuous-time normalized random walk, and we find the bound

$$
\kappa_{ij}^{(\mathrm{LLY})} \geq \frac{2(p_i/k_i + p_j/k_j)}{\omega_{ij}}. \tag{6.15}
$$

This again suggests the definition of a degree-normalized version of the link resistance curvature. The superscript LLY refers to the authors Lin, Lu and Yau of [161], where this variant of OR curvature was first considered.

6.3.2.3 Forman–Ricci curvature

In [99], Robin Forman introduced a notion of curvature for 'CW complexes'[6], which is a class of topological spaces with combinatorial descriptions that includes graphs; see for instance [124] for a definition. This so-called Forman–Ricci (FR) curvature is defined by generalizing the definition of the classical Ricci curvature in terms of the 'Bochner Laplacian' of a manifold to a definition for discrete spaces based on an analogous Laplacian on these spaces; see also [137]. The FR curvature is expressed in terms of local combinatorial data around a considered point, and Sreejith et al. [223] translated Forman's general definition to graphs, as

$$\kappa_{ij}^{(\mathrm{FR})} = 2w_i \left(1 - \frac{1}{2}\sum_{k\sim i}\sqrt{\frac{w_{ij}}{w_{ik}}}\right) + 2w_j \left(1 - \frac{1}{2}\sum_{k\sim j}\sqrt{\frac{w_{ij}}{w_{jk}}}\right) \tag{6.16}$$

where w_* and w_{**} are nonzero weights associated to the nodes and links of the graph. Expression (6.16) clearly resembles Definition (6.2) of κ in terms of the node resistance curvature and indeed, choosing unit weights yields the following relation:

Proposition 6.16. *The Forman–Ricci curvature with respect to unit weights is upper-bounded by the link resistance curvature as* $\kappa_{ij}^{(FR)}/\omega_{ij} \leq \kappa_{ij}$*, with equality if and only if* (i, j) *is a cut link.*

Proof. Starting from the definition of FR curvature (6.16) with unit weights $w = 1$, we find that

$$\kappa_{ij}^{(\mathrm{FR})} = 2\left(1 - \frac{1}{2}\sum_{k\sim i} 1\right) + 2\left(1 - \frac{1}{2}\sum_{k\sim j} 1\right) = 4 - d_i - d_j \overset{(a)}{\leq} \omega_{ij}\kappa_{ij}$$

where (a) is the lower bound for κ_{ij} in Property 6.5. This completes the proof. □

With the inverse degree as node weights $w_i = 1/k_i$ and unit link weights $c = 1$, we find the bound $\kappa_{ij}^{(\mathrm{FR})}/\omega_{ij} \leq 2(p_i/k_i + p_j/k_j)/\omega_{ij}$, which again features the degree-normalized version of the link resistance curvature similar to equation (6.15).

The relation between the link resistance curvature and Forman–Ricci and Ollivier–Ricci curvature can be summarized by the two-sided bound

$$\kappa_{ij}^{(\mathrm{OR})} \geq \kappa_{ij} \geq \kappa_{ij}^{(\mathrm{FR})}/\omega_{ij} \text{ with equality for cut links.}$$

[6]The acronym CW stands for 'closure-finite' and 'weak topology' in reference to two important properties of these objects [124].

Importantly, these are OR and FR curvatures with respect to specific data: ω as a metric for OR and $w = 1$ as weights for FR. In addition to serving as an argument for the interpretation of link resistance curvature as a discrete curvature, this result is relevant in the context of other works that investigate the relation between both curvatures [137, 207, 234].

6.3.2.4 Steinerberger's curvature

Steinerberger [226] introduced a discrete curvature based on the shortest-path distance matrix D of a graph, defined as the vector $\mathbf{w}$ that solves $D\mathbf{w} = n\mathbf{u}$, the equilibrium equation. If multiple solutions exist, then any solution that maximizes $\min_i w_i$ is chosen, and if the system has no solutions, then the Moore–Penrose pseudoinverse $\mathbf{w} = nD^{\dagger}\mathbf{u}$ may be used instead. Replacing the shortest-path distance matrix D by the resistance matrix Ω retrieves the equilibrium definition for $\mathbf{p}$, and we find the relation $\mathbf{w} = n\mathbf{p}/(2\sigma^2)$ between Steinerberger's curvature and the node resistance curvature.

6.3.3 Zero curvature in random geometric graphs

A natural question for discrete curvatures is whether they converge to the classical continuous curvature on discrete structures that represent finer and finer discretizations of some continuous space. For instance, Cheeger, Müller & Schrader [50] showed a curvature convergence result for piecewise flat spaces that approximate an underlying manifold and van der Hoorn et al. [246] recently proved convergence of Ollivier–Ricci curvature to Ricci curvature in a specific continuum limit of random geometric graphs sampled from Riemannian manifolds. As the third argument for our interpretation of the resistance curvature $\mathbf{p}$ as a discrete curvature, we show numerical and theoretical evidence that the resistance curvature in random geometric graphs in $\mathbb{R}^2$ is equal to the underlying zero curvature of the Euclidean plane.

We recall that a random geometric graph (RGG) is a random graph constructed on a domain $\mathbb{D} \subseteq \mathbb{R}^2$ in the plane from which points are sampled by a Poisson point process with a homogeneous rate λ; the expected number of points is equal to $N = \lambda \mathrm{Area}(\mathbb{D})$ so we may equivalently fix a desired N. These points are then taken as the nodes of a graph and pairs of nodes are linked if they lie at most a certain connection radius (distance) r apart from each other. All together, an RGG can thus be parametrized by $(\mathbb{D}, N, r)$. See [195] for more information on RGGs and their properties and Figure 6.4 for an illustration of their construction and some examples.

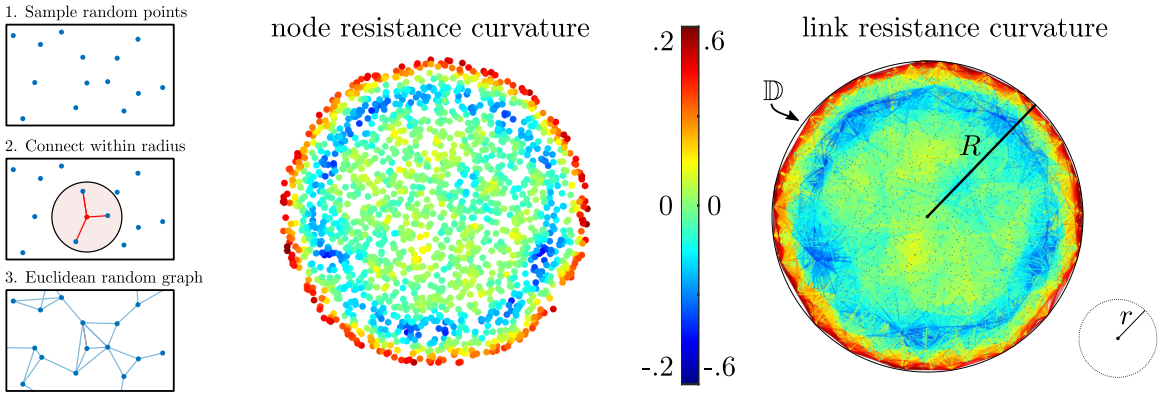


Figure 6.4: (Left panels) Illustration of the construction of an RGG: points are randomly sampled from a domain $\mathbb{D} \subseteq \mathbb{R}^2$ and connected if they lie within a certain connection radius r. (Right figure) The resistance curvatures are calculated in an RGG on a disc with radius $R/r = 4$ and expected number of nodes $N = 2000$, with red/green/blue indicating positive/zero/negative values. Visually, both the node and link resistance curvatures are close to zero near the center of the disc but become negative and then positive going closer to the boundary. More detailed experimental results are shown in Figure 6.5 and the mechanism behind the boundary effect is explained in the main text and Section 6.3.3.

To simplify the setup for further analysis, we consider RGGs on the disc $\mathbb{D} = \{x \in \mathbb{R}^2 : \|x\| \leq R\}$ of radius R, such that nodes at the same radial distance D_i from the boundary are statistically equivalent, and the parameters reduce to $(R/r, N)$.

As a first numerical result, Figure 6.4 shows the resistance curvatures in an RGG on the disc with $R/r = 4$ and $N = 2000$ nodes. Near the center of the disc, the resistance curvatures are close to zero, while moving towards the boundary, they first become negative and then positive at the boundary. Figure 6.5 illustrates this result in more detail. The mean and standard deviations of p_i are plotted with respect to D_i/r with data aggregated over 100 samples of an RGG with $R/r = 5$ and $N = 5000$. The main observation in Figure 6.5 is again that p_i is close to zero in the center of the domain — moreover, it appears to be a zero-mean random variable — and becomes negative and then positive when moving closer to the boundary.

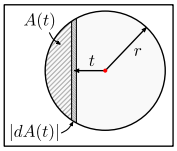
We now develop a model for the resistance curvature in Euclidean RGGs that partially explains these observations. While a full understanding of the relative resistance in RGGs is lacking, in particular concerning dependencies between the relative resistance of different links, we find that the numerical observations in Figures 6.4 and 6.5 can be explained in great detail using a simple heuristic for the effective resistance.

In [251, 252], von Luxburg et al. showed that in RGGs with increasing number of nodes, the effective resistance between any pair of nodes i, j converges to the sum of their inverse degrees: $d_i^{-1} + d_j^{-1}$. While the degree of a node depends on the specific random graph realization, we know by properties of Poisson sampling that the expected degree of a given node i is determined by the area of overlap between the domain and an r-radius disc centered at i; more precisely $\mathbb{E}(d_i) = \lambda S(i)$, where $S(i) := \text{Area}(\{x \in \mathbb{D} : \|x - i\| \leq r\})$, with expectation taken over the random graph ensemble. Combining these two results, we propose the following heuristic $\hat{\omega}$ for the effective resistance in RGGs:

$$\hat{\omega}_{ij} := \frac{1}{\lambda S(i)} + \frac{1}{\lambda S(j)} \text{ for all links } i \sim j. \tag{6.17}$$

Importantly, this heuristic reduces the relative resistance of a link to a purely geometric and local quantity that is only determined by how the r-radius region around the link overlaps with the domain $\mathbb{D}$. In the derivation in Section 6.3.3.1, we show that this heuristic, combined with the assumption that $r \ll R$, allows one to calculate the expected resistance curvature $\hat{p}$ as a function of the distance to the boundary D as

$$\mathbb{E}(\hat{p}(D)) = \frac{1}{2}\left(1 - \int_{t=-r}^{\min(D,r)} \frac{|dA(t)|}{A(\max(-r, t-D))}\right) \tag{6.18}$$



where $A(t)$ is the area of a circular segment at height t and $|dA(t)|$ the change of segment size — as illustrated in the figure on the right — and with expectation taken over the RGG ensemble. While we were not able to find a general closed-form expression for the integral in (6.18), the expression can be evaluated numerically and is shown in Figure 6.5 by the red line. Importantly, Figure 6.5 shows that formula (6.18) matches very well with the experimentally observed resistance curvatures (the black line). This agreement suggests that the heuristic (6.17) captures some of the mechanisms behind the experimentally observed convergence of resistance curvature in random geometric graphs.

Due to the appearance of min and max operations, expression (6.18) is a piecewise function of the boundary distance D. In the derivation at the end of this section, we show that there are three possible regimes for D as a result of the different local geometries around a node — i.e. how much of the connection discs around the node and its neighbours overlap with the domain (see also Figure 6.5). We find the following cases: (A) the *boundary regime* $D \leq r$ where a node as well as some of its

neighbours are influenced by the boundary, (B) the *near-boundary regime* $r < D \leq 2r$ where a node is not influenced directly by the boundary (i.e. no overlap between the connection disc and the boundary), but some of its neighbours are, and (C) the *bulk regime* $D > 2r$ where a node is at least two connection radii away from the boundary such that neither the node nor any of its neighbours are influenced by the boundary. Most importantly, in the bulk regime, expression (6.18) simplifies to $\mathbb{E}(\hat{p}(D)) = 0$, which implies zero node resistance curvature $\hat{p}$ in expectation in the bulk of RGGs on the disc. In the limit[7] of $r/R \to 0$, this bulk regime will take up all but a vanishing fraction of the domain and almost all nodes will thus have zero expected resistance curvature.

We remark that the derivation for expression (6.18) is independent of the specific shape of the domain $\mathbb{D}$ (i.e. it need not be a disc) and that for RGGs in higher dimensions, $A(t)$ is replaced by the volume of a higher-dimensional spherical cap at height t.

6.3.3.1 Derivation of the boundary function (6.18)

The model for the expected node resistance curvature is based on our heuristic for the effective resistance

$$\hat{\omega}_{ij} := \frac{1}{\lambda S(i)} + \frac{1}{\lambda S(j)}, \qquad \text{(Eq. (6.17))}$$

where $S(i)$ is the area of the intersection of the domain $\mathbb{D}$ and an r-radius disc around i. In general, this intersection can take many forms depending on the geometry of the boundary and the location of i with respect to the boundary. To overcome this complexity, we make two further assumptions. First, we assume that the curvature of the boundary (as a 1D curve in $\mathbb{R}^2$) is negligible with respect to the curvature of the r-disc around i; if the largest curvature of the boundary is $1/r_{\mathbb{D}}$, then we assume $r \ll r_{\mathbb{D}}$. In the case of a circular domain of radius R, as in the experiments, we have $r_{\mathbb{D}} = R$ such that the assumption is $r \ll R$. With this assumption, the boundary can be approximated locally as a straight line (i.e. when considering the neighbourhood around a node). Second, we assume that every point that lies at distance less than $2r$ from the boundary of $\mathbb{D}$ has a unique closest point on the boundary. This may be formalized using the concept of 'reach' [1, 89] by requiring that the boundary has reach at least $2r$. The reach of a domain $\mathbb{D} \subset \mathbb{R}^d$ is the largest number τ such that

[7] While the ratio r/R needs to be small to reduce the boundary effect, the expected number of nodes N needs to be large enough with respect to r such that the Poisson distribution (for the node degrees) concentrates around its mean and to be consistent with the conditions of von Luxburg et al. for our heuristic (6.20).

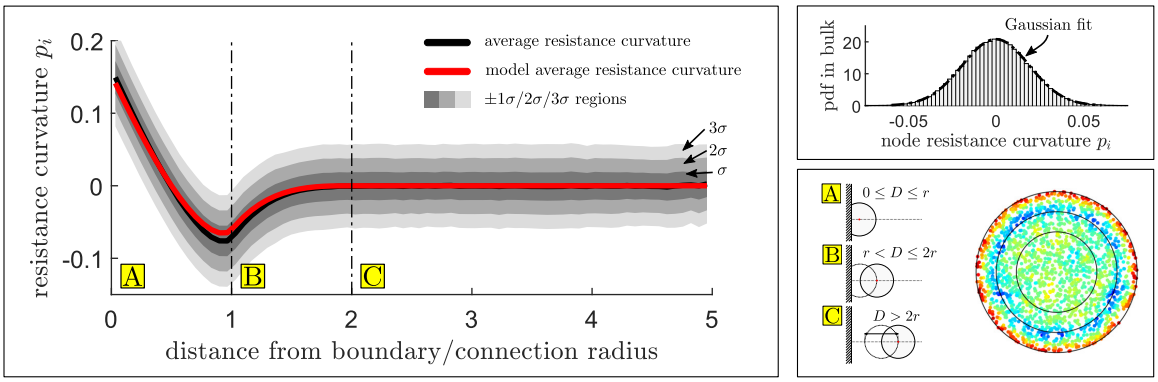


Figure 6.5: (Left panel) We sample 100 graphs from an RGG on a disc with $R/r = 5$ and $N = 5000$ and calculate the node resistance curvature p_i for all nodes. The plot shows the mean and variation around the mean of these curvatures (vertical axis) as a function of the distance from the boundary divided by the connection radius D_i/r (horizontal axis); the mean and variation are calculated by grouping the $100 \times N$ points according to their boundary distance in 75 equal-distance bins. The red line shows the heuristic $\mathbb{E}(\hat{p}(D))$ (expression (6.18)) which is in close correspondence with the black line, the experimental approximation for $\mathbb{E}(p(D))$. The largest deviation seems to occur when $D/r \approx 1$, but we have no explanation for this observation. The vertical black lines delineate the boundary (A), near-boundary (B) and bulk regimes. (Bottom right panel) illustration of the local geometries around a node that give rise to the regimes A–C. (Top right panel) The distribution of p_i for nodes in the bulk regime and the maximum likelihood Gaussian fit of this distribution; the good correspondence of this fit suggests that p_i could be a zero-mean Gaussian random variable in the bulk.

all points at distance less than τ from the boundary of $\mathbb{D}$ have a unique closest point on this boundary. For a fixed domain $\mathbb{D}$, these two assumptions will automatically be satisfied if we let r be small enough.

As Figure 6.6 illustrates, with these further assumptions, $S(i)$ is determined by the intersection of a disc and a half-plane (due to the straight boundary) and equal to

$$S(i) = \begin{cases} A(-D_i) & \text{if } 0 \le D_i \le r \\ \pi r^2 & \text{if } D_i > r \end{cases} \quad \text{or, in short } S(i) = A(\max\{-r, -D_i\}) \tag{6.19}$$

where $A(t)$ is the area of a circular segment[8] at height t. In other words, the function S only depends on the distance from the boundary, and only variations of node position in the direction towards or away from the boundary can change S; we may thus write $S(D_i)$. We note that $A(-r) = \pi r^2$ and $A(x) + A(-x) = \pi r^2$ for all $x \in [-r, r]$.

[8]This is equal to $A(t) = r^2 \arccos(t/r) - t\sqrt{r^2 - t^2}$ with $|dA(t)| = 2\sqrt{r^2 - t^2}$.

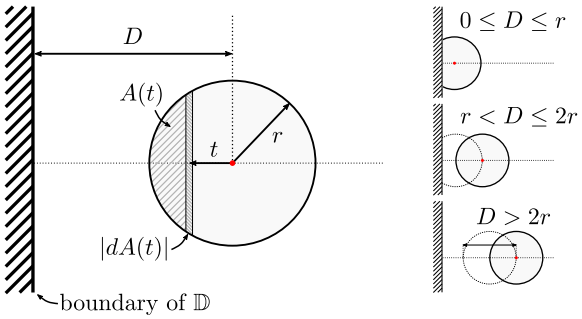


Figure 6.6: Illustration of the local neighbourhood around a node in an RGG. By assuming that the boundary curvature is negligible with respect to the curvature of the connection disc around each node, as $r \ll r_{\mathbb{D}}$, the boundary can be taken as a straight line. (Left figure) A node (red point) at distance $D > r$ from the boundary with the connection disc of radius r around this node, an example of a circle segment at height t with area $A(t)$ and $|dA(t)|$, which is the intersection between the connection disc and points at distance $D-t$ from the boundary. (Right figure) The three different node regimes determined by their distance D to the boundary.

Now, we fix a distance D and consider one sample graph G of the RGG on $\mathbb{D}$ and we add a node i at distance $D_i = D$ from the boundary. For this node, we can write

$$\begin{aligned}
\hat{p}_i &= 1 - \frac{1}{2}\sum_{j\sim i} \hat{\omega}_{ij} \quad \text{(we consider unweighted RGGs)} \\
&= 1 - \frac{1}{2}\sum_{j\sim i} \left(\frac{1}{\lambda S(D_i)} + \frac{1}{\lambda S(D_j)} \right) \quad \text{(by (6.17) for } \hat{\omega}) \\
&= 1 - \frac{d_i}{2\lambda S(D)} - \frac{1}{2}\sum_{j\sim i} \frac{1}{\lambda S(D_j)} \quad (D_i = D \text{ by construction}) \\
&= 1 - \frac{d_i}{2\lambda S(D)} - \frac{1}{2}\sum_{D'} \frac{\#\text{ neighbours of } i \text{ at distance } D' \text{ from boundary}}{\lambda S(D')},
\end{aligned}$$

where in the last expression D' ranges over $\{D_j : j \sim i\}$ with respect to the fixed node i. As noted, S only depends on variations in the direction orthogonal to the straight boundary. This allows us to express the position of the neighbours of i using the difference $D - D_j$; i.e. the distance between i and j measured in the direction orthogonal to the boundary. We can write

$$\hat{p}_i = 1 - \frac{d_i}{2\lambda S(D)} - \frac{1}{2}\sum_{t} \frac{\#\text{ neighbours of } i \text{ at distance } D-t \text{ from the boundary}}{\lambda S(D-t)} \tag{6.20}$$

where we sum over $\{t = D - D_j : j \sim i\}$. We now consider the expectation of $\hat{p}$ over a random realization of the RGG in which node i is included at the same point in the

domain, at distance D from the boundary. First, the expectation of the degree d_i is proportional to $S(D)$ as a property of the Poisson point process:

$$\mathbb{E}(d_i) = \lambda S(D).$$

Second, as shown in Figure 6.6 and as a property of the Poisson process, we find that the expected number of neighbours of node i at distance $D - t$ from the boundary is given by $\lambda |dA(t)|$. We can thus write the expectation of (6.20) as

$$\mathbb{E}(\hat{p}_i) = \frac{1}{2}\left(1 - \int_{-r}^{\min(r,D)} \frac{|dA(t)|}{S(D-t)}\right), \tag{6.21}$$

where the range for $t = D - D_j$ is $[-r, \min(r, D)]$. The lower bound of this range corresponds to the distance from i to the neighbours furthest away from the boundary. The upper bound of the range corresponds to the distance to the neighbours closest to the boundary. With expression (6.19) for the function S, we retrieve formula (6.18) for the expected resistance curvature in the section above. Equation (6.21) can be integrated numerically for different values of D, as shown in Figure 6.5.

We can further simplify expression (6.21) by distinguishing three different cases for D. Most importantly, this shows that nodes that are sufficiently far away from the boundary will have zero curvature $\hat{p}_i = 0$ in expectation and thus provides a (heuristic) explanation for the zero resistance curvature tendency in RGGs. The different regimes for D are as follows:

Bulk regime ($D > 2r$): In this case, we get $\min(r, D) = r$ and thus neighbours in the range $t \in [-r, r]$ around i. For these neighbours, we have $S(D - t) = \pi r^2$ for all t since $D - t > r$. Using $\int_{-r}^{r} |dA(t)| = A(-r) = \pi r^2$, we then find that

$$\mathbb{E}(\hat{p}_i) = \frac{1}{2}\left(1 - \frac{1}{\pi r^2}\int_{-r}^{r} |dA(t)|\right) = 0 \text{ for } D > 2r.$$

Near-boundary regime ($r < D \leq 2r$): In this case, we get $\min(r, D) = r$ and thus neighbours in the range $t \in [-r, r]$ around i. For these neighbours, we find (a) if $-r \leq t \leq D - r$ then $r \leq D - t \leq D + r$ and thus by (6.19) that $S(D-t) = \pi r^2$ and (b) if $D - r < t \leq r$ then $D - r \leq D - t < r$ and thus by (6.19) that $S(D-t) = A(t-D)$.

Expression (6.21) then yields

$$\begin{aligned}\mathbb{E}(\hat{p}_i) &= \frac{1}{2}\left(1 - \frac{1}{\pi r^2}\int_{-r}^{D-r}|dA(t)| - \int_{D-r}^{r}\frac{|dA(t)|}{A(t-D)}\right)\\ &= \frac{1}{2}\left(1 - \frac{A(-r)-A(D-r)}{\pi r^2} - \int_{D-r}^{r}\frac{|dA(t)|}{A(t-D)}\right)\\ &= \frac{A(D-r)}{2\pi r^2} - \frac{1}{2}\int_{D-r}^{r}\frac{|dA(t)|}{A(t-D)} \text{ for } r < D \leq 2r.\end{aligned}$$

Boundary regime ($0 \leq D \leq r$): In this case, we get $\min(r, D) = D$ and thus neighbours in the range $t \in [-r, D]$ around i. Similar to the near-boundary regime, we then find that

$$\begin{aligned}\mathbb{E}(\hat{p}_i) &= \frac{1}{2}\left(1 - \frac{1}{\pi r^2}\int_{-r}^{D-r}|dA(t)| - \int_{D-r}^{D}\frac{|dA(t)|}{A(t-D)}\right)\\ &= \frac{A(D-r)}{2\pi r^2} - \frac{1}{2}\int_{D-r}^{D}\frac{|dA(t)|}{A(t-D)} \text{ for } 0 \leq D \leq r.\end{aligned}$$

To summarize, we find the following piecewise expression:

$$\mathbb{E}(\hat{p}(D)) = \begin{cases} 0 & \text{if } D > 2r \\ \dfrac{A(D-r)}{2\pi r^2} - \dfrac{1}{2}\displaystyle\int_{D-r}^{r}\frac{|dA(t)|}{A(t-D)} & \text{if } r < D \leq 2r \\ \dfrac{A(D-r)}{2\pi r^2} - \dfrac{1}{2}\displaystyle\int_{D-r}^{D}\frac{|dA(t)|}{A(t-D)} & \text{if } 0 \leq D \leq r \end{cases}$$

The analysis of the resistance curvature in RGGs based on $\hat{\omega}$ not only explains why we may expect zero resistance curvature in the bulk of RRGs, but it also provides an explanation for the 'boundary effect' in Figures 6.4 and 6.5, where we found experimentally that the resistance curvature changes from zero to negative and then positive curvature when nearing the boundary. The derivation of $\mathbb{E}(\hat{p})$ shows how this boundary effect originates from the different possible local geometries around a node. This analysis could be relevant in other cases; for instance, in [263], where a similar boundary effect was described in the context of the magnitude of graphs (see also [154]), and [40], where a related boundary effect was used for boundary detection and related data-analysis tasks.

6.4 Discrete Ricci flow

Ricci flow is an important curvature-related concept in differential geometry [52]. Intuitively, the Ricci flow describes a curvature-dependent evolution of the metric of

a Riemannian manifold, where the metric decreases in directions of positive curvature while increasing in directions of negative curvature. The overall effect of this flow is that the curvature is 'smoothed out' and for this reason the Ricci flow is also thought of as a form of nonlinear heat equation for the metric. Due to its importance in the continuous setting, there have been several proposals to define a Ricci flow for discrete curvatures, with various applications: Jin et al. [133] defined a Ricci flow for meshes in computer graphics, Ni et al. [188] used a Ricci flow inspired evolution of networks for community detection and network alignment and Weber et al. considered Ricci flows related to the Forman–Ricci curvature [255]. Recently, Bai et al. studied a Ricci flow based on the Lin–Lu–Yau curvature [8] and Cushing et al. studied a Ricci flow based on Bakry–Émery curvature [64]. Here, we show that there is a natural Ricci flow associated to the resistance curvature with interesting features.

6.4.1 Resistance Ricci flow

Ollivier proposed in [189, Problem N] to look at the differential equation $\frac{d}{dt}d(i,j) = -\kappa_{ij}^{(\mathrm{OR})}d(i,j)$ as a discrete version of Ricci flow based on the Ollivier–Ricci curvature. If we introduce the link resistance curvature κ_{ij} and resistance distance $d(i,j) = \omega_{ij}$ in this expression, we find

$$\frac{d\omega_{ij}}{dt} = -2(p_i + p_j) \text{ for all } i \neq j \text{ in the same component} \tag{6.22}$$

which we call the *resistance Ricci flow*. Since ω is metric, we furthermore assume $d\omega_{ii}/dt = 0$. Ideally, equation (6.22) would describe the evolution of a resistance matrix $\Omega(t)$ for all $t \geq 0$ and starting from any initial resistance matrix $\Omega(0) = \Omega_0$. Unfortunately, however, $\Omega(t)$ is in general not guaranteed to be a resistance matrix since the examples below show that the flow can result in graphs with negative and diverging link weights in finite time. Nonetheless, we find that the resistance Ricci flow satisfies a number of interesting properties. As the flow (6.22) affects each connected component independently, we will assume G to be connected in the rest of Section 6.4.

6.4.1.1 Gradient flow

If X is a state space (e.g. a vector space) and $P : X \to \mathbb{R}$ a potential defined on this state space, then the evolution $dx/dt = -\nabla P$ for a flow $x(t) \in X$, assuming that it is well-defined, is called a *gradient flow*. This terminology reflects that the resulting flow follows the direction of steepest descent of the potential P, as given

by the negative gradient. As a first result, we find that the resistance Ricci flow is a gradient flow:

Proposition 6.17. *The resistance Ricci flow (6.22) is a gradient flow of the potential $2n\sigma^2$ defined on symmetric zero-diagonal matrices Ω:*

$$\frac{d\Omega}{dt} = -\nabla_\Omega(2n\sigma^2), \tag{6.23}$$

Proof. We follow the terminology and approach for matrix differentiation used in [198]. The gradient ∇_Ω of a real function f defined on resistance matrices is the matrix with entries

$$(\nabla_\Omega f)_{ij} := \frac{df}{d\omega_{ij}} = \operatorname{tr}\left(\left[\frac{\partial f}{\partial \Omega}\right]^T \frac{\partial \Omega}{\partial \omega_{ij}}\right),$$

where the second equality is the chain rule; see [198]. First, due to the structure of resistance matrices (zero diagonal and symmetry), we find

$$\frac{\partial \Omega}{\partial \omega_{ij}} = (1 - \mathbf{1}_{\{i=j\}})(\mathbf{e}_i\mathbf{e}_j^T + \mathbf{e}_j\mathbf{e}_i^T).$$

Second, following expression (6.6) for the resistance radius, we have $2n\sigma^2 = \frac{1}{2}\operatorname{tr}(\Omega Q \Omega)$. The partial derivative of the resistance radius is thus given by $\partial \operatorname{tr}(\frac{1}{2}\Omega Q\Omega)/\partial\Omega = \frac{1}{2}(Q\Omega + \Omega Q)$ and the gradient can be written as

$$\begin{aligned}
(\nabla_\Omega \operatorname{tr}(\tfrac{1}{2}\Omega Q\Omega))_{ij} &= \operatorname{tr}\left(\tfrac{1}{2}(1 - \mathbf{1}_{\{i=j\}})(Q\Omega + \Omega Q)(\mathbf{e}_i\mathbf{e}_j^T + \mathbf{e}_j\mathbf{e}_i^T)\right) \\
&= (1 - \mathbf{1}_{\{i=j\}})\left(\mathbf{e}_i^T Q\Omega\mathbf{e}_j + \mathbf{e}_j^T Q\Omega\mathbf{e}_i\right) \\
&= (1 - \mathbf{1}_{\{i=j\}})(2p_i + 2p_j - 4.\mathbf{1}_{\{i=j\}}) \quad \text{(by the } Q\Omega \text{ identity (3.7))} \\
&= \begin{cases} 2(p_i + p_j) & \text{if } i \neq j \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

This confirms that the gradient corresponds to minus the resistance Ricci flow (6.22), which completes the proof of Proposition 6.17. □

We remark that Proposition 6.17 should be seen as an 'instantaneous' result: at a time t where $\Omega(t)$ is a resistance matrix and thus σ^2 is well-defined, the resistance Ricci flow (6.22) is equal to the gradient flow of $2n$ times the resistance radius.

The gradient form (6.23) relates the resistance Ricci flow to the classical diffusion equation, which is a gradient flow of $\mathbf{x}(t)$ for the potential $\frac{1}{2}\operatorname{tr}(\mathbf{x}^T Q\mathbf{x})$. Furthermore, a general property of gradient dynamics is that the potential is a decreasing function of time and thus $d\sigma^2(t)/dt < 0$ for the resistance Ricci flow.

6.4.1.2 Flow of Laplacians

Expression (6.22) for the resistance Ricci flow in terms of changing effective resistances does not provide much insight into how the resistance Ricci flow affects the structure of a graph. Instead, the flow can be formulated in terms of Laplacian matrices:

Proposition 6.18. *The resistance Ricci flow* (6.22) *is equivalent to the following flow of Laplacian matrices:*

$$\frac{dQ}{dt} = 2Q\,\mathrm{diag}(\mathbf{p})Q, \text{ for some initial Laplacian } Q(0) = Q_0. \tag{6.24}$$

Proof. We start from expression (6.24) and show that this flow of Laplacian matrices is equivalent to (6.22) as a flow of effective resistances. First, the derivative of the inverse of a matrix satisfies $dA^{-1}/dt = -A^{-1}(dA/dt)A^{-1}$. Writing the Laplacian pseudoinverse as $Q^\dagger = (Q + \mathbf{u}\mathbf{u}^T/n)^{-1} - \mathbf{u}\mathbf{u}^T/n$ (see (2.6)) and considering the flow $dQ/dt = 2Q\,\mathrm{diag}(\mathbf{p})Q$, yields

$$\begin{aligned}\frac{dQ^\dagger}{dt} &= -Q^\dagger \frac{dQ}{dt} Q^\dagger \\ &= -2Q^\dagger Q\,\mathrm{diag}(\mathbf{p})QQ^\dagger \\ &= -2\left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right)\mathrm{diag}(\mathbf{p})\left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right).\end{aligned}$$

For the change of effective resistances, we then find

$$\frac{d\omega_{ij}}{dt} = (\mathbf{e}_i - \mathbf{e}_j)^T \frac{dQ^\dagger}{dt}(\mathbf{e}_i - \mathbf{e}_j) = -2(p_i + p_j) \text{ for all } i \neq j$$

which confirms that (6.24) and (6.22) are equivalent expressions for the resistance Ricci flow. □

Proposition 6.18 shows that the resistance Ricci flow — which we defined based on the proposal of Ollivier — corresponds to a quadratic flow of Laplacian matrices. By considering the off-diagonal entries of expression (6.24), we can write the evolution of individual link weights as:

$$\frac{dc_{ij}}{dt} = -2\sum_{\substack{k\sim i,j \\ k\neq i,j}} c_{ik}c_{kj}p_k + 2c_{ij}\left(k_i p_i + k_j p_j\right) \text{ for all } i \neq j, \tag{6.25}$$

where the equation for the diagonal (dk_i/dt) follows from the conserved zero row sum. The change of link weights in the resistance Ricci flow is thus caused by two types of processes: there is one term for each shared neighbour k of i and j, which can increase or decrease the link weight based on its resistance curvature p_k, and a second

term based on the degrees and curvatures of i and j. Other works such as [4] have arrived at similar dynamical rules when modeling dynamic network structures, with a balance between positive link-reinforcing terms and competing link-decaying terms. As a possible variation on the resistance Ricci flow, one might consider the dynamics $dQ/dt = 2Q\,\mathrm{diag}(\mathbf{q})Q$, where q is some function on the nodes. For instance, if q depends on a process taking place on the graph as in [4], this type of equation models the codependent evolution of structure and processes.

Finally, we remark that the Laplacian expression (6.24) shows that the resistance Ricci flow can result in links being added to or removed from a graph. This is in contrast with other proposals of discrete Ricci flow, which leave the graph structure intact and only evolve the link weights [8, 64, 255]. Already after a short time, the resistance Ricci flow can result, perhaps undesirably from the graph-theoretic perspective, in a fully connected graph, where only the link weight information is relevant.

To illustrate the resistance Ricci flow, we discuss two examples.

Example 6.19 (**node transitive graphs**)**.** The node resistance curvature in node transitive graphs is constant and equal to $p_i = 1/n$. Following the definition of the resistance Ricci flow, we then find that $d\omega_{ij}/dt = -2/n$ is equal for all pairs of nodes $i \neq j$ such that transitivity is conserved by the flow and thus $p_i(t) = 1/n$ as long as it is well-defined. As a consequence, the resistance Ricci flow for node transitive graphs simplifies to $dQ/dt = 2Q^2/n$. Diagonalizing this equation and solving the differential equation $dx/dt = 2x^2 \Rightarrow x(t) = x_0/(1 - 2tx_0)$ for each eigenvalue, we obtain the solution

$$Q(t) = \left[I - \tfrac{2}{n} t Q_0\right]^{-1} Q_0 \qquad \text{for } t < \frac{n}{2\mu_{\max}(Q_0)}.$$

This flow *diverges in finite time* for $t \to n/(2\mu_{\max})$. The resistance Ricci flow for node transitive graphs is formally related to a matrix flow studied in [172] as a model for structural balance in social networks, which considers the differential equation $dA/dt = A^2$ for a matrix A. The entries $(A)_{ij}$ of this matrix represent pairwise affinities or rivalries, depending on the sign, and finite-time divergence was interpreted in context of the model as the onset of a structurally balanced configuration of social relations [172].

Example 6.20 (**path graph**)**.** All nodes in a path graph have zero node resistance curvature, except for the end nodes, which have $p = 1/2$. From equation (6.25) for

the evolution of link weights, we then find that the only change occurs for the end links, whose link weight evolves according to

$$\frac{dc}{dt} = c^2 \Rightarrow c(t) = \frac{c_0}{1 - c_0 t} \quad \text{for } t < \frac{1}{c_0},$$

and diverges to $+\infty$ for $t \to 1/c_0$. While this result may seem pathological, we can interpret the 'infinite affinity' (which corresponds to $\omega \to 0$) as merging the end node and its neighbour into a single node (see Section 3.3.2). This analysis shows that the resistance Ricci flow of a path graph evolves by merging the end nodes of the path with their neighbours at $t = 1/c_0$, resulting in a path of decreasing length until just a single node remains.

An alternative 'normalized flow' can be defined based on the degree-normalized link resistance curvature as $d\omega_{ij}/dt = -2(p_i/k_i + p_j/k_j)$ or equivalently $dQ/dt = 2Q\,\mathrm{diag}(\mathbf{p}/\mathbf{k})Q$ where $\mathbf{p}/\mathbf{k}$ denotes entrywise division of vectors. This flow appears to share many properties with the unnormalized flow (6.22) with some apparent additional advantages that follow from the normalization.

6.4.2 Local resistance Ricci flow

To develop some intuition for the resistance Ricci flow, we define and analyze a simplified 'local' version of the normalized resistance Ricci flow; this local flow only affects the graph structure and weights around a given node x. As we show, the flow transfers weight from the direct links with its neighbours to the links in between its neighbours and results in an increase of the node resistance curvature p_x (or decreases for the reverse flow).

By linearity, the normalized resistance Ricci flow can be decomposed into a sum of n terms associated to individual nodes:

$$\frac{dQ}{dt} = \sum_{x \in \mathcal{N}} 2p_x Q \frac{\mathbf{e}_x \mathbf{e}_x^T}{k_x} Q\,. \tag{6.26}$$

We isolate one term from this decomposition and define the *local (resistance Ricci) flow* for node x as the following Laplacian flow:

$$\frac{dQ}{dt} = -Q \frac{\mathbf{e}_x \mathbf{e}_x^T}{k_x} Q \text{ with } Q(0) = Q_0. \tag{6.27}$$

In the rest of Section 6.4.2, we will consider this flow with respect to an arbitrary but fixed node x.

The local resistance Ricci flow can be solved exactly and its solution is given in terms of the Schur complement of the Laplacian. If we let matrix $Q(t)/\{x\}$ be the Schur complement of $Q(t)$ with respect to $\{x\}^c$ and with an additional zero row and column corresponding to index x, then we find the following solution

Proposition 6.21. *The solution of the local resistance Ricci flow* (6.27) *is*

$$Q(t) = e^{-t}Q_0 + (1 - e^{-t})Q_0/\{x\}, \text{ for all } t > 0. \tag{6.28}$$

Proof. We derive the evolution of the link weights by looking at the entries of the local flow (6.27). For neighbours $i \sim x$, we find

$$\frac{dc_{ix}}{dt} = -c_{ix} \Rightarrow c_{ix}(t) = e^{-t}c_{ix}(0) \text{ for all } i \sim x, \text{ and thus } k_x(t) = e^{-t}k_x(0).$$

For pairs of neighbours $i, j \sim x$, we find

$$\frac{dc_{ij}}{dt} = \frac{c_{ix}c_{jx}}{k_x} = e^{-t}\frac{c_{ix}(0)c_{jx}(0)}{k_x(0)} \Rightarrow c_{ij}(t) = c_{ij}(0) + (1 - e^{-t})\frac{c_{ix}(0)c_{jx}(0)}{k_x(0)}.$$

For pairs of non-neighbours, we simply have $dc_{ij}/dt = 0$ and thus $c_{ij}(t) = c_{ij}(0)$. These expressions for the link weights and thus the off-diagonal entries of $Q(t)$ agree entrywise with the proposed solution (6.28). Since zero row and column sums are conserved by definition of the local flow, this also uniquely determines the diagonal entries and completes the proof. □

The solution (6.28) confirms that the local resistance Ricci flow is indeed *local*: links and their weights only change in the direct neighbourhood of the node x.

While the appearance of the Schur complement might be surprising at first, it can be better understood by noting that the flow equations can be expressed as

$$\frac{dQ(t)}{dt} = Q(t)/\{x\} - Q(t) = Q_0/\{x\} - Q(t),$$

where $Q(t)/\{x\} = Q_0/\{x\}$ holds for all $t > 0$ because the effective resistances between non-neighbours of x are unchanged by the local flow (see Property 6.23 below). Since $Q_0/\{x\}$ is constant, the local flow is thus a (matrix) differential equation of the form $dA/dt = C - A$ which agrees with the solution in Proposition 6.21.

An important consequence of Proposition 6.21 is that the local flow is well-defined:

Property 6.22. *The solution $Q(t)$ to the local resistance Ricci flow is the Laplacian matrix of a connected graph for all finite times $t > 0$. At $t = \infty$, the matrix $Q_0/\{x\}$ is the Laplacian of a non-connected graph that has one isolated node.*

Proof. For finite times, the matrix $Q(t)$ described by (6.28) is given by the Laplacian matrix Q_0 with some link weights changed in the neighbourhood of x, but with no links removed and all weights positive; this is thus a Laplacian matrix. The result for $t = \infty$ follows immediately from Proposition 6.21. □

The local resistance Ricci flow thus gives rise to a graph $G(t)$, which is a connected weighted graph for finite times t, and in the limit of $t = \infty$ is equal to the x^c-Kron reduction of G_0 with an additional disconnected node, labeled x. With the transformation $\theta = e^{-t}$, which lies in $\theta \in [0, 1]$ for $t \geq 0$, it becomes more apparent that the local resistance Ricci flow graph $G(t)$ graph interpolates between G_0 and its x^c-Kron reduction as $G(\theta) = \theta G_0 + (1 - \theta)G_0/\{x\}$. This is illustrated in Figure 6.7 for a star graph.

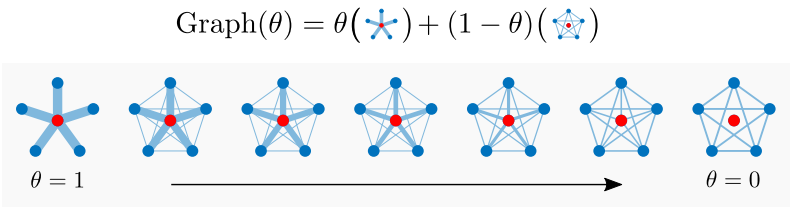


Figure 6.7: The local resistance Ricci flow on a star graph evolves from the star graph with the central node x (in red) connected to its neighbours at $\theta = 1$ to a graph with x disconnected and the x^c-Kron reduction of the star graph at $\theta = 0$. The width of a link is proportional to its weight.

The effective resistances of the graph $G(t)$ are found as follows:

Property 6.23. *The effective resistances in $G(t)$ are equal to*

$$\omega_{ix}(t) = \omega_{ix}(0) + (e^t - 1)k_x(0)^{-1} \text{ for all } i \sim x \quad \text{and} \quad \omega_{ij}(t) = \omega_{ij}(0) \text{ if } i, j \neq x.$$

Proof. Following the proof of Proposition 6.18, we know that

$$\frac{dQ^\dagger}{dt} = -Q^\dagger \frac{dQ}{dt} Q^\dagger = \left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right) \frac{\mathbf{e}_x \mathbf{e}_x^T}{k_x} \left(I - \frac{\mathbf{u}\mathbf{u}^T}{n}\right)$$

and thus that $d\omega_{ij}/dt = 0$ with $\omega_{ij}(t) = \omega_{ij}(0)$ whenever $i, j \neq x$, and that

$$\frac{d\omega_{ix}}{dt} = k_x^{-1} = e^t k_x(0)^{-1} \text{ for all } i \neq x.$$

This differential equation is solved by $\omega_{ix}(t) = \omega_{ix}(0) + (e^t - 1)/k_x(0)$, which completes the proof. □

Property 6.23 shows that the resistance distance between x and the rest of the nodes increases uniformly, while all other distances remain constant; this confirms

that $G(t)/\{x\} = G_0/\{x\}$. Geometrically, this uniform increase of the effective resistance between i and all other nodes means that the vertex $\varphi(\mathbf{e}_x)$ of the hyperacute simplex S moves perpendicularly away from the x^c-face of S.

Finally, the resistance curvatures in $G(t)$ can be solved as follows:

Property 6.24. *The resistance curvatures in $G(t)$ for finite times are equal to*

$$p_x(t) = e^{-t}p_x(0) + (1-e^{-t})\frac{1}{2} \quad \text{and} \quad p_i(t) = p_i(0) - \frac{c_{ix}(0)}{k_x(0)}\left[p_x(t) - p_x(0)\right].$$

Proof. Making use of the differential equations for the link weights and effective resistances in the proof of Proposition 6.21 and Property 6.23, the change of node resistance curvature of x equals

$$\frac{dp_x}{dt} = -\frac{1}{2}\sum_{i\sim x}\left(\frac{dc_{ix}}{dt}\omega_{ix} + c_{ix}\frac{d\omega_{ix}}{dt}\right) = -\frac{1}{2}\sum_{i\sim x}\left(-c_{ix}(t)\omega_{ix}(t) + \frac{c_{ix}(t)}{k_x(t)}\right) = \frac{1}{2} - p_x(t).$$

This differential equation is solved as $p_x(t) = e^{-t}p_x(0) + (1-e^{-t})\frac{1}{2}$ as required. For the neighbours of x, we find

$$\begin{aligned}
\frac{dp_i}{dt} &= -\frac{1}{2}\sum_{j:j\sim i}\left(\frac{dc_{ij}}{dt}\omega_{ij} + c_{ij}\frac{d\omega_{ij}}{dt}\right) \\
&= -\frac{1}{2}\left(\frac{dc_{ix}}{dt}\omega_{ix} + c_{ix}\frac{d\omega_{ix}}{dt}\right) - \frac{1}{2}\sum_{j\sim x}\left(\frac{dc_{ij}}{dt}\omega_{ij}\right) \quad \text{(flow is local)} \\
&= -\frac{1}{2}\left(-c_{ix}(t)\omega_{ix}(t) + \frac{c_{ix}(t)}{k_x(t)}\right) - \frac{1}{2}\sum_{j\sim x}\frac{c_{ix}(t)c_{jx}(t)}{k_x(t)}\omega_{ij}(t) \\
&= \frac{c_{ix}(t)}{k_x(t)}\left(\frac{1}{2}\left[k_x(t)\omega_{ix}(t) - \sum_{j\sim x}c_{jx}(t)\omega_{ij}(t)\right] - \frac{1}{2}\right) \\
&= \frac{c_{ix}(t)}{k_x(t)}\left(p_x(t) - \frac{1}{2}\right) \qquad \text{(see Eq. (6.5) for } p\text{)}.
\end{aligned}$$

Introducing the solutions for $c_{ix}(t), k_x(t)$ and $p_x(t)$, this yields

$$\begin{aligned}
\frac{dp_i}{dt} = e^{-t}\frac{c_{ix}(0)}{k_x(0)}\left(p_x(0) - \frac{1}{2}\right) \Rightarrow p_i(t) &= p_i(0) + \frac{c_{ix}(0)}{k_x(0)}\left(1-e^{-t}\right)\left(p_x(0) - \frac{1}{2}\right) \\
&= p_i(0) - \frac{c_{ix}(0)}{k_x(0)}[p_x(t) - p_x(0)]
\end{aligned}$$

as required. This completes the proof. □

The solutions for $p_x(t)$ and $p_i(t)$ respect the unit-sum property of resistance curvatures. We note that while $p_x(t) \to \frac{1}{2}$ for increasing times, at $t = \infty$, the resistance curvatures are determined by the graph $G_0/\{x\}$ in which x is an isolated node, with

$p_x = 1$. Consequently, there is a discontinuity in the resistance curvatures at $t = \infty$, which originates from node x being disconnected.

Following Proposition 6.21 and Properties 6.22–6.24 above, the local resistance Ricci flow describes a local change of a graph around a node x; the flow shifts link weight from the direct incident links on x to links between the neighbours of x, with an increase of the node resistance curvature of x as a result. Similarly, if we consider the *reverse local flow* by changing the sign of (6.27), we find a shift of link weight from links between neighbours to direct incident links with a decreasing node resistance curvature as a result. The decomposition (6.26) of the resistance Ricci flow in terms of local flows and the analysis of the local resistance Ricci flow above thus suggest the following intuition: the resistance Ricci flow changes the neighbourhood around each node x (by shifting link weight between its neighbours and incident links) such that p_x increases if it is negative and decreases if it is positive. These processes occur concurrently for all nodes, and the change of resistance curvature for a node x is thus also influenced by the local flow of all of its neighbouring nodes. A full understanding of the resistance Ricci flow would thus require a better understanding of the interplay between the different local processes.

While studied here in the context of resistance Ricci flow, the flow (6.27) and the resulting graphs $G(t)$ are interesting objects in their own right. Following Proposition 6.21, the local flow equation can be seen as a continuous version of Kron reduction with respect to a node x. Below are some suggestions (not meant to be precise) on how one may further study this continuous Kron reduction. For a graph G with Laplacian Q one might define an "infinitesimal dx^c-Kron reduction" $G' = G/d\{x\}$ with Laplacian $Q' = Q + dt(Q/\{x\} - Q)$; this infinitesimal Kron reduction would have many properties in common with the normal Kron reduction such as

$$c'_{ij} = c_{ij} + \frac{c_{ix}c_{jx}}{k_x}dt,$$

and the composition property would still hold. For matrices more generally, one could similarly consider an infinitesimal version of the Schur complement $A/d\mathcal{V}^c$ and the related flow $dA/dt = A(t)/\mathcal{V}^c - A(t)$ whenever these Schur complements are well defined.

6.5 Positive curvature

In differential geometry, spaces with positive curvature (bounded from below by some $K > 0$ at every point) enjoy many strong properties. Inspired by these results, this

section discusses some properties of graphs with positive node resistance curvature. More precisely, we consider the following class of graphs:

Definition 6.25. graph G is *positively* (resp. *nonnegatively*) *curved* if its node resistance curvature is positive (resp. nonnegative).

By extension, a resistance metric space is positively curved if the associated graph is positively curved. In line with the characterization of resistance matrices in Theorem 3.9, the resistance matrix of a positively curved graph can be characterized as a (real, symmetric) invertible matrix Ω that satisfies

$$(\mathbf{u}^T\Omega^{-1}\mathbf{u})(\Omega^{-1})_{ij} \geq (\Omega^{-1}\mathbf{u}\mathbf{u}^T\Omega^{-1})_{ij} > 0 \text{ for all } i \neq j.$$

Some examples of positively curved graphs, including node transitive graphs, were already discussed in Section 6.2. We give some further examples here, which highlight that positively curved graphs appear commonly.

Example 6.26 (**graphs with** $n \leq 3$)**.** Graphs with $n \leq 3$ nodes are always nonnegatively curved. The graph on one node has $p = 1$ by definition. A graph on two nodes has $p = 1/2$ for both nodes if it is connected and $p = 1$ otherwise. Disconnected graphs on three nodes consist of components of at most 2 nodes, which are positively curved. For connected graphs on three nodes, assume for contradiction that one of the nodes has $p_i < 0$. Then by the unit-sum property of p, we have that the other two must sum to a value strictly larger than 1. This is impossible because $p \leq 1/2$ for any connected node, so any graph on three nodes has nonnegative curvature.

Example 6.27 ($\mathcal{V}^\star$**-Kron reduction**)**.** By definition, the maximum variance support $\mathcal{V}^\star_\omega$ yields a graph $G^\star = G/\mathcal{V}^{\star c}$ which is positively curved; this graph can be found by repeatedly taking the Kron reduction with respect to the positively curved nodes in a graph until all nodes are positively curved (Theorem 5.23). For any graph G, this produces a canonical 'embedded' positively curved graph $G^\star$.

Example 6.28 (**Kron reduction of positively curved graph**)**.** Following Proposition 3.25 for the node resistance curvature and Kron reduction, we find:

Proposition 6.29. *The Kron reduction of a positively (nonnegatively) curved graph is positively (nonnegatively) curved.*

Proof. By the composition property of Kron reduction, it suffices to prove the result for elementary Kron reductions. Let G be a graph with positive curvature, x an arbitrary node and G' the x^c-Kron reduction of G. Following expression (3.15) for the resistance curvature in Kron reduction graphs, the resistance curvatures in G' satisfy

$$p'_i = p_i + \frac{c_{ix}}{k_x} p_x \geq p_i \text{ for all } i \neq x.$$

Thus if $p > 0$, then also $p' > 0$ (and similarly for nonnegative curvature), which completes the proof. □

In other words, for every positively curved graph, we get many more examples of positively curved graphs. For instance, any Kron reduction of a node transitive graph is positively curved.

Positively curved graphs (and the associated resistance metric spaces) can be seen as a further refinement in the hierarchy of metric spaces and associated geometric objects in Proposition 4.22, especially since they are again closed under Kron reduction. In the next subsections we discuss some strong properties of positively curved graphs.

6.5.1 Connectivity and toughness

Positive curvature has strong implications on the structural properties of a graph; we recall that a graph is biconnected if it contains no cut nodes. Following bounds on the resistance curvature, we find the following restriction:

Proposition 6.30. *Positively curved graphs are biconnected.*

Proof. Since $p_i \leq 1 - \beta(G_i \backslash \{i\})/2$ by Property 6.2 and $\beta(G_i \backslash \{i\}) \geq 2$ for cut nodes — as removing a cut node increases the number of connected components — cut nodes have nonpositive curvature. A positively curved graph can thus not contain any cut nodes. □

A biconnected graph can also be characterized as a graph in which 'removing one node results in one component'. The relevant notion that generalizes this characterization is *toughness*: a graph G is t-tough if removing any k nodes yields at most k/t components [14]. More precisely: a graph G is t-tough for some $t > 0$ if $\beta(G \backslash \mathcal{V}) \leq |\mathcal{V}|/t$ for any subset of nodes $\mathcal{V} \subseteq \mathcal{N}$ for which $\beta(G \backslash \mathcal{V}) \geq 2$; the complete graph is defined to be t-tough for all $t > 0$. For example, if a graph is 1-tough, this

means that removing any k nodes from the graph results in at most k components[9] — and in particular that the graph is biconnected. The *toughness* $\tau(G)$ of a graph is defined as the largest t such that G is t-tough[10]. Intuitively, the larger the toughness of a graph, the more difficult it is to disconnect this graph in many components, as larger sets of nodes need to be removed. Graph toughness was introduced by Chvátal in the context of Hamiltonian circuits and many results are known relating $\tau(G)$ to the existence of large cycles in G; see for instance [14].

In his study of simplices, where positive curvature corresponds to simplices whose circumcenter lies in the simplex interior, Fiedler found a result [96, Thm. 3.4.18] which implies the following relation between toughness and positively curved graphs:

Theorem 6.31. *Positively curved connected graphs are* 1*-tough.*

Proof. Let G be a connected positively curved graph and $R \subseteq \mathcal{N}$ a subset of nodes such that removing R results in at least $l \geq 2$ disconnected components[11]. Such a set always exists (e.g. let $R = \mathcal{N} \backslash \{i, j\}$ for some $i \not\sim j$) except if G is the complete graph, which is 1-tough. We will prove that $l \leq |R|$, which means that G is 1-tough.

(graph preprocessing step) Let $L_1, L_2, \ldots, L_l \subseteq \mathcal{N}$ denote the subsets of nodes that form connected components after removing the set of nodes R and let $\ell_1, \ell_2, \ldots, \ell_l$ be an arbitrary node in each of the subsets (i.e. $\ell_i \in L_i$ for $1 \leq i \leq l$). We consider the Kron reduction G' of G with respect to the remaining nodes in L_i; i.e.

$$G' := G / \bigcup_{i=1}^{l} (L_i \backslash \{\ell_i\}) \,.$$

The graph G' is a graph on $n' = |R| + l$ nodes and each of the nodes ℓ_i is connected only to the nodes in R; i.e. the node set $\{\ell_i\}_{i=1}^{l}$ is an independent set with no links between them. Indeed, by Proposition 3.19 on the structural implications of Kron reduction and the fact that any path from ℓ_i to ℓ_j for distinct $i, j \in [1, \ell]$ must pass through R since this is a cut node set, we know that $\ell_i \not\sim \ell_j$ in G'. The graph G' thus has two types of links: links within R and links between R and the nodes $\{\ell_i\}_{i=1}^{l}$. Figure 6.8 illustrates the construction of G' from G.

(proof of toughness) Second, by Proposition 6.29, we know that Kron reduction conserves positive curvature and thus that $p' > 0$ since $p > 0$. Summing the curvature

[9] We note that the t-tough property is different from k-connectivity, which says that removing any k nodes does not disconnect the graph.

[10] The toughness of the complete graph is defined to be $\tau = \infty$ and we note that this includes K_2, the path on two nodes and K_1, the graph on one node.

[11] This is a technical requirement for the definition of toughness.

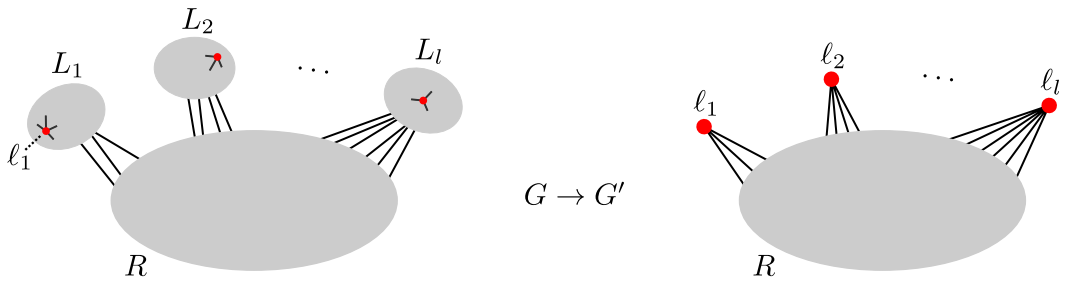


Figure 6.8: Sketch of the graph preprocessing step. The nodes of a graph are partitioned into a set R of nodes and sets $L_1, \ldots, L_l$ determined by the connected components of $G \backslash R$. An arbitrary node $\ell_1, \ldots, \ell_l$ is selected in each of these components. In the Kron reduction G', only the node ℓ_i is left in component L_i, resulting in the graph structure on the right: an arbitrary graph on the nodes R and the set of nodes ℓ_i connected to R. If G is positively curved, then so is G'.

of the nodes in R in G' and by the definition of node resistance curvature, we find

$$0 < \sum_{i \in R} p'_i = |R| - \frac{1}{2} \sum_{\substack{j \sim i \\ i \in R, j \notin R}} c'_{ij} \omega'_{ij} - \sum_{\substack{j \sim i \\ i,j \in R}} c'_{ij} \omega'_{ij}$$

where the relative resistances between nodes in R appear twice — once for each node at the end of the link. Grouping terms and invoking Foster's Theorem, we then find that

$$0 < |R| - \frac{1}{2} \sum_{j \sim i} c'_{ij} \omega'_{ij} - \frac{1}{2} \sum_{\substack{j \sim i \\ i,j \in R}} c'_{ij} \omega'_{ij} = |R| - \frac{1}{2} \left(|R| + l - 1 \right) - \frac{1}{2} \sum_{\substack{j \sim i \\ i,j \in R}} c'_{ij} \omega'_{ij} \leq \frac{1}{2} \left(|R| - l + 1 \right)$$

where the last inequality follows from nonnegativity of the relative resistance: $c'_{ij} \omega'_{ij} \geq 0$. This inequality implies $l < |R| + 1$ and since $l, |R|$ and 1 are integers and the inequality is strict, we thus have $l \leq |R|$ as required. Note that if we relax the connectedness constraint, we obtain the inequality $l < |R| + \beta$, which does not imply 1-toughness if $\beta > 1$. □

To our knowledge, no similar connectivity results exist for the known notions of discrete curvature. Fiedler suggested that Theorem 6.31 might be a complete characterization of positively curved graphs; i.e. in our terminology, that all 1-tough graphs have positive resistance curvature. However, comparing the computational complexity of deciding whether a graph is 1-tough and deciding whether it is positively curved gives the following result:

Proposition 6.32. *If $P \neq NP$, then not all 1-tough graphs are positively curved.*

Proof. In [15], it is shown that the decision problem "Is G 1-tough?" is NP-hard. Calculating the resistance curvature p for an n-node graph can be done in polynomial time, where in a straightforward implementation calculating the pseudoinverse Laplacian is the most time-consuming step with a worst-case time complexity of $O(n^3)$. Thus, the problem "Is G positively curved?" is in the computational complexity class P. As a result, if $P \neq NP$, both problems must be distinct and since all positively curved graphs are 1-tough by Theorem 6.31, not all 1-tough graphs can be positively curved. □

In other words, positively curved graphs are a strict subset of 1-tough graphs if $P \neq NP$. However, since it is unclear how different weightings of a graph affect the resistance curvature, the question "does every 1-tough graph have a weighting with positive curvature?" remains unsettled.

6.5.2 Submodular resistance radius

We recall the definition of the resistance radius as a function on subsets of the nodes:

$$\sigma^2(\mathcal{V}) := \frac{1}{2}\left(\mathbf{u}^T[\Omega_{\mathcal{V}\mathcal{V}}]^{-1}\mathbf{u}\right)^{-1} \text{ for all } \mathcal{V} \subseteq \mathcal{N} \qquad \text{(Eq. (3.19))}$$

with $\sigma^2(\{i\}) = \sigma^2(\emptyset) = 0$. The resistance radius function is nondecreasing $A \subseteq B \Rightarrow \sigma^2(A) \leq \sigma^2(B)$ and satisfies a partial inclusion–exclusion property: $\sigma^2(A \cup B) = \sigma^2(A) + \sigma^2(B) - \sigma^2(A \cap B)$ if $A \cap B$ is an $A - B$ cut set. For nonnegatively curved graphs, we find the following strengthening of the inclusion–exclusion property:

Theorem 6.33. *Let G be a graph with nonnegative resistance curvature. Then the resistance radius function satisfies*

$$\sigma^2(A \cup B) \leq \sigma^2(A) + \sigma^2(B) - \sigma^2(A \cap B) \tag{6.29}$$

for all $A, B \subseteq \mathcal{N}$ with nonempty intersection $A \cap B \neq \emptyset$.

The proof of Theorem 6.33 is given in Section 6.5.2.1. A set function $f : \{\mathcal{V} \subseteq \mathcal{N}\} \to \mathbb{R}$ that satisfies $f(A \cup B) \leq f(A) + f(B) - f(A \cap B)$ for all subsets A, B is called *submodular*; see for instance [163] for a survey of basic results and examples of submodular functions. One interpretation of submodularity is as a discrete analog of convexity with the important analogous property that certain discrete optimization problems with submodular objective functions can be solved efficiently. A second, equivalent characterization of submodular functions is

$$f(Y \cup \{z\}) - f(Y) \leq f(X \cup \{z\}) - f(X)$$

for all subsets $X \subseteq Y$ and elements z. This expression shows that submodular functions also embody the principle of diminishing returns: the increase in the set function when adding an element z to a small set (X) is larger than the increase when adding it to a larger set (Y).

Following Theorem 6.33, the resistance radius is submodular for graphs with nonnegative resistance curvature, up to the additional requirement that $A \cap B \neq \emptyset$. This requirement cannot be relaxed in general, as this would produce a contradiction if A and B each consist of a single node and $A \cap B = \emptyset$. We will call this *submodularity**; i.e. defined as submodularity where the subsets A and B must intersect.

A natural question following Theorem 6.33 is to try and characterize graphs with submodular* resistance radius. Is this a 'trivial' property that is satisfied by many (or even most/all) graphs, or does a submodular* resistance radius have any implications on the resistance curvature of the graph. In particular, since nonnegative curvature leads to submodularity*, it might be the case that submodularity* implies some lower bound on the curvature, or even nonnegative curvature. We find the following result:

Proposition 6.34. *Let G be a graph with submodular* resistance radius. Then the nodes have resistance curvature $p \geq -1$, and furthermore $p' \geq -1$ in any Kron reduction G'.*

Proof. Let G be a connected graph with submodular* resistance radius. If G has $n = 1$ node, then $p = 1 > -1$ and the proposition is satisfied; we thus assume that $n \geq 2$. Let z be any node and $x \sim z$ one of its neighbours and define $A = \mathcal{N}\backslash\{z\}$ and $B = \{x, z\}$, for which $A \cup B = \mathcal{N}$ and $A \cap B = \{x\}$. By submodularity* of G and $A \cap B \neq \emptyset$, we then know that

$$\sigma^2(\mathcal{N}) - \sigma^2(\mathcal{N}\backslash\{z\}) \leq \sigma^2(\{x, z\}) - \sigma^2(\{x\}) \Leftrightarrow \frac{p_z^2}{k_z} \leq \frac{\omega_{xz}}{4},$$

following expression (3.16) for the difference between resistance radii in the left-hand side and $\sigma^2(\{x\}) = 0$ and $\sigma^2(\{x, z\}) = \omega_{xz}/4$ in the right-hand side. If z has a single neighbour, we know that $p_z = 1/2 > -1$. If z has at least two neighbours, the above inequality holds for any neighbour and summing over all neighbours $x \sim z$ multiplied by $c_{xz} > 0$, we then find

$$p_z^2 = \sum_{x:x\sim z} c_{xz}\frac{p_z^2}{k_z} \leq \frac{1}{4}\sum_{x:x\sim z} c_{xz}\omega_{xz} = \frac{1}{2}(1 - p_z) \Leftrightarrow 2p_z^2 + p_z - 1 \leq 0,$$

where we used the definition of resistance curvature. This quadratic function in p_z is nonpositive if $p_z \in [-1, 1/2]$; this implies that $p_z \geq -1$, as required. The same

approach applies to any Kron reduction $G' = G/\mathcal{V}^c$ of G by letting $A = \mathcal{V}\backslash\{z\}$. This holds for any node z and thus completes the proof. □

This lower-bound on the resistance curvature is not trivial since nodes with $p < -1$ can exist in general graphs; for instance any node with $d_i \geq 5$ in a tree graph. In other words, a submodular* resistance radius puts some additional constraints on the graph structure. A full characterization of graphs with submodular* resistance radius would be interesting, and we propose the following problem:

Problem 6.35. Characterize graphs with submodular* resistance radius. Is submodularity* equivalent to nonnegative resistance curvature?

6.5.2.1 Proof of Theorem 6.33

To conclude, we give the proof of Theorem 6.33.

Proof of Theorem 6.33. We recall that we are trying to prove inequality (6.29) for the resistance radius σ^2. We will first show another inequality (equation (6.30) below) for σ^2 and then show that this implies the desired inequality and thus the theorem.

We start by proving the following inequality:

$$\sigma^2(V\cup\{x,y\})-\sigma^2(V\cup\{x\}) \leq \sigma^2(V\cup\{y\})-\sigma^2(V) \text{ for all } x \neq y \text{ and } \emptyset \neq V \subseteq \mathcal{N}\backslash\{x,y\}. \tag{6.30}$$

The graph must have $n \geq 3$ nodes for equation (6.30) to apply. Let x, y be two distinct nodes, $V = \mathcal{N}\backslash\{x,y\}$ and let $G' = G/\{x\}$ be the x^c-Kron reduction of G. By expression (3.16) for the relation between resistance radii and elementary Kron reductions, we then find that

$$\sigma^2(V\cup\{x,y\}) - \sigma^2(V\cup\{x\}) = \frac{p_y^2}{k_y} \quad \text{and} \quad \sigma^2(V\cup\{y\}) - \sigma^2(V) = \frac{{p'_y}^2}{k'_y}. \tag{6.31}$$

If $x \not\sim y$ in G, then we have $p'_y = p_y$ and $k'_y = k_y$ by properties of the Kron reduction and thus $p_y^2/k_y = p_y'^2/k'_y$. If $x \sim y$ in G, then we have $p'_y = p_y + c_{xy}p_x/k_x \geq p_y$ by expression (3.16) and nonnegative curvature of G, and $k'_y = k_y - c_{xy}^2/k_x < k_y$ and thus $p_y^2/k_y < p_y'^2/k'_y$. For any x, y, we thus have $p_y^2/k_y \leq p_y'^2/k'_y$ (with equality if and only if $x \not\sim y$), which by (6.31) confirms the inequality (6.30) for any triple V, x, y whose union is $\mathcal{N}$. For a general choice of V, x, y, we can take the corresponding $(V \cup \{x,y\})$-Kron reduction of the graph, which is nonnegatively curved by closure with respect to Kron reduction (Proposition 6.29). In this Kron reduction, the union

of V, x, y is the full node set and we may thus follow the same steps as above to confirm inequality (6.30) in general.

Next, we show that inequality (6.30) implies submodularity*

$$\sigma^2(A \cup B) \leq \sigma^2(A) + \sigma^2(B) - \sigma^2(A \cap B) \qquad \text{(Eq. (6.29))}$$

for all $A, B \subseteq \mathcal{N}$ with $A \cap B \neq 0$ as in the theorem. What follows is an adaptation of a proof given by Schrijver in [212, Thm. 44.1] for the equivalence between different notions of submodularity. Let $A, B \subseteq \mathcal{N}$ with $A \cap B \neq \emptyset$, and define the symmetric difference $A \Delta B := (A \cup B) \backslash (A \cap B)$, which are the nodes unique to A or B (analogous to an XOR relation). We will prove the implication (6.30)$\Rightarrow$(6.29) by induction on $|A \Delta B|$, the number of nodes in the symmetric difference.

(induction base cases) If $|A \Delta B| \leq 2$, then (i) either $A \subseteq B$ (or the other way around), in which case $A \cup B = A, A \cap B = B$ and thus equality holds for (6.29), or (ii) if they are not a subset, then let $V = A \cap B$ and $x = A \backslash B, y = B \backslash A$ and then (6.30) combined with $V = A \cap B \neq \emptyset$ implies (6.29).

These are the base cases for induction. Now, assuming inequality (6.29) holds for all pairs of subsets A, B with nonempty intersection and $|A \Delta B| \leq k$ where $k \geq 2$, we show that it also holds for $|A \Delta B| = k + 1$.

(induction step) Let $A, B \subseteq \mathcal{N}$ with $A \cap B \neq \emptyset$ (in particular this implies $A, B \neq \emptyset$) and with $|A \Delta B| = k + 1 \geq 3$. If $A \subseteq B$, we know that the inequality holds with equality, so assume that this is not the case. Then $A \backslash B$ has at least two elements (if not, interchange A and B) and let $a \in A \backslash B$ be one of the elements. First, we define $B' = (B \cup A) \backslash \{a\}$, for which we have[12] $|A \Delta B'| < |A \Delta B| = k + 1$ and thus $|A \Delta B'| \leq k$. By the induction assumption and the fact that $A \cap B' = A \backslash \{a\}$ is nonempty, we thus have the inequality

$$\sigma^2(A \cup B') - \sigma^2(A) \leq \sigma^2(B') - \sigma^2(A \cap B') \Leftrightarrow \sigma^2(A \cup B) - \sigma^2(A) \leq \sigma^2((A \cup B) \backslash \{a\}) - \sigma^2(A \backslash \{a\}).$$

Next, we define $A' = A \backslash \{a\}$, for which we have $|A' \Delta B| < |A \Delta B| = k + 1$ and thus $|A' \Delta B| \leq k$. By the induction assumption and the fact that $A' \cap B = A \cap B$ is nonempty, we then have the inequality

$$\sigma^2(A' \cup B) - \sigma^2(A') \leq \sigma^2(B) - \sigma^2(A' \cap B) \Leftrightarrow \sigma^2((A \cup B) \backslash \{a\}) - \sigma^2(A \backslash \{a\}) \leq \sigma^2(B) - \sigma^2(A \cap B).$$

[12] Note that this inequality requires that $A \backslash B$ has at least two elements since otherwise $B' = B$. This is why the induction base case is $k \leq 2$ and why we thus require (6.30) to settle the $k = 2$ case.

Combining the two inequalities above, we find

$$\sigma^2(A \cup B) - \sigma^2(A) \leq \sigma^2(B) - \sigma^2(A \cap B).$$

This concludes the induction step and since A, B were any two subsets of nodes with nonempty intersection, it confirms that (6.30)⇒(6.29) and completes the proof. □

Chapter 7

Conclusion and outlook

This thesis dealt with the effective resistance on graphs, both reviewing and unifying parts of the existing theory as well as developing several new results, especially related to geometry. If anything, we hope to have convinced the reader that the effective resistance is an interesting and rich concept in graph theory and beyond, with many more properties left to be discovered.

The first part of the thesis, Chapters 2–4, discussed the basic theory of effective resistances with a focus on matrix-theoretic properties. We discussed the Fiedler–Bapat identity as a central result and introduced Kron reductions — both of which find numerous applications throughout the thesis. We then derived the famous result that the effective resistance is a metric and classified the resistance distance as a strict negative type metric, and geometrically as the squared Euclidean distance between vertices of a hyperacute simplex. This led to Fiedler's graph–simplex correspondence, which gives a geometric perspective on much of the later results. In addition to providing a unified and self-contained treatment of the theory, we gave several new proofs and results.

In the second part, Chapters 5 and 6, we introduced new results and applications of the effective resistance. Broadly speaking, our new contributions are aligned with the following three central definitions:

- **Graph variance** $\operatorname{var} : \Delta \to \mathbb{R}$ (Definition 5.1). To address a methodological gap in the analysis of distributions and other data defined on the nodes of a graph, we propose generalized notions of variance and covariance based on distances between the nodes. This is not only a practical tool but leads to some interesting theoretical questions. We study the maximum-variance problem and characterize its solutions for different classes of metric spaces, where we find

that the resistance distance admits a greedy solution of the maximum-variance problem.

- **Resistance curvature $\mathbf{p}$** (Definition 6.1). As summarized in Section 6.2.3, the resistance curvature appears naturally in the different settings considered in this thesis: the Fiedler–Bapat identity, the graph–simplex correspondence and the maximum-variance problem. We derive some basic theoretical results about $\mathbf{p}$, we discuss a several alternative definitions and find an interpretation of $\mathbf{p}$ in the context of discrete curvature. This interpretation leads to a new class of graphs with strong properties, *positively curved graphs* characterized by $\mathbf{p} > 0$, and to two variants of *discrete Ricci flow* that describe a flow of graphs $G(t)$.

- **Resistance radius σ^2** (Definition 3.19). Similar to the resistance curvature, the resistance radius appears naturally in the different settings in this thesis. We derive some basic theoretical results of this graph invariant and find particularly interesting results when considering the resistance radius as a set function $\sigma^2 : \{\mathcal{V} \subseteq \mathcal{N}\} \to \mathbb{R}$ defined based on the Kron reductions of a graph.

A detailed summary of the contributions with references to specific equations, theorems and results was provided in Section 1.3 and the introductions of the respective chapters. As a quick 'one-line summary', many of the central concepts in this thesis make an appearance in Theorem 3.9:

$$\Omega \text{ is a resistance matrix } \Leftrightarrow (\mathbf{u}^T\Omega^{-1}\mathbf{u})(\Omega^{-1})_{ij} \geq (\Omega^{-1}\mathbf{u}\mathbf{u}^T\Omega^{-1})_{ij} \text{ for all } i \neq j.$$

This condition follows from the Fiedler–Bapat identity and reflects consistency with the definition of Laplacian matrices, with link weight positivity $c > 0$ front and center; we furthermore recognize the definitions of resistance curvature $\mathbf{p} \propto \Omega^{-1}\mathbf{u}$ and resistance radius $\sigma^2 \propto (\mathbf{u}^T\Omega^{-1}\mathbf{u})^{-1}$, and the condition is easily extended to positively curved graphs by requiring the righthandside to be positive.

Similarly, Kron reduction embodies many of the themes in this thesis. It first appears as the resistance-preserving graph operation, where closure of graphs with respect to Kron reduction reflects (via the graph–simplex correspondence) the geometric closure of hyperacute simplices with respect to taking faces. In Chapter 5 furthermore, the maximum variance is expressed in terms of Kron reductions of a graph and both the resistance curvature and resistance radius evolve simply with respect to Kron reduction.

The results in this thesis lead to many new questions. First, there are some 'loose ends' where further work is desirable: further studying node-merged graphs (Definition 3.27), the proposed covariance measure (Definition 5.7), the link resistance curvature κ (Definition 6.4) and the resistance Ricci flow (Eq. (6.22)). Second, while we describe connections to various applications and other theories — such as graph embeddings, Leinster's magnitude, triangulations and models of dynamical graphs — this is just the first step and much work can be done to further discover what we can learn from these connections. Finally, we list a number of follow-up questions and research directions that we find particularly interesting and/or promising:

- Given the importance of the Fiedler–Bapat identity in this thesis, it would be interesting to generalize the identity to other settings. For instance, for a symmetric matrix A with $\mathbf{u}^T A^\dagger \mathbf{u} \neq 0$, we may define an analogous identity based on the matrix $-2A^\dagger + 2\frac{A^\dagger \mathbf{u}\mathbf{u}^T A^\dagger}{\mathbf{u}^T A^\dagger \mathbf{u}}$. It would be interesting to see how much of the theory of resistance matrices (e.g. based on submatrices and Schur complements) can be reproduced in this setting.

- The relation between the resistance radius (and curvature) and Leinster's magnitude as described in Section 3.4 is compelling. For the theory of effective resistances, this relation suggests to study the matrix with entries $\exp(-\omega_{ij}t)$ and its inverse, while for the theory of magnitude, it suggests to study the matrix $-2Z^{-1} + 2\frac{Z^{-1}\mathbf{u}\mathbf{u}^T Z^{-1}}{\mathbf{u}^T Z^{-1}\mathbf{u}}$ and the associated Fiedler–Bapat identity for invertible similarity matrices Z. In particular, following the results on submatrices of the extended resistance matrix in Section 3.3.1, this may lead to new insights on the magnitude of submatrices of Z.

- It would be interesting to further develop some practical aspects of the graph (co)variance, for instance, trying to understand the influence of uncertainties in the estimation of distributions and network structure or developing better null models. More generally, we believe that it would be promising to develop a theory of distributions on the nodes of a graph. In Chapter 5, we studied the variance $\mathrm{var} : \Delta \to \mathbb{R}$, but many other characterizations of distributions might be useful. A starting point might be to define classes of distributions on the nodes, in analogy with well-studied classical distributions such as normal distributions.

- Theorem 6.33 says that nonnegatively curved graphs have submodular* resistance radius set functions, while Proposition 6.34 says that a submodular* resistance radius implies a nontrivial lower bound on the resistance curvature. Together, these results suggest to study whether the following equivalence holds:

$$\sigma^2 \text{ is submodular}^\star \quad \overset{?}{\Leftrightarrow} \quad \mathbf{p} \geq 0.$$

  This would be a new result in the study of discrete curvatures.

- The local resistance Ricci flow leads to a continuous variant of Kron reduction and, when applied to general matrices, a continuous variant of the Schur complement. Given the importance of Kron reduction and Schur complements in graph and matrix theory, a further characterization and study of these results seems a promising direction for future research.

- While finishing this thesis, we learned about the concept of Choquet capacity [51] and fuzzy measures [231] (also related to Lovász extensions [163]) and we believe that this could be a useful new perspective on the resistance radius set function. For instance, there is a relation to equilibrium equations (in the definition of capacities) and monotonicity and submodularity play important roles.

To conclude, we restate our hope that this thesis may be useful to anyone trying to better understand the effective resistance, and that it may enable and stimulate further research into this fascinating topic.

Appendix A

Linear algebra

A.1 Matrices and vectors

The basic objects in matrix theory are matrices. A real $k \times d$ matrix A is an array of real numbers $(A)_{ij}$ indexed by $i \in [1, k]$ and $j \in [1, d]$, for some $k, d \in \mathbb{N}$ as

$$A = \begin{pmatrix} (A)_{11} & \dots & (A)_{1d} \\ \vdots & \ddots & \vdots \\ (A)_{k1} & \dots & (A)_{kd} \end{pmatrix}.$$

We also write $A \in \mathbb{R}^{k \times d}$, where k and d are called the row and column dimensions. The transpose A^T of a matrix is defined by 'flipping' the matrix around its diagonal as $(A^T)_{ij} = (A)_{ji}$. A column vector is a $k \times 1$ matrix and is written in bold as $\mathbf{f} = (f_1, \dots, f_k)^T$ with its entries denoted by $f_i = (\mathbf{f})_i$. A row vector is a $1 \times d$ matrix written as $\mathbf{g}^T$. In particular, a row and column vector can correspond to the row or column of another matrix. A 1×1 matrix is called a scalar and corresponds to a single real number. The vector $\mathbf{e}_i$ is the i^{th} unit vector with $(\mathbf{e}_i)_j = 1$ if $i = j$ and zero otherwise and the vector $\mathbf{u} = (1, \dots, 1)^T$ is the all-one column vector.

Two matrices can be multiplied if their dimensions agree; a $k \times d$ matrix A and a $d \times \ell$ matrix B multiply to give the $k \times \ell$ matrix AB with entries

$$(AB)_{ij} = \sum_{r=1}^{d} (A)_{ir}(B)_{rj} \text{ for all } i \in [1, k] \text{ and } j \in [1, \ell].$$

In particular, a matrix multiplied with a column vector yields a column vector $A\mathbf{f}$ and a row vector multiplied with a matrix yields a row vector $\mathbf{g}^T A$. Finally, $\mathbf{g}^T A\mathbf{f}$ is called a quadratic form and yields a scalar. These are of course just the basics, and we refer for instance to [130] for more on matrix theory.

Matrix theory is a particular (practical) way to study linear algebra. A real $k \times d$ matrix A represents a linear map $T : V \to W$, where V is a d-dimensional vector space and W a k-dimensional vector space over the real numbers. Fixing a basis for V and W, the matrix entries of A are given by the map T as follows: for basis vectors $j \in V$ and $i \in W$, the matrix entry $(A)_{ij}$ is given by the i^{th} coordinate of the vector $T(j) \in W$. Similarly, a column vector corresponds to a vector $\mathbf{f} \in V$ with its entries $f_1, \ldots, f_d$ given by the coordinates in the fixed basis and the matrix–vector product $A\mathbf{f}$ corresponds to the vector $T(\mathbf{f})$ expressed in the fixed basis of W. While any matrix implicitly represents a linear map T, we will always work in a fixed basis and thus use the language of matrices instead of linear maps.

Throughout this thesis we mainly work with vectors in $\mathbb{R}^n$, which correspond to functions on the n nodes of a graph. Let $f : \mathcal{N} \to \mathbb{R}$ be a real function on the nodes; the collection of all such functions is a vector space with functions as vectors, elementwise addition and and real scalar multiplication. This is called the (real) *function space* on $\mathcal{N}$ and is sometimes denoted $C(\mathcal{N})$. The function space on $\mathcal{N}$ is isomorphic to $\mathbb{R}^n$ and thus we may write $\mathbf{f} \in \mathbb{R}^n$ to denote a function on the nodes. The nodes are fixed as basis vectors such that the entries of $\mathbf{f}$ as a column vector are given by $f_i = f(i)$ for a node $i \in \mathcal{N}$ and we have the $n \times 1$ column vector $\mathbf{f} = (f(1), \ldots, f(n))^T$. Similarly, an $n \times n$ matrix A can be seen as a linear map between node functions $A : C(\mathcal{N}) \to C(\mathcal{N})$, where an entry $(A)_{ij}$ corresponds to the node pair $i, j \in \mathcal{N}$.

We endow the function space $\mathbb{R}^n$ with the standard inner product. The inner product of two vectors $\mathbf{f}, \mathbf{g}$ is defined by[1] $\langle \mathbf{f}, \mathbf{g} \rangle = \sum_{i \in \mathcal{N}} f(i)g(i)$ and we note that this inner product is independent of the basis in which these vectors are expressed. From the inner product, we obtain a dual vector space[2] and since this is in bijection with $\mathbb{R}^n$, each vector $\mathbf{f}$ has a dual vector, which is denoted by $\mathbf{f}^T$. The matrix representations of a linear map T and its dual $T^\star$ are related by the matrix transpose as A for T and A^T for $T^\star$. The inner product between two vectors can be written as the matrix product $\langle \mathbf{f}, \mathbf{g} \rangle = \mathbf{f}^T\mathbf{g}$. The inner product also induces the following structures on $\mathbb{R}^n$:

- a norm: $\|\mathbf{f}\| = \sqrt{\mathbf{f}^T\mathbf{f}}$,

[1] Another inner product that is often used on $C(\mathcal{N})$ is $\langle \mathbf{f}, \mathbf{g} \rangle = \sum_{i=1}^n f(i)g(i)/k_i$ with degree k_i or alternatively with the combinatorial degree d_i.

[2] This is the vector space whose vectors are linear maps from $\mathbb{R}^n \to \mathbb{R}$ and with elementwise addition and real scalar multiplication; in other words, a dual vector $\mathbf{f}^\star$ acts on a vector $\mathbf{g}$ as $\mathbf{f}^\star(\mathbf{g}) \in \mathbb{R}$ and may for that reason be written as a row vector.

- a metric: $\|\mathbf{f} - \mathbf{g}\|$,
- angles: $\cos^{-1}(\mathbf{f}^T\mathbf{g}/\sqrt{\|\mathbf{f}\|.\|\mathbf{g}\|})$,
- orthogonality: $\mathbf{f} \perp \mathbf{g}$ if $\mathbf{f}^T\mathbf{g} = 0$, and correspondingly we may define orthogonal subspaces such as $\text{span}(\mathbf{f})^\perp$ in $\mathbb{R}^d$.

For further results on matrix theory, see for instance [130].

A.2 Eigenvalues of symmetric matrices

The eigenvalues and eigenvectors of a matrix A are solutions to the eigenvalue problem: if $A\mathbf{f} = \lambda\mathbf{f}$, then λ is called an eigenvalue of A with corresponding eigenvector $\mathbf{f}$. In general, the eigenvalues of a matrix may be complex, but in the case of symmetric matrices things simplify significantly [130, Thm. 2.5.6]:

Theorem A.1 (spectral theorem for symmetric matrices)**.** *Let A be a symmetric $n\times n$ matrix. The eigenvalues of A are real and there exists a basis of $\mathbb{R}^n$ of eigenvectors, where eigenvectors that correspond to different eigenvalues are orthogonal.*

In particular, each eigenvalue λ_k is associated with a subspace E_k of $\mathbb{R}^n$ and subspaces that belong to different eigenvalues are orthogonal; E_k is called the eigenspace of λ_k and the dimension of this space is the geometric multiplicity[3] of λ_k. The total sum of geometric multiplicities is n.

If $\lambda_1, \ldots, \lambda_n$ and $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are the eigenvalues and corresponding orthonormal eigenvectors of a symmetric matrix A, then we may write the following spectral decomposition:

$$A = \sum_{k=1}^{n} \lambda_k \mathbf{x}_k \mathbf{x}_k^T.$$

A symmetric matrix with nonnegative eigenvalues is called *positive semidefinite* and with positive eigenvalues it is called *positive definite*; these are sometimes abbreviated as PSD and PD, respectively. An equivalent characterization of positive (semi)definite matrices is the following [130, Def. 4.1.9 & Thm 4.1.10]:

Proposition A.2. *Let A be a symmetric $n\times n$ matrix. Then A is positive semidefinite if and only if $\mathbf{f}^T A\mathbf{f} \geq 0$ for all $\mathbf{f} \in \mathbb{R}^n$ and A is positive definite if and only if $\mathbf{f}^T A\mathbf{f} > 0$ for all nonzero $\mathbf{f} \in \mathbb{R}^n$.*

[3] For a symmetric matrix A, the geometric multiplicity (the dimension of the corresponding eigenspace) of an eigenvalue is equal to its algebraic multiplicity (the multiplicity of this eigenvalue as a root in the 'characteristic polynomial' $\det(A - \lambda I)$).

This also relates to the Courant–Fischer–Weyl Theorem [130, Thm. 4.2.6], from which we will use the following corollary [130, Cor. 4.2.12]:

Corollary A.3. *Let A be a symmetric matrix. If $\mathbf{f}^T A\mathbf{f} \geq 0$ for all $\mathbf{f}$ in a k-dimensional subspace, then A has at least k nonnegative eigenvalues. If the inequality is strict for nonzero $\mathbf{f}$, then A has at least k positive eigenvalues.*

Finally, positive semidefinite matrices are closely related to Gram matrices. Let $\mathbf{b}_1, \ldots, \mathbf{b}_n$ be n vectors in $\mathbb{R}^d$. Then the $n \times n$ matrix A with entries $(A)_{ij} = \mathbf{b}_i^T \mathbf{b}_j$ is called the Gram matrix of these vectors. Gram matrices are related to positive semidefinite matrices as follows [130, §7.2]:

Proposition A.4. *Let A be a matrix. Then A is a Gram matrix if and only if A is positive semidefinite.*

For a positive semidefinite matrix A, we can thus always write a decomposition $A = B^T B$. The vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ (i.e. the columns of B) can be constructed from the spectral decomposition of A as $(\mathbf{b}_i)_k = (\mathbf{x}_k)_i \sqrt{\lambda_k}$ for $1 \leq i \leq n$ and $1 \leq k \leq d$ with $\{\mathbf{x}_k, \lambda_k\}_{k=1}^d$ the eigenvalue-eigenvector pairs of A.

A.3 The Moore–Penrose pseudoinverse

We make use of the Moore–Penrose pseudoinverse of a matrix as defined in [196]. Let A be a real matrix. Then the Moore–Penrose pseudoinverse is defined by

$$AA^\dagger A = A, \quad A^\dagger AA^\dagger = A^\dagger \quad \text{and} \quad AA^\dagger, A^\dagger A \text{ are symmetric.}$$

The Moore–Penrose exists and is unique [196, Thm. 1] and corresponds to the usual inverse for invertible matrices. We note that from the definition it follows immediately that the Moore–Penrose pseudoinverse preserves symmetry and positive (semi)definiteness. Furthermore, it can be checked from the definition and uniqueness, that the Moore–Penrose pseudoinverse of a symmetric matrix A with eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding orthonormal eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_n$ has eigendecomposition

$$A^\dagger = \sum_{k:\lambda_k \neq 0} \lambda_k^{-1} \mathbf{x}_k \mathbf{x}_k^T.$$

We will make use the following property of the Moore–Penrose pseudoinverse:

Proposition A.5. *Let A be a matrix. Then $AA^\dagger$ is an orthogonal projection matrix onto* $\ker(A)^\perp$.

Proof. Let $P = A^\dagger A$. By definition of the Moore–Penrose pseudoinverse, we find $P^T = P$ and

$$AA^\dagger A = A \Rightarrow A^\dagger AA^\dagger A = A^\dagger A \Leftrightarrow P^2 = P.$$

This shows that P is an orthogonal projection matrix. Next, we find

$$\begin{aligned} P = A^\dagger A \quad &\text{such that} \quad A\mathbf{f} = 0 \Rightarrow P\mathbf{f} = 0, \\ AP = AA^\dagger A = A \quad &\text{such that} \quad P\mathbf{f} = 0 \Rightarrow A\mathbf{f} = 0, \end{aligned}$$

which implies $\ker(P) = \ker(A)$ and thus that the image of P is $\ker(A)^\perp$, as required. □

Bibliography

[1] E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, and L. Wasserman. Estimating the reach of a manifold. *Electronic Journal of Statistics*, 13(1):1359 – 1399, 2019.

[2] E. Abbe. Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research*, 18(177):1–86, 2018.

[3] E. Angriman, M. Predari, A. van der Grinten, and H. Meyerhenke. Approximation of the Diagonal of a Laplacian's Pseudoinverse for Complex Network Analysis. In F. Grandoni, G. Herman, and P. Sanders, editors, *28th Annual European Symposium on Algorithms (ESA 2020)*, volume 173 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 6:1–6:24, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[4] T. Aoki, K. Yawata, and T. Aoyagi. Self-organization of complex networks as a dynamical system. *Physical Review E*, 91(1):012908, 2015.

[5] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. *Journal of the ACM*, 56(2):1–37, 2009.

[6] G. I. Atabekov. *Linear Network Theory*. Pergamon Press, Oxford, UK, 1965.

[7] J. C. Baez and B. Fong. A compositional framework for passive linear networks. *Theory and Applications of Categories*, 33(38):1158–1222, 2018.

[8] S. Bai, Y. Lin, L. Lu, Z. Wang, and S.-T. Yau. Ollivier Ricci-flow on weighted graphs. 2021. arXiv:2010.01802 [math.DG].

[9] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de Probabilités XIX 1983/84*, Lecture Notes in Mathematics, pages 177–206. Springer, Berlin Heidelberg, Germany, 1985.

[10] R. B. Bapat. Resistance matrix of a weighted graph. *MATCH Communications in Mathematical and in Computer Chemistry*, 50:73–82, 2004.

[11] R. B. Bapat. *Graphs and Matrices*. Springer, London, UK, 2010.

[12] R. B. Bapat and S. K. Neogy. On a quadratic programming problem involving distances in trees. *Annals of Operations Research*, 243(1-2):365–373, 2014.

[13] R. B. Bapat and S. Sivasubramanian. Identities for minors of the Laplacian, resistance and distance matrices. *Linear Algebra and its Applications*, 435(6):1479–1489, 2011.

[14] D. Bauer, H. J. Broersma, and E. Schmeichel. Toughness in graphs — A survey. *Graphs and Combinatorics*, 22(10):1–35, 2006.

[15] D. Bauer, S. L. Hakimi, and E. Schmeichel. Recognizing tough graphs is NP-hard. *Discrete Applied Mathematics*, 28(3):191–195, 1990.

[16] F. Bauer, F. Chung, Y. Lin, and Y. Liu. Curvature aspects of graphs. *Proceedings of the American Mathematical Society*, 145(5):2033–2042, 2017.

[17] E. Bendito, A. Carmona, and A. M. Encinas. Potential theory for Schrödinger operators on finite networks. *Revista Matematica Iberoamericana*, 21(3):771–818, 2005.

[18] E. Bendito, A. Carmona, A. M. Encinas, and J. M. Gesto. A formula for the Kirchhoff index. *International Journal of Quantum Chemistry*, 108(6):1200–1206, 2008.

[19] E. Bendito, A. Carmona, A. M. Encinas, and J. M. Gesto. Characterization of symmetric M-matrices as resistive inverses. *Linear Algebra and its Applications*, 430(4):1336–1349, 2009.

[20] M. Berger. *Differential Geometry: Manifolds, Curves, and Surfaces*, volume 115 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, NY, US, 1988.

[21] A. Berman and R. J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences*. Society for Industrial and Applied Mathematics, Philadelphia, PA, US, 1994.

[22] N. Biggs. *Algebraic Graph Theory*, volume 67 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, UK, 2nd edition, 1974.

[23] N. Biggs. Algebraic potential theory on graphs. *Bulletin of the London Mathematical Society*, 29(6):641–682, 1997.

[24] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008–12, 2008.

[25] L. M. Blumenthal. *Theory and Applications of Distance Geometry*. Clarendon Press, Oxford, UK, 1953.

[26] A. I. Bobenko, J. M. Sullivan, P. Schröder, and G. M. Ziegler, editors. *Discrete Differential Geometry*, volume 38 of *Oberwolfach Seminars*. Birkhäuser, Basel, Switzerland, 2008.

[27] M. Boguñá, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M. A. Serrano. Network geometry. *Nature Reviews Physics*, 3(2):114–135, 2021.

[28] B. Bollobás. *Graph Theory*. Springer, New York, NY, US, 1979.

[29] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, UK, 2nd edition, 2001.

[30] A.-I. Bonciocat and K.-T. Sturm. Mass transportation and rough curvature bounds for discrete spaces. *Journal of Functional Analysis*, 256(9):2944–2966, 2009.

[31] S. P. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[32] J. Brandts, S. Korotov, and M. Křížek. *Simplicial Partitions with Applications to the Finite Element Method*. Springer Monographs in Mathematics. Springer, Cham, Switzerland, 2020.

[33] J. Brandts, S. Korotov, M. Křížek, and J. Šolc. On nonobtuse simplicial partitions. *SIAM Review*, 51(2):317–335, 2009.

[34] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, 30:107–117, 1998.

[35] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. 2022. arXiv:2104.13478 [CS.LG].

[36] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst. Geometric deep learning: Going beyond Euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[37] R. L. Brooks, C. A. B. Smith, A. H. Stone, and W. T. Tutte. The dissection of rectangles into squares. *Duke Mathematical Journal*, 7(1):312–340, 1940.

[38] A. E. Brouwer and W. H. Haemers. *Spectra of Graphs*. Springer, New York, NY, US, 2011.

[39] F. Buckley and F. Harary. *Distance in Graphs*. Addison-Wesley, Redwood City, CA, US, 1990.

[40] E. Bunch, J. Kline, D. Dickinson, S. Bhat, and G. Fung. Weighting vectors for machine learning: Numerical harmonic analysis applied to boundary detection. 2021. arXiv:2106.00827 [cs.LG].

[41] R. Burton and R. Pemantle. Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *The Annals of Probability*, 21(3):1329–1371, 1993.

[42] M. W. Cadotte, K. Carscadden, and N. Mirotchnick. Beyond species: Functional diversity and the maintenance of ecological processes and services. *Journal of Applied Ecology*, 48(5):1079–1087, 2011.

[43] H. Cai, V. W. Zheng, and K. Chang. A comprehensive survey of graph embedding: Problems, techniques, and applications. *IEEE Transactions on Knowledge & Data Engineering*, 30(09):1616–1637, 2018.

[44] G. Carlsson. Topology and data. *Bulletin (new series) of the American Mathematical Society*, 46(2):255–308, 2009.

[45] A. Cevallos, F. Eisenbrand, and R. Zenklusen. Max-Sum Diversity Via Convex Programming. In S. Fekete and A. Lubiw, editors, *32nd International Symposium on Computational Geometry (SoCG 2016)*, volume 51 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:14, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum für Informatik.

[46] A. K. Chandra, P. Raghavan, W. L. Ruzzo, R. Smolensky, and P. Tiwari. The electrical resistance of a graph captures its commute and cover times. *Computational Complexity*, 6(4):312–340, 1996.

[47] H.-C. Chang. *Tightening Curves and Graphs on Surfaces*. PhD thesis, University of Illinois at Urbana-Champaign, Illinois, 2018.

[48] G. Chartrand, D. Erwin, G. L. Johns, and P. Zhang. Boundary vertices in graphs. *Discrete Mathematics*, 263(1):25–34, 2003.

[49] P. Chebotarev. A class of graph-geodetic distances generalizing the shortest-path and the resistance distances. *Discrete Applied Mathematics*, 159(5):295–302, 2011.

[50] J. Cheeger, W. Müller, and R. Schrader. On the curvature of piecewise flat spaces. *Communications in Mathematical Physics*, 92(3):405–454, 1984.

[51] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1954.

[52] B. Chow and D. Knopf. *The Ricci Flow: An Introduction*, volume 110 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, US, 2004.

[53] T. Chu, Y. Gao, R. Peng, S. Sachdeva, S. Sawlani, and J. Wang. Graph sparsification, spectral sketches, and faster resistance computation, via short cycle decompositions. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 361–372, 2018.

[54] C. Chuan-Chong and K. Khee-Meng. *Principles and Techniques in Combinatorics*. World Scientific, Singapore, 1992.

[55] F. R. K. Chung. *Spectral Graph Theory*. American Mathematical Society, Providence, RI, US, 1997.

[56] F. R. K. Chung and S.-T. Yau. Logarithmic Harnack inequalities. *Mathematical Research Letters*, 3(6):793–812, 1996.

[57] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006.

[58] Y. Colin de Verdière. *Spectres de Graphes*. Société Mathématique de France, 1998.

[59] Y. Colin de Verdière, I. Gitler, and D. Vertigan. 1996.

[60] M. Coscia. Pearson correlations on complex networks. *Journal of Complex Networks*, 9(6), 2021. cnab036.

[61] H. S. M. Coxeter. The circumradius of the general simplex. *Mathematical Gazette*, 15(210):229–231, 1930.

[62] D. E. Crabtree and E. V. Haynsworth. An identity for the Schur complement of a matrix. *Proceedings of the American Mathematical Society*, 22(2):364–366, 1969.

[63] E. B. Curtis, D. Ingerman, and J. A. Morrow. Circular planar graphs and resistor networks. *Linear Algebra and its Applications*, 283(1):115–150, 1998.

[64] D. Cushing, S. Kamtue, S. Liu, F. Münch, N. Peyerimhoff, and H. B. Snodgrass. Bakry–Émery curvature sharpness and curvature flow in finite weighted graphs. I. Theory. 2022. arXiv:2204.10064 [math.CA].

[65] D. M. Cvetković, M. Doob, and H. Sachs. *Spectra of Graphs: Theory and Application*, volume 87 of *Pure and Applied Mathematics*. Academic Press, New York, NY, US, 1980.

[66] P. Dankelmann. Average distance in weighted graphs. *Discrete Mathematics*, 312(1):12–20, 2012.

[67] J.-C. Delvenne, S. N. Yaliraki, and M. Barahona. Stability of graph communities across time scales. *Proceedings of the National Academy of Sciences*, 107(29):12755–12760, 2010.

[68] B. Dennis, G. Patil, O. Rossi, S. Stehman, and C. Taillie. A bibliography of literature on ecological diversity and related methodology. *Ecological Diversity in Theory and Practice*, 1:319–353, 1979.

[69] M. Desbrun, A. N. Hirani, M. Leok, and J. E. Marsden. Discrete exterior calculus, 2005. arXiv:10.48550 [math.DG].

[70] K. Devriendt. Effective resistance is more than distance: Laplacians, simplices and the Schur complement. *Linear Algebra and its Applications*, 639:24–49, 2022.

[71] K. Devriendt and R. Lambiotte. Nonlinear network dynamics with consensus–dissensus bifurcation. *Journal of Nonlinear Science*, 31:18, 2021.

[72] K. Devriendt and R. Lambiotte. Discrete curvature on graphs from the effective resistance. *Journal of Physics: Complexity*, 3(2):025008, 2022.

[73] K. Devriendt, S. Martin-Gutierrez, and R. Lambiotte. Variance and covariance of distributions on graphs. *SIAM Review*, 64(2):343–359, 2022.

[74] K. Deweese. *Bridging the Theory–Practice Gap of Laplacian Linear Solvers*. PhD thesis, University of California Santa Barbara, California, 2018.

[75] M. Deza and E. Deza. *Encyclopedia of Distances*. Springer, Berlin Heidelberg, Germany, 4th edition, 2016.

[76] F. Di Giovanni, G. Luise, and M. M. Bronstein. Heterogeneous manifolds for curvature-aware graph embedding. In *ICLR 2022 Workshop on Geometrical and Topological Representation Learning*, 2022.

[77] R. Diestel. *Graph Theory*. Springer, Berlin Heidelberg, Germany, 2010.

[78] F. Dörfler and F. Bullo. Kron reduction of graphs with applications to electrical networks. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60(1):150–163, 2013.

[79] F. Dörfler, J. W. Simpson-Porco, and F. Bullo. Electrical networks and algebraic graph theory: Models, properties, and applications. *Proceedings of the IEEE*, 106(5):977–1005, 2018.

[80] P. G. Doyle. Electric currents in infinite networks. 2007. arXiv:0703899 [math.PR].

[81] P. G. Doyle and J. L. Snell. *Random Walks and Electric Networks*. Washington D.C., CO, US, 1984.

[82] P. Drineas and M. W. Mahoney. Effective resistances, statistical leverage, and applications to linear equation solving. 2010. arXiv:1005.3097 [cs.NA].

[83] D. Dubey and S. K. Neogy. On solving a non-convex quadratic programming problem involving resistance distances in graphs. *Annals of Operations Research*, 287(2):643–651, 2018.

[84] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[85] W. Ellens, F. M. Spieksma, P. Van Mieghem, A. Jamakovic, and R. E. Kooij. Effective graph resistance. *Linear Algebra and its Applications*, 435(10):2491–2506, 2011.

[86] G. V. Epifanov. Reduction of a plane graph to an edge by star–triangle transformations. *Doklady Akademii Nauk SSSR*, 166:19–22, 1966.

[87] M. Erbar and J. Maas. Ricci curvature of finite Markov chains via convexity of the entropy. *Archive for Rational Mechanics and Analysis*, 206(3):997–1038, 2012.

[88] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5(1):17–60, 1960.

[89] H. Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.

[90] M. Fiedler. Ueber die qualitative Lage des Mittelpunktes der umgeschriebenen Hyperkugel im n-Simplex. *Commentationes Mathematicae Universitatis Carolinae*, 2(1):3–51, 1961.

[91] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[92] M. Fiedler. Aggregation in graphs. In *Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. I*, volume 18 of *Colloq. Math. Soc. János Bolyai*, pages 315–330. North-Holland, Amsterdam-New York, 1978.

[93] M. Fiedler. A geometric approach to the Laplacian matrix of a graph. In R. A. Brualdi, S. Friedland, and V. Klee, editors, *Combinatorial and Graph-Theoretical Problems in Linear Algebra*, pages 73–98, New York, NY, 1993. Springer.

[94] M. Fiedler. Some characterizations of symmetric inverse M-matrices. *Linear Algebra and its Applications*, 275-276:179 – 187, 1998. Proceedings of the Sixth Conference of the International Linear Algebra Society.

[95] M. Fiedler. Ultrametric sets in Euclidean point spaces. *The Electronic Journal of Linear Algebra*, 3:23–30, 1998.

[96] M. Fiedler. *Matrices and Graphs in Geometry*, volume 139 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, UK, 2011.

[97] H. Flanders. Infinite networks: II — Resistance in an infinite grid. *Journal of Mathematical Analysis and Applications*, 40(1):30–35, 1972.

[98] B. Fong and D. I. Spivak. *An Invitation to Applied Category Theory: Seven Sketches in Compositionality*. Cambridge University Press, Cambridge, UK, 2019.

[99] R. Forman. Bochner's method for cell complexes and combinatorial Ricci curvature. *Discrete & Computational Geometry*, 29(3):323–374, 2003.

[100] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010.

[101] R. M. Foster. The average impedance of an electrical network. *Contributions to Applied Mechanics (Reissner Anniversary Volume)*, pages 333–340, 1949.

[102] F. Fouss, M. Saerens, and M. Shimbo. *Algorithms and Models for Network Data and Link Analysis*. Cambridge University Press, Cambridge, UK, 2016.

[103] K. Françoisse, I. Kivimäki, A. Mantrach, F. Rossi, and M. Saerens. A bag-of-paths framework for network data analysis. *Neural Networks*, 90:90–111, 2017.

[104] A. Ghosh, S. Boyd, and A. Saberi. Minimizing effective resistance of a graph. *SIAM Review*, 50(1):37–66, 2008.

[105] R. Ghrist. Barcodes: The persistent topology of data. *Bulletin (new series) of the American Mathematical Society*, 45(1):61–76, 2008.

[106] D. Glickenstein. Geometric triangulations and discrete Laplacians on manifolds. 2005. arXiv:10.48550 [math.MG].

[107] C. D. Godsil and G. Royle. *Algebraic Graph Theory*, volume 207 of *Graduate Texts in Mathematics*. Springer, New York, NY, US, 2001.

[108] J. Gorard. Some relativistic and gravitational properties of the Wolfram model. *Complex Systems*, 29(2):599–654, 2020.

[109] A. Gosztolai and A. Arnaudon. Unfolding the multiscale structure of networks with dynamical Ollivier–Ricci curvature. *Nature Communications*, 12(1):4561, 2021.

[110] J. Gower. Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra and its Applications*, 67:81–97, 1985.

[111] R. L. Graham and L. Lovász. Distance matrix polynomials of trees. *Advances in Mathematics*, 29(1):60–88, 1978.

[112] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer–Verlag Limited, 2008.

[113] T. N. E. Greville. Note on the generalized inverse of a matrix product. *SIAM Review*, 8(4):518–4, 1966.

[114] E. Grippo and E. A. Jonckheere. Effective resistance criterion for negative curvature: Application to congestion control. In *2016 IEEE Conference on Control Applications (CCA)*, pages 129–136, 2016.

[115] M. Gromov. Hyperbolic groups. In *Essays in Group Theory*, pages 75–263. Springer, New York, NY, 1987.

[116] A. Grover and J. Leskovec. Node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 855–864, New York, NY, 2016. Association for Computing Machinery.

[117] V. Gurvich. Metric and ultrametric spaces of resistances. *Discrete Applied Mathematics*, 158(14):1496–1505, 2010.

[118] I. Gutman and B. Mohar. The quasi-Wiener and the Kirchhoff indices coincide. *Journal of Chemical Information and Computer Sciences*, 36(5):982–985, 1996.

[119] A. D. Gvishiani and V. A. Gurvich. Metric and ultrametric spaces of resistances. *Russian Mathematical Surveys*, 42(6):235–236, 1987.

[120] W. H. Haemers. Interlacing eigenvalues and graphs. *Linear Algebra and its Applications*, 226-228:593–616, 1995.

[121] W. L. Hamilton. Graph representation learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 14(3):1–159, 2020.

[122] J. Hansen and R. Ghrist. Toward a spectral theory of cellular sheaves. *Journal of Applied and Computational Topology*, 3:315–358, 2019.

[123] F. Harary. *Graph Theory*. Addison-Wesley Series in Mathematics. Addison-Wesley, Reading, MA, US, 1969.

[124] A. Hatcher. *Algebraic Topology*. Cambridge University Press, Cambridge, UK, 2002.

[125] E. V. Haynsworth. Determination of the inertia of a partitioned Hermitian matrix. *Linear Algebra and its Applications*, 1(1):73–81, 1968.

[126] A. N. Hirani. *Discrete Exterior Calculus*. PhD thesis, California Institute of Technology Pasadena, California, 2003.

[127] P. Hjorth, P. Lisoněk, S. Markvorsen, and C. Thomassen. Finite metric spaces of strictly negative type. *Linear Algebra and its Applications*, 270(1-3):255–273, 1998.

[128] M. Hollander, E. Chicken, and D. A. Wolfe. *Nonparametric Statistical Methods*. Wiley Series in Probability and Statistics. Hoboken, NJ, US, 3rd edition, 2014.

[129] M. Homs-Dones, K. Devriendt, and R. Lambiotte. Nonlinear consensus on networks: Equilibria, effective resistance, and trees of motifs. *SIAM Journal on Applied Dynamical Systems*, 20(3):1544–1570, 2021.

[130] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, UK, 2nd edition, 2013.

[131] V. F. Ivanoff. The circumradius of a simplex. *Mathematics Magazine*, 43(2):71–72, 1970.

[132] R. K. Iyer, S. Jegelka, and J. A. Bilmes. Curvature and optimal algorithms for learning and minimizing submodular functions. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26, page 2742–2750. Curran Associates, Inc., 2013.

[133] M. Jin, J. Kim, F. Luo, and X. Gu. Discrete surface Ricci flow. *IEEE Transactions on Visualization and Computer Graphics*, 14(5):1030–1043, 2008.

[134] E. Jonckheere, E. Grippo, and R. Banirazi. Curvature, entropy, congestion management and the power grid. In *2019 IEEE Conference on Control Technology and Applications (CCTA)*, pages 535–542, 2019.

[135] J. Jost. *Riemannian Geometry and Geometric Analysis*. Universitext. Springer, Berlin Heidelberg, Germany, 1995.

[136] J. Jost and S. Liu. Ollivier's Ricci curvature, local clustering and curvature-dimension inequalities on graphs. *Discrete & Computational Geometry*, 51(2):300–322, 2013.

[137] J. Jost and F. Münch. Characterizations of Forman curvature. 2021. arXiv:2110.04554 [math.DG].

[138] S. Kamtue. Combinatorial, Bakry–Émery, Ollivier's Ricci curvature notions and their motivation from Riemannian geometry. 2018. arXiv:1803.08898 [math.CO].

[139] R. M. Karp. Reducibility among combinatorial problems. In R. E. Miller and J. W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. 1972.

[140] J. G. Kemeny. Generalization of a fundamental matrix. *Linear Algebra and its Applications*, 38:193–206, 1981.

[141] M. Kempton, G. Lippner, and F. Münch. Large-scale Ricci curvature on graphs. *Calculus of Variations and Partial Differential Equations*, 59(5):116, 2020.

[142] J. Kigami. *Analysis on Fractals*. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, UK, 2001.

[143] G. Kirchhoff. Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Annalen der Physik*, 148(12):497–508, 1847.

[144] G. Kirchhoff. On the solution of the equations obtained from the investigation of the linear distribution of galvanic currents. *IRE Transactions on Circuit Theory*, 5(1):4–7, 1958.

[145] S. Kirkland, M. Neumann, and B. L. Shader. Characteristic vertices of weighted trees via perron values. *Linear and Multilinear Algebra*, 40(4):311–325, 1996.

[146] D. J. Klein. Graph geometry, graph metrics and Wiener. *MATCH Communications in Mathematical and in Computer Chemistry*, 35(7):7–27, 1997.

[147] D. J. Klein. Distances and volumina for graphs. *Journal of Mathematical Chemistry*, 23:179–195, 1998.

[148] D. J. Klein and O. Ivanciuc. Graph cyclicity, excess conductance, and resistance deficit. *Journal of Mathematical Chemistry*, 30(3):271–287, 2001.

[149] D. J. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12(1):81–95, 1993.

[150] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi. The regularity lemma and its applications in graph theory. In *Lecture Notes in Computer Science*, pages 84–112, Berlin, 2002. Springer.

[151] W. Kook and K.-J. Lee. Simplicial networks and effective resistance. *Advances in Applied Mathematics*, 100(100):71–86, 2018.

[152] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Physical Review E*, 82:036106, 2010.

[153] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang. Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences*, 110(52):20935–20940, 2013.

[154] T. Leinster. The magnitude of metric spaces. 2010. arXiv:1012.5857 [math.MG].

[155] T. Leinster. *Basic Category Theory*. Cambridge University Press, Cambridge, UK, 2014.

[156] T. Leinster. *Entropy and Diversity: The Axiomatic Approach*. Cambridge University Press, Cambridge, UK, 2021.

[157] T. Leinster and C. A. Cobbold. Measuring diversity: The importance of species similarity. *Ecology*, 93(3):477–489, 2012.

[158] T. Leinster and M. W. Meckes. Maximizing diversity in biology and beyond. *Entropy*, 18(3):88, 2016.

[159] D. A. Levin, Y. Peres, E. L. Wilmer, J. Propp, and D. B. Wilson. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, US, 2nd edition, 2017.

[160] L. Liberti, C. Lavor, N. Maculan, and A. Mucherino. Euclidean distance geometry and applications. *SIAM Review*, 56(1):3–69, 2014.

[161] Y. Lin, L. Lu, and S.-T. Yau. Ricci curvature of graphs. *Tohoku Mathematical Journal*, 63(4):605 – 627, 2011.

[162] Y. Lin and S.-T. Yau. Ricci curvature and eigenvalue estimate on locally finite graphs. *Mathematical Research Letters*, 17:343–356, 03 2010.

[163] L. Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art*, pages 235–257. Springer, Berlin Heidelberg, Germany, 1983.

[164] L. Lovász. *Graphs and Geometry*, volume 65 of *Colloquium Publications*. American Mathematical Society, Providence, RI, US, 2019.

[165] L. Lovász. Random walks on graphs: A survey. In *Combinatorics, Paul Erdős is eighty*, pages 353–398, Budapest, Hungary, 1993. János Bolyai Math. Soc.

[166] I. Lukovits, S. Nikolić, and N. Trinajstić. Resistance distance in regular graphs. *International Journal of Quantum Chemistry*, 71(3):217–225, 1999.

[167] R. Lyons. Determinantal probability measures. *Publications Mathématiques de l'IHÉS*, 98:167–212, 2003.

[168] R. Lyons. Distance covariance in metric spaces. *The Annals of Probability*, 41(5):3284–3305, 2013.

[169] R. Lyons and Y. Peres. *Probability on Trees and Networks*, volume 42 of *Cambridge Series on Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, UK, 2016.

[170] S. Mac Lane. *Categories for the Working Mathematician*, volume 5 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, NY, US, 2nd edition, 1978.

[171] A. Madry, D. Straszak, and J. Tarnawski. Fast generation of random spanning trees and the effective resistance metric. In *Proceedings of the 2015 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2019–2036. 2015.

[172] S. A. Marvel, J. Kleinberg, R. D. Kleinberg, and S. H. Strogatz. Continuous-time model of structural balance. *Proceedings of the National Academy of Sciences*, 108(5):1771–1776, 2011.

[173] M. W. Meckes. Positive definite metric spaces. *Positivity*, 17(3):733–757, 2012.

[174] K. S. Menger. New foundation of Euclidean geometry. *American Journal of Mathematics*, 53:721, 1931.

[175] R. Merris. Laplacian matrices of graphs: A survey. *Linear Algebra and its Applications*, 197-198:143–176, 1994.

[176] C. Mihail and G. Milena. The Markov chain simulation method for generating connected power law random graphs. In *Proceedings of the Fifth Workshop on Algorithm Engineering and Experiments*, volume 111, page 16. SIAM, Philadelphia, 2003.

[177] B. Mohar, Y. Alavi, G. Chartrand, O. Oellermann, and A. Schwenk. The Laplacian spectrum of graphs. *Graph Theory, Combinatorics and Applications*, 2:871–898, 1991.

[178] B. Mohar and C. Thomassen. *Graphs on Surfaces*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, US, 2001.

[179] D. J. H. Moore and G. E. Sharpe. Metric transformation of an $(m+1)$-terminal resistive network into a hyperacute angled simplex Pm in Euclidean space Em. In *Proceedings of the Eleventh Midwest Symposium on Circuit Theory, Notre Dame, Indiana*, pages 184–192, 1968.

[180] Mosseri, R. and Sadoc, J. F. The Bethe lattice : A regular tiling of the hyperbolic plane. *Journal de Physique Lettres*, 43(8):249–252, 1982.

[181] F. Münch and R. K. Wojciechowski. Ollivier Ricci curvature for general graph Laplacians: Heat equation, Laplacian comparison, non-explosion and diameter bounds. *Advances in Mathematics*, 356:106759, 2019.

[182] K. A. Murgas, E. Saucan, and R. Sandhu. Quantifying cellular pluripotency and pathway robustness through Forman–Ricci curvature. 2021. bioRxiv doi: 10.1101/2021.10.03.462918.

[183] L. Najman and P. Romon, editors. *Modern Approaches to Discrete Curvature*, volume 2184 of *Lecture Notes in Mathematics*. Springer International Publishing, Berlin, Germany, 2017.

[184] M. E. J. Newman. *Networks*. Oxford University Press, Oxford, UK, 2nd edition, 2018.

[185] M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical Review E*, 69:026113, 2004.

[186] C.-C. Ni, Y.-Y. Lin, J. Gao, and X. D. Gu. Network alignment by discrete Ollivier–Ricci flow. In *Graph Drawing and Network Visualization*, pages 447–462. Springer International Publishing, Berlin, 2018.

[187] C.-C. Ni, Y.-Y. Lin, J. Gao, X. D. Gu, and E. Saucan. Ricci curvature of the Internet topology. In *2015 IEEE Conference on Computer Communications (INFOCOM)*, pages 2758–2766. IEEE, 2015.

[188] C.-C. Ni, Y.-Y. Lin, F. Luo, and J. Gao. Community detection on networks with Ricci flow. *Scientific Reports*, 9(1):9984, 2019.

[189] Y. Ollivier. A survey of Ricci curvature for metric spaces and Markov chains. In *Probabilistic approach to geometry*, pages 343–381. Mathematical Society of Japan, 2010.

[190] J. Pach. The beginnings of geometric graph theory. In L. Lovász, I. Z. Ruzsa, and V. T. Sós, editors, *Erdős Centennial*, pages 465–484. Springer, Berlin Heidelberg, Germany, 2013.

[191] J. L. Palacios. Closed-form formulas for Kirchhoff index. *International Journal of Quantum Chemistry*, 81(2):135–140, 2001.

[192] J. L. Palacios. Resistance distance in graphs and random walks. *International Journal of Quantum Chemistry*, 81(1):29–33, 2001.

[193] T. P. Peixoto. Parsimonious module inference in large networks. *Physical Review Letters*, 110:148701, 2013.

[194] P. Peng, D. Lopatta, Y. Yoshida, and G. Goranci. Local algorithms for estimating effective resistance. In *Proceedings of the 27th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, page 1329–1338, New York, NY, 2021. Association for Computing Machinery.

[195] M. Penrose. *Random Geometric Graphs*. Oxford University Press, Oxford, UK, 2003.

[196] R. Penrose. A generalized inverse for matrices. *Mathematical Proceedings of the Cambridge Philosophical Society*, 51(3):406–413, 1955.

[197] B. Perozzi, R. Al-Rfou, and S. Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, page 701–710, New York, NY, USA, 2014. Association for Computing Machinery.

[198] K. B. Petersen and M. S. Pedersen. The matrix cookbook, 2008. http://www2.imm.dtu.dk/pubdb/p.php?3274, accessed on 29/05/22.

[199] N. Peyerimhoff. Curvature notions on graphs Leeds summer school, 18-19 July 2019. https://www.maths.dur.ac.uk/users/norbert.peyerimhoff/peyerimhoff-lecture-notes.pdf, accessed on 29/05/22.

[200] M. A. Porter, J.-P. Onnela, and P. J. Mucha. Communities in networks. *Notices of the American Mathematical Society*, 56(9):1082–1097, 2009.

[201] H. Qiu and E. R. Hancock. Clustering and embedding using commute times. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):1873–1890, 2007.

[202] P. Ramaciotti Morales, R. Lamarche-Perrin, R. Fournier-S'niehotta, R. Poulain, L. Tabourier, and F. Tarissan. Measuring diversity in heterogeneous information networks. *Theoretical Computer Science*, 859:80–115, 2021.

[203] G. Ranjan and Z.-L. Zhang. Geometry of complex networks and topological centrality. *Physica A*, 392(17):3833–3845, 2013.

[204] C. R. Rao. Diversity: Its measurement, decomposition, apportionment and analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 44(1):1–22, 1982.

[205] T. Regge. General relativity without coordinates. *Il Nuovo Cimento*, 19(3):558–571, 1961.

[206] V. Salnikov, D. Cassese, R. Lambiotte, and N. S. Jones. Co-occurrence simplicial complexes in mathematics: Identifying the holes of knowledge. 3(1):37, 2018.

[207] A. Samal, R. P. Sreejith, J. Gu, S. Liu, E. Saucan, and J. Jost. Comparative analysis of two discretizations of Ricci curvature for complex networks. *Scientific Reports*, 8(1):8650, 2018.

[208] E. Saucan, A. Samal, and J. Jost. A simple differential geometry for complex networks. *Network Science*, 9(S1):106–133, 2021.

[209] D. Schleuter, M. Daufresne, F. Massol, and C. Argillier. A user's guide to functional diversity indices. *Ecological Monographs*, 80(3):469–484, 2010.

[210] I. J. Schoenberg. Remarks to Maurice Fréchet's article "Sur la définition axiomatique d'une classe d'espace distanciés vectoriellement applicable sur l'espace de Hilbert". *Annals of Mathematics*, 36(3):724–732, 1935.

[211] I. J. Schoenberg. Metric spaces and positive definite functions. *Transactions of the American Mathematical Society*, 44(3):522–536, 1938.

[212] A. Schrijver. *Combinatorial Optimization : Polyhedra and Efficiency*, volume 24 of *Algorithms and Combinatorics*. Springer, Berlin, Germany, 2003.

[213] G. E. Sharpe. On the $(m+1)$-terminal resistive-network problem. *Proceedings of the Institution of Electrical Engineers*, 116:503–509(6), 1969.

[214] G. E. Sharpe and D. J. H. Moore. Transfer impedances and the no-amplification property of resistive networks. *Proceedings of the IEEE*, 56(6):1116–1117, 1968.

[215] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Processing Magazine*, 30(3):83–98, 2013.

[216] J. Sia, E. Jonckheere, and P. Bogdan. Ollivier–Ricci curvature-based method to community detection in complex networks. *Scientific Reports*, 9(1):9800, 2019.

[217] J. Sigbeku, E. Saucan, and A. Monod. Curved Markov chain Monte Carlo for network learning. In *Complex Networks & Their Applications X*, pages 461–473, Berlin, 2022. Springer International Publishing.

[218] D. A. Spielman. Graphs, vectors, and matrices. *Bulletin (new series) of the American Mathematical Society*, 54(1):45–61, 2017.

[219] D. A. Spielman and N. Srivastava. Graph sparsification by effective resistances. *SIAM Journal on Computing*, 40(6):1913–1926, 2011.

[220] D. A. Spielman and S.-H. Teng. Nearly-linear time algorithms for graph partitioning, graph sparsification, and solving linear systems. In *Proceedings of the Thirty-Sixth Annual ACM Symposium on Theory of Computing; 13-16 June 2004*, STOC '04, page 81–90, New York, NY, 2004. Association for Computing Machinery.

[221] D. A. Spielman and S.-H. Teng. Spectral sparsification of graphs. *SIAM Journal on Computing*, 40(4):981–1025, 2011.

[222] D. A. Spielman and S.-H. Teng. Nearly linear time algorithms for preconditioning and solving symmetric, diagonally dominant linear systems. *SIAM Journal on Matrix Analysis and Applications*, 35(3):835–885, 2014.

[223] R. P. Sreejith, K. Mohanraj, J. Jost, E. Saucan, and A. Samal. Forman curvature for complex networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(6):063206, 2016.

[224] C. St. J. A. Nash-Williams. Random walk and electric currents in networks. *Mathematical Proceedings of the Cambridge Philosophical Society*, 55(2):181–194, 1959.

[225] S. Steinerberger. The boundary of a graph and its isoperimetric inequality, 2022. arXiv:2201.03489 [stat.CO].

[226] S. Steinerberger. Curvature on graphs via equilibrium measures, 2022. arXiv:2202.01658 [stat.CO].

[227] A. Stirling. A general framework for analysing diversity in science, technology and society. *Journal of the Royal Society, Interface*, 4(15):707, 2007.

[228] G. E. Subak-Sharpe. On the structure of well-conditioned positive resistance networks. In *Proceedings of the 33rd Midwest Symposium on Circuits and Systems*, pages 293–296. IEEE, 1990.

[229] G. E. Subak-Sharpe. On a remarkable measurement property of resistive networks. In *1992 IEEE International Symposium on Circuits and Systems*, volume 4, pages 1751–1755. IEEE, 1992.

[230] G. E. Subak-Sharpe and G. P. H. Styan. A necessary condition for the realization of a resistive n-port based on network size and on the concept of weighted terminal valency. In *2000 IEEE International Symposium on Circuits and Systems (ISCAS)*, volume 1, pages 487–490. IEEE, 2000.

[231] M. Sugeno. *Theory of Fuzzy Integrals and its Applications*. PhD thesis, Tokyo Institute of Technology, Tokyo, 1974.

[232] L. Sun, W. Wang, J. Zhou, and C. Bu. Some results on resistance distances and resistance matrices. *Linear and Multilinear Algebra*, 63(3):523–533, 2015.

[233] G. J. Székely, M. L. Rizzo, and N. K. Bakirov. Measuring and testing dependence by correlation of distances. *The Annals of Statistics*, 35(6):2769 – 2794, 2007.

[234] P. Tee and C. A. Trugenberger. Enhanced Forman curvature and its relation to Ollivier curvature. *Europhysics Letters*, 133(6):60006, 2021.

[235] S.-H. Teng. The Laplacian paradigm: Emerging algorithms for massive graphs. In J. Kratochvíl, A. Li, J. Fiala, and P. Kolman, editors, *Theory and Applications of Models of Computation*, pages 2–14, Berlin Heidelberg, Germany, 2010. Springer.

[236] P. Tetali. Random walks and the effective resistance of networks. *Journal of Theoretical Probability*, 4(1):101–109, 1991.

[237] C. Thomassen. Resistances and currents in infinite electrical networks. *Journal of Combinatorial Theory. Series B*, 49(1):87–102, 1990.

[238] J. Topping, F. Di Giovanni, B. P. Chamberlain, X. Dong, and M. M. Bronstein. Understanding over-squashing and bottlenecks on graphs via curvature. In *International Conference on Learning Representations*, 2022.

[239] L. Torres, K. S. Chan, and T. Eliassi-Rad. GLEE: Geometric Laplacian Eigenmap Embedding. *Journal of Complex Networks*, 8(2), 2020. cnaa007.

[240] N. Trinajstić. *Chemical Graph Theory*. CRC Press, Boca Raton, FL, US, 2nd edition, 1992.

[241] C. A. Trugenberger. Combinatorial quantum gravity: Geometry from random bits. *The Journal of High Energy Physics*, 2017(9):1–8, 2017.

[242] W. T. Tutte. *Graph Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, UK, 2001.

[243] M. Tyloo, T. Coletta, and P. Jacquod. Robustness of synchrony in complex networks and generalized Kirchhoff indices. *Physical Review Letters*, 120:084101, 2018.

[244] M. Tyloo and P. Jacquod. Global robustness versus local vulnerabilities in complex synchronous networks. *Physical Review E*, 100:032303, 2019.

[245] R. van der Hofstad. *Random Graphs and Complex Networks*, volume 43 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, UK, 2017.

[246] P. van der Hoorn, W. J. Cunningham, G. Lippner, C. Trugenberger, and D. Krioukov. Ollivier–Ricci curvature convergence in random geometric graphs. *Physical Review Research*, 3(1):013211, 2021.

[247] P. Van Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, Cambridge, UK, 2010.

[248] P. Van Mieghem, K. Devriendt, and H. Cetinay. Pseudoinverse of the Laplacian and best spreader node in a network. *Physical Review E*, 96:032311, 2017.

[249] N. Vishnoi. $Lx = b$ Laplacian solvers and their algorithmic applications. *Foundations and Trends in Theoretical Computer Science*, 8:1–141, 2012.

[250] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.

[251] U. von Luxburg, A. Radl, and M. Hein. Getting lost in space: Large sample analysis of the resistance distance. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems*, volume 23, pages 2622–2630. Curran Associates, Inc., 2010.

[252] U. von Luxburg, A. Radl, and M. Hein. Hitting and commute times in large random neighborhood graphs. *Journal of Machine Learning Research*, 15(52):1751–1798, 2014.

[253] D. Wang, P. Cui, and W. Zhu. Structural deep network embedding. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, page 1225–1234, New York, NY, 2016. Association for Computing Machinery.

[254] X. Wang, J. L. A. Dubbeldam, and P. Van Mieghem. Kemeny's constant and the effective graph resistance. *Linear Algebra and its Applications*, 535:231 – 244, 2017.

[255] M. Weber, J. Jost, and E. Saucan. Forman–Ricci flow for change detection in large dynamic data sets. *Axioms*, 5(4):26, 2016.

[256] M. Weber, J. Stelzer, E. Saucan, A. Naitsat, G. Lohmann, and J. Jost. Curvature-based methods for brain network analysis. 2019. arXiv:1707.00180 [q-bio.NC].

[257] M. Weber, M. Zaheer, A. S. Rawat, A. K. Menon, and S. Kumar. Robust large-margin learning in hyperbolic space. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 17863–17873. Curran Associates, Inc., 2020.

[258] E. Wehrhahn. On the derivation of the TWNs from the Kirchhoff's laws. In *1993 IEEE International Symposium on Circuits and Systems*, pages 2553–2555, 1993.

[259] T. Weihrauch. A characterization of effective resistance metrics. *Potential Analysis*, 51(3):437–467, 2018.

[260] L. Weinberg. *Network Analysis and Synthesis*. McGraw-Hill Electrical and Electronic Engineering Series. McGraw-Hill, New York, NY, US, 1962.

[261] G. Westendorp. Webpage: A formula for the N-circumsphere of an N-simplex., 2013. https://westy31.home.xs4all.nl/Circumsphere/ncircumsphere.htm, accessed on 29/05/2022.

[262] H. Wiener. Structural determination of paraffin boiling points. *Journal of the American Chemical Society*, 69(1):17–20, 1947.

[263] S. Willerton. Heuristic and computer calculations for the magnitude of metric spaces. 2009. arXiv:0910.5500 [math.MG].

[264] W. Xiao and I. Gutman. Resistance distance and Laplacian spectrum. *Theoretical Chemistry Accounts*, 110(4):284–289, 2003.

[265] G. F. Young, L. Scardovi, and N. E. Leonard. Robustness of noisy consensus dynamics with directed communication. In *Proceedings of the 2010 American Control Conference*, pages 6312–6317. IEEE, 2010.

[266] G. F. Young, L. Scardovi, and N. E. Leonard. A new notion of effective resistance for directed graphs — part I: Definition and properties. *IEEE Transactions on Automatic Control*, 61(7):1727–1736, 2016.

[267] F. Zhang, editor. *The Schur Complement and its Applications*. Springer-Verlag, New York, NY, US, 2005.

[268] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun. Graph neural networks: A review of methods and applications. *AI Open*, 1:57–81, 2020.

[269] J. Zhou, Z. Wang, and C. Bu. On the resistance matrix of a graph. *The Electronic Journal of Combinatorics*, 23(1), 2016. P1.41.

[270] T. Zhou, J. Ren, M. Medo, and Y.-C. Zhang. Bipartite network projection and personal recommendation. *Physical Review E*, 76:046115, 2007.