



**Master Thesis**

University of Cologne – Institute for Theoretical Physics

Forschungszentrum Jülich – IEK-STE

---

# Optimal geometries of transport networks with linear congestion

---

*Author:*
Matthias Dahlmanns

*First Supervisor:*
Prof. Dr. Dirk Witthaut

*Second Supervisor:*
Prof. Dr. Andreas Schadschneider

submitted:

September 28, 2020

Modern society relies on the proper functioning of various networks to provide elementary goods such as water, communication, or electric power. Urban transport systems are becoming increasingly important, in particular due to the ongoing effort to reduce mobility-based carbon emissions and mitigate climate change. On the other hand, an increasing share of the global population lives in cities which results in an rapidly growing usage of public transport systems. As a result, public transport systems are increasingly prone to congestion, raising the question of how to optimise the system to cope with this challenge.

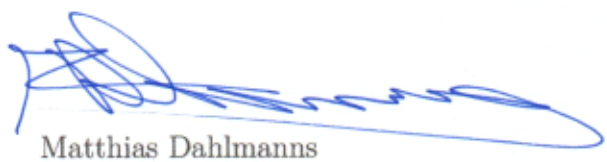
This thesis investigates an elementary model for transportation networks, where flows are determined to minimise the total travel time. Including a linear congestion term, the model interpolates between ordinary shortest-paths flows and current flow following Kirchhoff's and Ohm's laws.

This framework is applied to study three fundamental problems in transportation networks. First, we investigate fundamental flow pattern such as the rerouting of flows after the failure of a single link. Then we turn to optimal networks and investigate how optimal structures change when fundamental model parameters are varied, showing that the formation of loops is actually discontinuous. Finally, we turn to more coarse scales and investigate the optimal structure of multimodal traffic networks.

# Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Jülich, 28.09.2020



Matthias Dahlmanns

# Contents

# I. Introduction

Modern society relies on the proper functioning of various networks to provide elementary goods such as water, communication, or electric power. In the last decade, network science emerged as a scientific discipline connecting concepts and methods from statistical physics, nonlinear dynamics and applied mathematics with applications in biological, technical and socioeconomic systems.

The efficiency and robustness of networks is crucial for a wide range of systems such as in biological organisms, the vascular system supplies cells with water and nutrients [1, 2, 3] and the global spreading of epidemics is governed by network structures [4] as well as food webs [5]. An increasing number of human-build networks distribute everyday goods in our modern life: Electrical power grids [6], hydraulic networks [7] as well as as well as natural gas networks [8] and other pipeline grids are crucial for our standard of living. Street traffic [9, 10] and public transport networks [11, 12, 13] significantly improve the mobility of people. The world wide web transfers informations via the internet [14]. Efficient logistic networks are crucial to run globalised ware circulations [15]. Furthermore, we find efficient logistic networks build by other species [16]. In economy, there are also different types of networks to study the interactions of customers and suppliers [17, 18] as well as financial markets [19].

Urban transport systems are becoming increasingly important, in particular due to the ongoing effort to reduce mobility-based carbon emissions and mitigate climate change. On the other hand, an increasing share of the global population lives in cities which results in an rapidly growing usage of public transport systems. As a result, urban transport systems are increasingly prone to congestion, raising the question of how to optimise the system to cope with this challenge. This Master thesis focuses on the robustness and optimal geometry of transportation networks subject to congestion. The network model is inspired by urban transport networks such as subway or tram systems, but the analysis can also be transferred to other systems such as street traffic networks.

A variety of models is developed to describe congested flows in specific situations [9, 20]. In this thesis, however, we use a fundamental model that includes congestion as a linear coupling of the local flow in the network to the cost function. This one-dimensional model allows us to study optimal geometries for very general scenarios.

It turns out that the flow equations in this model can be formulated as a *quadratic*

*programme* (QP), that is a mathematically well studied class of optimisation problems [21]. Furthermore, we find when driving our model parameter to its extreme values, that it interpolates between two well known cases. We observe on one hand ordinary shortest-path flows [22] and on the other hand current flows determined by Ohm's and Kirchhoff's law as we find it in electrical current networks [23, 24].

First, we formulate the linear congestion model in chapter II, before we study fundamental flow patterns which are generated by this model in chapter III. In particular we exploit the rerouting pattern on a regular lattice. Subsequently, we discuss optimal network geometries considering economic constrains for two scenarios: In chapter IV, we study the optimal topology of a generic network with fluctuating flows and pursue the question when it is optimal to build a tree-like network and when to include a loop. In chapter V, we discuss the distance-to-centre problem for radial symmetric cites and discuss the impact of congestion and human behaviour on the optimal geometry. In chapter VI, we summarise the results and give an outlook.

# II. A Linear Congestion Model

In this section we propose a model that adopts congestion by adding a linear term to the flow cost function, which we denote by $\tau$. The cost refers to the total travel time which the travellers inside the system try to minimise. In this thesis, we use the expressions of cost and travel time synonymously. Mathematically seen, this model interpolates between ordinary shortest-path flows and current flows determined by Ohm's and Kirchhoff's laws.

We establish the model and the basic notation in section II.1 and discuss both the shortest path limit and the Ohmic limit in sections II.2 and II.3, respectively. We then generalise the formalism to finite congestion using KKT conditions in section II.4. It turns out that this method is not suitable to find explicit solutions for the flows, but we can use numerical solver by formulating our problem as a *Quadratic Programme* (QP) that is a standard class of mathematical optimisation problems [21]. It is, however, numerically expensive to solve a QP which high degrees of freedom. Hence, we introduce a method to reduce the numerical complexity of the solution in II.5 and discuss further possibilities to simplify the model using additional assumptions in II.6. We use the library *ALGLIB* [25] to write a general solver for our model for which we analyse the scaling of the runtime in section II.7.

## II.1. Formulation of the model

We consider a transportation network, consisting of a set of $N$ nodes which are connected by a set of $L$ edges on which travellers can move between nodes. We assume the network to be directed, that means for an edge $l$ which starts at node $n$ and ends at node $m$, denoted by $(n, m)$ or $l$ interchangeably, the flow $F_l$ along the edge can only flow from $n$ to $m$. This implies in particular $F_l \geq 0$.

We note that most connections in real transportation networks are bidirectional: Streets, subway or rail networks usually allow for transport in both directions. Such a connection is represented by two edges, one for each direction.

We assume that the time $\tau_l$ which a traveller needs to travel along an edge $l$ increases

linearly with the local flow $F_l$, i.e.

$$\tau_l(F_l) = t_l(1 + \eta F_l)\,. \tag{II-1}$$

We introduce the congestion parameter $\eta$ which is the inverse of the flow intensity which doubles the link travel time compared to the free-flow travel time $t_l$. For the sake of simplicity, we assume that $\eta$ is equal for all edges[1]. This user link cost function $\tau_l(F_l)$ is a special case of the Bureau of Public Road (BPR) functions [26], which read in their general form

$$\tau_l(F_l) = t_l^0\left[1 + \eta\left(\frac{F_l}{u_l}\right)^{\beta}\right]$$

and are widely used in the literature [27, 28]. For our model, we use $\beta = 1$ and $u_l = 1$. The velocity $v_l$ on a link of length $d_l$ then reads

$$v_l(F_l) = \frac{d_l}{\tau_l} = \frac{v_{l,0}}{1 + \eta\, F_l}\,, \tag{II-2}$$

with $v_{l,0} = d_l/t_l$ being the velocity on the empty link.

Traffic flow is fundamentally different from the flow of physical quantities such as water in a hydraulic network or current in an electric network. In the latter case, we only have to keep track of the net flow over a link or the net in or outflow at a node in the network. In the case of traffic networks, we must take into account that travellers have individual starting points and destinations and cannot be set off against each other. Assume that travellers can enter or leave the network only at nodes, which correspond e.g. in public transport systems to stations. We introduce a set of flow layers such that each layer $(n)$ represents travellers with a node $i$ as destination. For each flow layer $(n)$ we associate an inflow vector $P_j^{(n)}$ with one component for each node $j$ in the network counting the number of travellers entering the network at $j$ with destination $i$. Travellers leaving the network are counted negatively. This implies $P_j^{(n)} \geq 0 \;\forall\; j \neq i$ and

$$P_i^{(n)} = -\sum_{j\neq i} P_j^{(n)}\,.$$

Note that one can also swap the signs in $P_j^{(n)}$ to obtain a flow layer with a single starting point instead of a single destination. We must only ensure that for each

[1] Note that one might introduce multiple types of edges with different $\eta$ for each type, e.g. when considering a subway network with different lines where it takes a certain time to change from one line to another because of waiting time or the need of walking from one platform to another. In this case, edges can either represent subway tracks or a connection from one line to another within a station. The impact of congestion might differ between subway tracks and the connection time between lines within a station. In this thesis, however, we will only consider networks where $\eta$ is equal for all edges

layer $(n)$ there is exactly one component $P_i^{(n)}$ with opposite sign from all the other components of the same layer – given the layer has any inflow at all. Introducing the incidence matrix

$$I_{il} = \begin{cases} +1 & , \text{ if } l \text{ ends in } i \\ -1 & , \text{ if } l \text{ starts in } i \\ 0 & , \text{ else} \end{cases}, \tag{II-3}$$

which encodes the topology and orientation of the edges $l$ in the network [29] allows us to write down the flow conservation law as

$$\sum_l I_{il} F_l^{(n)} = P_i^{(n)}, \tag{II-4}$$

where $F_l^{(n)}$ is the flow on edge $l$ caused by the flow layer $(n)$, i.e. the total flow $F_l$ on the edge is given by

$$F_l = \sum_n F_l^{(n)}.$$

The continuity equation (II-4) is also referred to as Kirchhoff's current law (KCL) in circuit theory. In this thesis, we will restrict ourselves to traffic flow problems with either a single starting point or a single destination. This restriction vastly simplifies the analysis as we do not have to distinguish between different travellers in this case. As we only have one flow layer, we can omit the flow layer index and simply refer to $F_l$ and $P_j$ as the flows and inflows, respectively. The total travel time or cost $\tau$ reads

$$\tau = \sum_l \tau_l(F_l) F_l = \sum_l t_l \left(F_l + \eta F_l^2\right). \tag{II-5}$$

We demand that the flows should arrange in such a way that they minimise $\tau$ under the boundary conditions imposed by (II-4). Considering furthermore the directedness constrains $F_l \geq 0$, the mathematical problem can be written as:

$$\min_{F_l} \quad \sum_l t_l \Big(F_l + \eta F_l^2\Big) \tag{II-6a}$$

$$\text{subject to} \quad \sum_l I_{il} F_l = P_i \tag{II-6b}$$

$$F_l \geq 0\,. \tag{II-6c}$$

The objective function is a quadratic polynomial in $L$ variables with $N$ equality and $L$ inequality constrains. Such a system is called a *Quadratic Program* (QP) [21]. In general, solving a QP is NP-hard [30]. But if the objective function is convex, as it is the case in (II-6), the problem can be solved in polynomial time [31].

Restricting ourselves to a single flow layer allows a further simplification. In this thesis we do not consider explicit one-way edges. That is, two nodes $n$ and $m$ either have

no direct connection, or they are linked by two edges $(m, n)$ and $(n, m)$ representing both possible directions. If there is only a single source or a single sink, then there will never be a counter-flow between two nodes $m$ and $n$. That is, at most one of the links $(m, n)$ and $(n, m)$ will have a non-vanishing flow.

This fact allows for an alternative description of the network as an undirected graph, representing each street or line only by one edge which can be crossed into both directions. Fixing an orientation for each edge, we can keep track of the direction of the flow by the sign of the flow $F_l$. In this case, we have to replace $F_l$ by $|F_l|$ in all expressions for the travel time $\tau$ and we can omit the inequality constrains $F_l \geq 0$ so that the optimisation problem for an undirected network reads

$$\min_{F_l} \quad \sum_l t_l \Big( |F_l| + \eta F_l^2 \Big) \tag{II-7a}$$

$$\text{subject to} \quad \sum_l I_{il} F_l = P_i \,. \tag{II-7b}$$

This alternative description will be used when discussing the solution in the Ohmic limit in section II.3 and for the analysis in chapter IV.

## II.2. Solution in the Shortest path limit

In the limit $\eta = 0$, the quadratic terms in the objective function vanish and we obtain the Linear Programme (LP)

$$\min_{F_l} \quad \sum_l t_l F_l \tag{II-8a}$$

$$\text{subject to} \quad \sum_l I_{il} F_l = P_i \tag{II-8b}$$

$$F_l \geq 0 \,, \tag{II-8c}$$

which is known as the *shortest path problem* [22], we will therefore refer to this limit as the shortest path limit in this thesis. A standard way to solve this problem is to apply Dijkstra's algorithm [32].

## II.3. Solution in the Ohmic limit

In the limit $\eta \to \infty$, the optimisation problem (II-6) reads

$$\begin{aligned} \min_{F_l} \ & \sum_l t_l F_l^2 \\ \text{subject to } & \sum_l I_{il} F_l - P_i = 0 \\ & F_l \geq 0 \,. \end{aligned}$$

Using the description in terms of undirected networks introduced above we can omit the constrains $F_l \geq 0$ and the problem reduces to

$$\min_{F_l} \ \sum_l t_l F_l^2 \tag{II-9a}$$

$$\text{subject to } \sum_l I_{il} F_l - P_i = 0 \tag{II-9b}$$

which is identical to the program which solves the flows in Ohmic networks such as electrical power grids [23, ch.IX, thm.1]. In this case, the problem becomes particularly simple to solve using the method of Lagrange multipliers [33].

In this formalism, we introduce for each constrain a Lagrange multiplier $\lambda_i$ and define the Lagrangian

$$\mathcal{L} = \tau - \sum_i \lambda_i \left( \sum_l I_{il} F_l - P_i \right) .$$

For the optimal set of flows $F_l$, the Lagrangian $\mathcal{L}$ must be stationary, i.e.

$$\partial_{F_l} \mathcal{L} = \partial_{F_l} \tau - \sum_i \lambda_i I_{il} = 2 t_l F_l - \sum_i \lambda_i I_{il} \stackrel{!}{=} 0 \,.$$

We therefore get

$$F_l = \frac{\sum_i \lambda_i I_{il}}{2 t_l} = \kappa_l \sum_i \lambda_i I_{il} \tag{II-10}$$

with $\kappa_l := (2t_l)^{-1}$. Let us rewrite this relation in a slightly different form. For a link $l = (n, m)$ that connects two nodes $n$ and $m$, the flow reads

$$F_{n \to m} = \frac{\lambda_n - \lambda_m}{2 t_{nm}} . \tag{II-11}$$

We thus recover Ohm's law if we identify $\lambda_n$ as the voltage at node $n$ and $2t_{nm}$ as the resistance of the link $(n, m)$. Thus, we will henceforth refer to the limit $\eta \to \infty$ as the Ohmic limit.

To determine the $\lambda_i$, we insert (II-10) into the boundary conditions (II-9b):

$$\sum_l I_{il}\kappa_l \sum_j \lambda_j I_{jl} - P_i = 0\,,$$

which simplifies to

$$\sum_j \kappa_{ij}(\lambda_i - \lambda_j) = P_i \tag{II-12}$$

when using the properties of $I_{il}$. Introducing the matrix

$$A_{ij} = -\kappa_{ij} + \left(\sum_k \kappa_{ik}\right)\delta_{ij}\,,$$

this problem writes

$$\sum_j A_{ij}\lambda_j = P_i\,, \tag{II-13}$$

which can be solved by multiplying from the left with the (pseudo) inverse matrix $(A^{-1})_{ij}$:

$$\lambda_i = \sum_j (A^{-1})_{ij} P_j \tag{II-14}$$

Finally, we obtain the flows $F_l$ in the Ohmic limit by inserting (II-14) into (II-10).

## II.4. Generalisation to finite congestion

The previous approach of solving the optimisation problems in the limit $\eta \to \infty$ via Lagrange multipliers can be generalised to the case of finite $\eta > 0$ and directed networks using the theorem of Karush-Kuhn-Tucker (KKT) [21, ch. 5.5.3] which generalises the formalism of Lagrange multipliers: Considering the optimisation problem (II-6) for a network with $N$ nodes and $L$ edges, we introduce the generalised Lagrangian

$$\mathcal{L}(F_l, \lambda_i, \mu_l) := \underbrace{\sum_l t_l \left(F_l + \eta F_l^2\right)}_{\equiv \tau(F_l)} - \sum_i \lambda_i \left(\sum_l I_{il} F_l - P_i\right) - \sum_l \mu_l F_l$$

and demand that its gradient vanishs.

The KKT conditions then read

$$\partial_{F_l}\mathcal{L} = t_l\left(1 + 2\eta F_l\right) - \sum_i \lambda_i I_{il} - \mu_l = 0 \tag{II-15a}$$

$$\sum_l I_{il} F_l - P_i = 0 \qquad i = 1, \dots, N \tag{II-15b}$$

$$F_l \geq 0 \qquad l = 1, \dots, L \tag{II-15c}$$

$$\mu_l \geq 0 \qquad l = 1, \dots, L \tag{II-15d}$$

$$\mu_l F_l = 0 \qquad l = 1, \dots, L\,. \tag{II-15e}$$

Solving (II-15a) for $F_l$ and considering (II-15c), we find

$$F_l = \begin{cases} \frac{1}{2\eta t_l} \sum_i \left(\lambda_i I_{il} + \mu_l - t_l\right) & , \text{ if positive} \\ 0 & , \text{ else}\,. \end{cases}$$

Using (II-15e), we can omit $\mu_l$ as it has to be zero if $F_l \neq 0$. Hence, using the Heaviside function

$$\Theta(x) = \begin{cases} 1 & , \ x \geq 0 \\ 0 & , \ x < 0 \end{cases},$$

we can write

$$F_l = \frac{\sum_i \lambda_i I_{il} - t_l}{2\eta t_l} \Theta\left(\frac{\sum_i \lambda_i I_{il} - t_l}{2\eta t_l}\right).$$

Considering $l \mathrel{\hat{=}} (n, m)$, i.e. let edge $l$ start at node $n$ and end at node $m$, and using the definition of $I_{il}$, we can write this as

$$F_{n\to m} = \frac{\lambda_n - \lambda_m - t_{n\to m}}{2\eta t_{n\to m}} \Theta\left(\frac{\lambda_n - \lambda_m - t_{n\to m}}{2\eta t_{n\to m}}\right).$$

Introducing the matrix of capacities

$$\kappa_{nm} := \begin{cases} \frac{1}{2\eta t_{n\to m}} & , \text{ if link } (n, m) \text{ exists} \\ 0 & , \text{ else} \end{cases} \tag{II-16}$$

this becomes

$$F_{n\to m} = \left[\kappa_{nm}\left(\lambda_n - \lambda_m\right) - \frac{1}{2\eta}\right] \Theta\left(\kappa_{nm}\left(\lambda_n - \lambda_m\right) - \frac{1}{2\eta}\right). \tag{II-17}$$

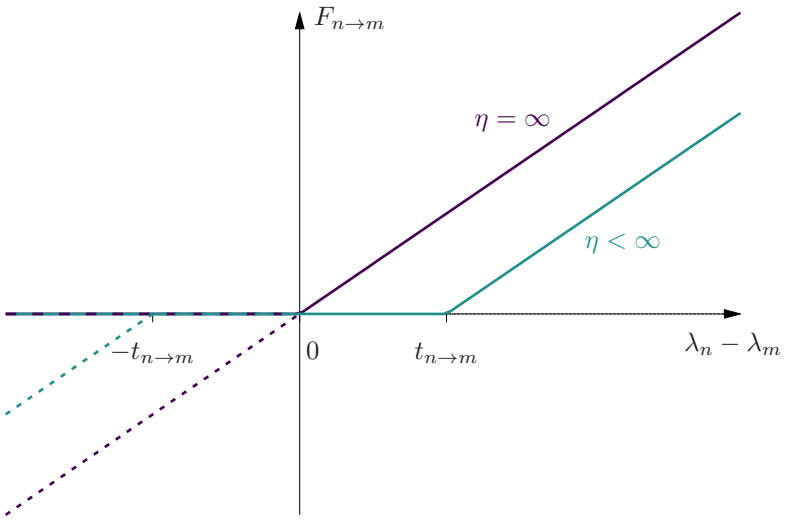



Figure II.1.: The flow $F_{n\to m}$ along an edge from node $n$ to node $m$ in dependence of the potential drop $\lambda_n - \lambda_m$ in the Ohmic limit $\eta = \infty$ and for finite congestion $0 < \eta < \infty$. The solid lines mark the flow for directed networks and the dotted lines for undirected networks, respectively.

Inserting (II-17) into (II-15b) yields

$$\begin{aligned} P_i &= \sum_l I_{il} F_l = \sum_{j:\,\exists(i,j)} F_{i\to j} - \sum_{j:\,\exists(j,i)} F_{j\to i} \\ &= \sum_j \left[\kappa_{ij}\left(\lambda_i - \lambda_j\right) - \frac{1}{2\eta}\right] \Theta\left(\kappa_{ij}\left(\lambda_i - \lambda_j\right) - \frac{1}{2\eta}\right) \\ &\quad - \sum_j \left[\kappa_{ji}\left(\lambda_j - \lambda_i\right) - \frac{1}{2\eta}\right] \Theta\left(\kappa_{ji}\left(\lambda_j - \lambda_i\right) - \frac{1}{2\eta}\right). \end{aligned} \tag{II-18}$$

Note that we recover (II-12) for undirected networks, i.e. $\kappa_{ij} = \kappa_{ji}$, in the Ohmic limit, i.e. $\eta \to \infty$ with $\kappa_{ij}^{-1} \propto \eta t_{i\to j}$ kept finite.

For finite congestion $\eta < \infty$, we can interpret the $\lambda_i$ as potentials like in the Ohmic limit. In Fig. II.1, the flow $F_{n\to m}$ on an edge from $n$ to $m$ is plotted over the potential drop $\lambda_n - \lambda_m$ along this edge. In the limit $\eta \to \infty$, we find a straight line, i.e. the flow is simply proportional to the potential drop in the Ohmic limit. For finite congestion, $\eta < \infty$, we observe a plateau for $|\lambda_n - \lambda_m| < t_{n\to m}$ caused by the Heaviside function in (II-17). Thus, a finite free-flow travel time $t_{n\to m}$ suppresses the flow on the edge up to an critical potential drop equals $t_{n\to m}$ while in the limit $t_{n\to m} \to 0$ the $-1/2\eta$ inside the Heaviside function becomes negligible such that the argument is always positive if the potential drop $\lambda_n - \lambda_m$ is.
However, the resulting equations (II-18) are nonlinear and even non-analytic in the potentials $\lambda_i$, which regularly causes problems for numerical solvers. Hence, this method is not used for actual computations in this thesis.

## II.5. Reducing the numerical complexity using an adapted Hardy Cross method

The optimisation problem (II-6) can be solved directly using a numerical QP solver. However, the solution becomes numerically very expensive for large systems as the runtime is polynomial in the number of edges $L$ of the system. To reduce the numerical complexity, we can make use of Kirchhoff's current law which reduces the degrees of freedom of the system. For simplicity we will use undirected networks in the following. We can later recover the case of directed networks by imposing the inequality constrains (II-6c).

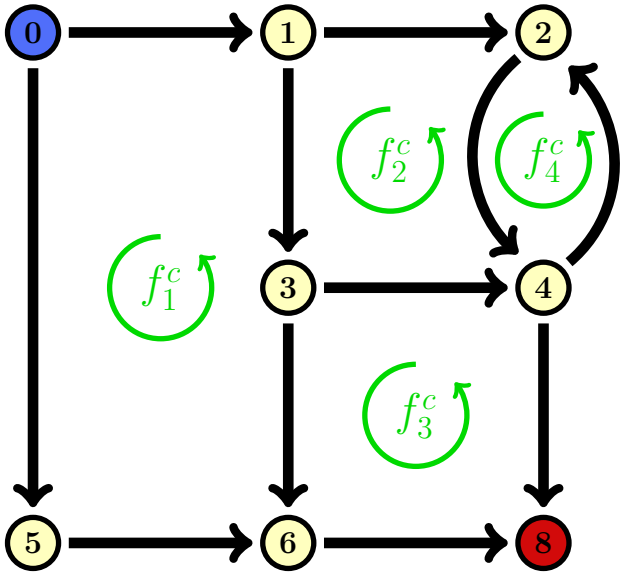



Figure II.2.: Sketch of a network with loops: Each loop adds a new degree of freedom (d.o.f.) on how the flows can arrange. The d.o.f. can be visualised as cycle flows $f_1^c$ which flow around each loop. Note that when having two edges connecting the same nodes we get also for each of these 2-loops a d.o.f. like $f_4^c$.

If the network is a tree, the flows are fully determined by flow conservation. For each closed loop in the network, a new degree of freedom (d.o.f.) arises to arrange the flows so that they are not violating Kirchhoff's current law. As sketched in fig. II.2, we can consider these degrees of freedom as cycle flows which are added to the initial flow on each edge with parallel orientation on the loop and subtracted from the flows on edges where the orientation is anti-parallel.

With these thoughts, we can partially solve the problem using a method inspired by the Hardy Cross method [34], which was originally developed to calculate flows in pipe networks.

The method involves the following steps:

1. find a flow solution $F_{l,0}$ which fulfils Kirchhoff's current law at each node.

2. add for each closed loop a cycle flow $f_i^c$ flowing along the edges which form the loop. If the orientation of $f_i^c$ and the edge are parallel, $f_i^c$ is added to the $F_{l,0}$. When they are anti-parallel, it is subtracted.

3. find the set of $f_i^c$ which minimise $\tau$. In directed networks there are the inequality boundary condition $F_l = F_{l,0} \pm f_i^c \geq 0$ which still must be fulfilled while the equality constrains remain in each case fulfilled.

To formalise this method, we have to introduce the circuit incidence matrix $C_{li}$ with the components

$$C_{li} = \begin{cases} +1 & \text{, if edge } l \text{ is parallel to loop } i \\ -1 & \text{, if edge } l \text{ is anti-parallel to loop } i \\ 0 & \text{, if edge } l \text{ is not in loop } i \end{cases} \tag{II-19}$$

Using this definition, the resulting flows $F_l(f_i^c)$ after adding the cycle flows $f_i^c$ to the initial guess $F_{l,0}$ read

$$F_l(f_i^c) = F_{l,0} + \sum_i C_{li} f_i^c \,. \tag{II-20}$$

To simplify calculations, we restrict from here on directed networks. The system can now be written as

$$\begin{aligned} &\min_{f_i^c} \; \tau(F_l(f_i^c)) - \tau(F_{l,0}) \\ &\text{subject to } F_l(f_i^c) = F_{l,0} + \sum_i C_{li} f_i^c \geq 0 \,, \end{aligned}$$

where we subtracted $\tau(F_{l,0})$ from the original objective function as this is a constant with respect to the cycle flows $f_i^c$ and simplifies the expression. When we insert (II-20) into (II-6a), we get

$$\begin{aligned} \tau(F_l(f_i^c)) - \tau(F_{l,0}) &= \sum_l t_l \left[ \left( F_l(f_i^c) + \eta F_l^2(f_i^c) \right) - \left( F_{l,0} + \eta F_{l,0}^2 \right) \right] \\ &= \sum_l t_l \left[ \left( F_{l,0} + \sum_i C_{li} f_i^c + \eta (F_{l,0} \sum_i C_{li} f_i^c)^2 \right) - \left( F_{l,0} + \eta F_{l,0}^2 \right) \right] \\ &= \sum_l t_l \left[ \sum_i C_{li} f_i^c + 2\eta F_{l,0} \sum_i C_{li} f_i^c + \eta \sum_{i,j} C_{li} f_i^c C_{lj} f_j^c \right] \\ &= \sum_{i,j} \eta \sum_l C_{li} C_{lj} f_i^c f_j^c + \sum_i \sum_l t_l \left( 1 + 2\eta F_{l,0} \right) C_{li} f_i^c \\ &= \sum_{i,j} f_i^c \beta_{ij} f_j^c + \sum_i \alpha_i f_i^c \,, \end{aligned}$$

were we introduced the symmetric matrix

$$\beta_{ij} = \eta \sum_l C_{li} C_{lj} \tag{II-21}$$

and the vector

$$\alpha_i = \sum_l \left(1 + 2\eta F_{l,0}\right) C_{li} \,. \tag{II-22}$$

We find the objective function to be again a quadratic polynomial so that the remaining system is also a QP:

$$\min_{f_i^c} \; \sum_{i,j} f_i^c \beta_{ij} f_j^c + \sum_i \alpha_i f_i^c \tag{II-23a}$$

$$\text{subject to } F_l(f_i^c) = F_{l,0} + \sum_i C_{li} f_i^c \geq 0 \,. \tag{II-23b}$$

While the initial QP was of order $L$ with $N$ equality and $L$ inequality constrains, the new QP is only of order $n_l$ with $L$ inequality constrains where $n_l$ is the number of loops in the network. On the other hand, the reduced QP is in general not convex so that optimised solving algorithms for convex QPs can not be applied here.

Considering real world networks like the Cologne tram network [35], there are much fewer loops than number of links, i.e. $n_l \ll L$, if we consider each connection between two stations as a link. Without considering different lines, in the Cologne tram network we find 219 connections and only 18 loops. In addition to these large loops we must also consider that each connection in this network consists of two edges – one for each direction. Thus, we get additional 219 loops consisting out of the anti-parallel edges connecting the same two nodes. In the end, this means that the original QP is of order $L = 438$ while the reduced QP with the adapted Hardy Cross method is only of order $n_l = 237$, i.e. the order is only a bit more than half of the original value.

## II.6. Further reductions of complexity

As examined in the previous section for the Cologne subway grid, 219 out of 237 degrees of freedom (d.o.f.) in the reduced QP, which remains to solve when using the adapted Hardy-Cross method, are caused by the anti-parallel edges in the network, i.e. the fact that along all links in the network one can go in both directions. Since we observe a large number of such bidirectional connections between nodes in most transport networks, it would be very efficient to handle these d.o.f. separately instead of solving them using a QP solver, as we could reduce the d.o.f. in the case of the Cologne tram network by more than 90% in this way.

We can easily determine the values of the cycle flows associated to a pair of anti-parallel edge like $f_4^c$ in Fig. II.2. We know that the flow on one of the two edges

which form the loop along which $f_4^c$ flows must be exactly zero as we could reduce the cost $\tau$ otherwise by adapting $f_4^c$ such that the smaller of these flows vanishes.

However, we cannot implement this argument into the QP solver without additional assumptions. One possibility to fix these d.o.f. is to assume that our initial guess already determined most of the flow directions correctly. If the flow directions on each link is the same for the initial guess and the optimal flows, then we could remove all edges which are going in the opposite direction to the initial flow so that all the 2-edge loops disappear. Since we use for our initial guess the Ohmic flow pattern, which corresponds to the $\eta \to \infty$ limit, we would except only small qualitative differences between initial guess and optimal flows.

But when we omit these edges, we must ensure that we detect where the optimal flow might switches the direction compared to the guess. As long as all flows in the optimal pattern are greater than 0, the boundary conditions (II-23b) are not active so that we can be sure that we have found the optimal flows. If there are edges where the flow vanishes, we have to reinsert the corresponding anti-parallel edge and to solve the QP again for the extended network. This procedure should be repeated until no flows are vanishing on edges where the counter part is disabled.

This method is very efficient if there are no or only few edges where the flow switches directions but it can be very inefficient if the procedure must be repeated frequently.

## II.7. Runtime analysis

Solving a *QP* in general is NP-hard [30]. But also the solution of *convex QP* can be numerically very complex. To get a feeling for the scaling of the QP solver runtime with the number of degrees of freedom (d.o.f.), the squaregrid illustrated in fig. III.5 was solved for different lattice sizes $D$ for a very large $\eta = 10^6$ and an intermediate $\eta = 10$. The runtimes are plotted in fig. II.3. The $d.o.f.$, i.e. the number of edges in the squaregrid, is given by

$$d.o.f.(D) = 4D^2 - 2$$

For large $\eta$, where the flow fills all of the network, the runtimes are $\mathcal{O}(d.o.f.^3)$ while for intermediate $\eta$ with a large number of vanishing flows, the complexity is $\mathcal{O}(d.o.f.^4)$. Using the library *QuadProg++* [36], which implements a convex QP solver using the algorithm of Goldfarb and Idnani [37], the complexity could be reduced to $\mathcal{O}(d.o.f.^{2.7})$ but at the price that the solver is less reliable in the sense that it cannot solve some networks.

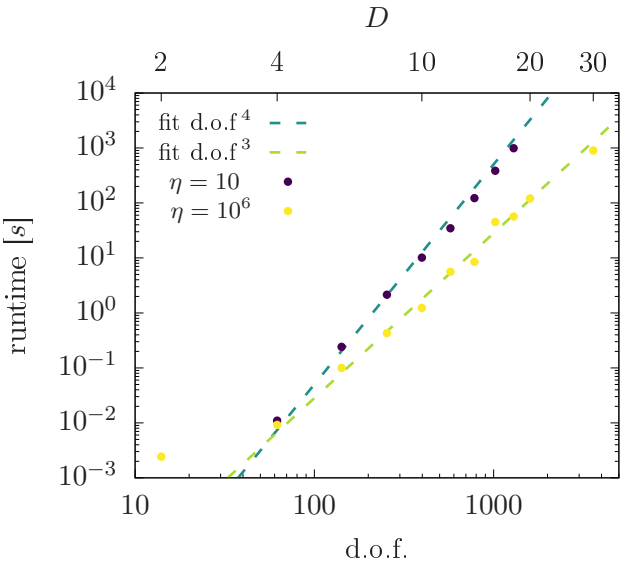



Figure II.3.: Runtimes to solve the $D \times (D+1)$ squarelattice as skechted in fig. III.5 for different values of $\eta$. The degrees of freedom (d.o.f.) of the QP, i.e. the number of edges in the network, which correspond to the lattice size $D$ are plotted on the lower x-axis.

# III. Fundamental flow patterns in the presence of congestion

In this chapter, we will analyse the optimal flow patterns for some elementary networks to elucidate the impact of congestion.

First, we consider two small scale benchmark networks which allow for an analytical treatment in section III.1. These network feature one source and one sink term linked by two or three different paths. We will focus on the question whether the flow is concentrated on the shortest path or branches onto different paths. This analysis is generalised to larger networks in section III.2.

In section III.3, we will then turn to a problem of high practical relevance. How does the network respond to the failure of a single link? How is the flow rerouted past the failure? In the absence of congestion ($\eta = 0$), flow rerouting will always proceed via the shortest path. In the Ohmic limit, different scaling laws have been obtained for the decay of flow changes with distance to the failing link [38, 39]. Here, we extend these results to networks that include congestion.

The results of this chapter are obtained by exploiting the linear congestion model for directed networks which is introduced in sec. II.1. For larger networks, we flows are computed using a numerical QP solver. The benchmarks allows us to check whether the numerical solver is working correctly.

## III.1. Benchmarks

The optimal flow pattern that minimise the total travel times cannot be obtained in closed form for most cases. However, for some fundamental networks, analytical solutions can easily be found. In this section we derive and discuss solutions for two general networks:

1. a network where a source and a sink are connected via two concurrent branches and

2. a $3 \times 2$ square grid with a source and sink next to each other in the centre.

By studying these simple geometries, we can already learn about the impact of congestion on the flow pattern. Furthermore, we can use these benchmarks to test numerical solvers by comparing their results with the analytical solutions.

### III.1.1. Two concurrent branches

Consider a network with one source and one sink connected by two branches with different positive weights $t_1$ and $t_2$, respectively. The strength of both the source and the sink are set to 1 such that, according to Kirchhoff's current law (II-6b), in total there is a unit flow from the source to the sink. The scenario is sketched in Fig. III.1.

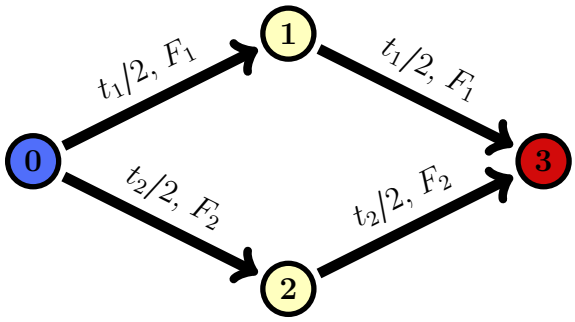



Figure III.1.: Sketch of a network with a source (0) and a sink (3). They are connected via two concurrent branches, one of them going via node (1) and the other via node (2), respectively. Each branch consists of edges with weight $t_1/2$ and $t_2/2$, respectively, and carry a flow $F_1/F_2$.

Although this is a very simple network, it represents a fundamental network structure that can frequently be observed as a building block in larger networks: It models the problem of having different paths connecting a single source-sink pair in a general setting. In the limit $\eta = 0$, the flow only takes the shortest path, i.e. the path with lowest cost. In Ohmic networks, characterised by $\eta \to \infty$, however, we can observe that the flow always distributes over all available branches. This raises the question which paths the flow takes when varying the congestion parameter $\eta$.

As flow will only go from left to the right in Fig. III.1, we can ignore the links in the opposite direction and proceed only with the remaining links. From Kirchhoff's current law we know that the flows on the edges $(0,1)$ and $(1,3)$ as well as on $(0,2)$ and $(2,3)$ must be identical and sum up to 1. Hence the flow can be fully described by two components $F_1$ and $F_2$ which are always non-negative with each component being the flow along one of the two branches. We can therefore write

$$F_2 = 1 - F_1 \tag{III-1}$$

and insert this into the QP for directed networks (II-6). Both the equality and inequality constraints are already taken into account by the previous reasoning. To

solve the QP, it therefore remains to insert $F_1$ and (III-1) into the objective function (II-6a):

$$\begin{aligned}\tau &= 2 \cdot \frac{t_1}{2}(F_1 + \eta F_1^2) + 2 \cdot \frac{t_2}{2}(F_2 + \eta F_2^2) \\ &= t_1(F_1 + \eta F_1^2) + t_2\big((1 - F_1) + \eta(1 - F_1)^2\big) \\ &= t_2 + (t_1 - t_2)F_1 + \eta t_1 F_1^2 + \eta t_2(1 - 2F_1 + F_1^2) \\ &= t_2(1 + \eta) + (t_1 - t_2(1 + 2\eta))F_1 + \eta(t_1 + t_2)F_1^2\end{aligned}$$

and to find the $F_1$ that minimises $\tau$. For $\eta > 0$, this expression is a parabola with positive quadratic coefficient $\eta(t_1 + t_2)$. Hence, it becomes minimal in the point where the derivative w.r.t $F_1$ vanishes:

$$\partial_{F_1}\tau = (t_1 - t_2(1 + 2\eta)) + 2\eta(t_1 + t_2)\tilde{F}_1 \stackrel{!}{=} 0\,,$$

so that we get

$$\tilde{F}_1 = \frac{t_2(1 + 2\eta) - t_1}{2\eta(t_1 + t_2)}\,.$$

We can rescale time such that $t_2 = 1$ without loss of generality. Furthermore we need to consider the constraint $0 \leq F_1 \leq 1$ as both flow components $F_1$ and $F_2$ are non-negative. Introducing an effective flow

$$\tilde{f}_1(t_1, \eta) = \frac{1 + 2\eta - t_1}{2\eta(1 + t_1)}$$

we thus obtain

$$F_1(t_1, \eta) = \begin{cases} 0 & , \tilde{f}_1(t_1, \eta) < 0 \\ \tilde{f}_1(t_1, \eta) & , \tilde{f}_1(t_1, \eta) \in [0, 1] \\ 1 & , \tilde{f}_1(t_1, \eta) > 1 \end{cases}\,. \tag{III-2}$$

Note that this expression also covers the case $\eta = 0$ as we find $\tau|_{\eta=0} = 1 + (t_1 - 1)F_1$, which is minimised for $t_1 > 1$ by the minimal $F_1$ value, i.e. $F_1 = 0$ and for $t_1 < 1$ by the maximal $F_1$, i.e. $F_1 = 1$. Evaluating (III-2) gives the same result. We note that for $\eta = 0$ and $t_1 = t_2$, $\tau$ is independent of $F_1$ so that there is no distinguished optimal flow. However, an infinitesimal variation of either $\eta$, $t_1$ or $t_2$ allows already to define the optimal flow so that this special case has physically no importance.

In Fig. III.2, the optimal flow $F_1$ is plotted for both varying values of $\eta$ (Fig. III.2a) and $t_1$ (Fig. III.2b). We find that the flow is split up into two non-vanishing components if $t_1 \in \left[\frac{1}{1+2\eta}, 1 + 2\eta\right]$. As the width of this interval increases with $\eta$, we can draw the conclusion that the range of relative branch weights $t_1$ grows linearly with $\eta$ while the curve inside this range is smoothly tilted from the step function at $\eta = 0$ to the Ohmic limit $F_1(t_1) = 1/(1 + t_1)$.

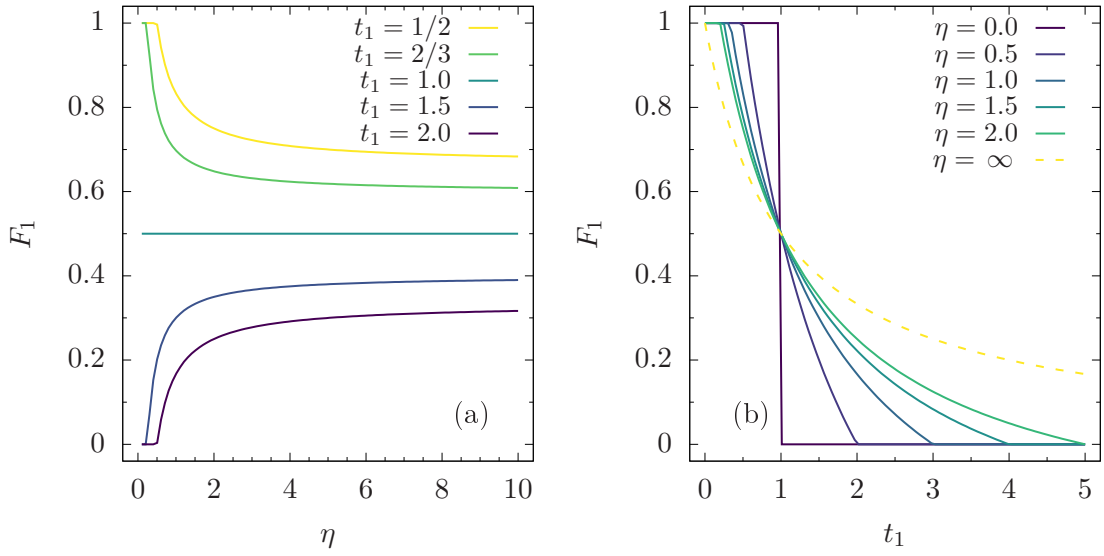



Figure III.2.: The optimal flow $F_1$ in a network with two concurrent branches in dependence of the congestion parameter $\eta$ and the ratio of the branch weights $t_1$ according to (III-2) plotted in (a) the $\eta$ direction and (b) the $t_1$ direction.

### III.1.2. $3 \times 2$ square lattice

Another general geometry for which we can easily find an analytical solution is the $3 \times 2$ square lattice as sketched in fig. III.3a. In this setting, we have a unit-strength pair of source and sink which are connected via three branches. There is one direct link and two paths which consist of a series of three links each. Due to the symmetry of the two long branches, we can assume that the flows are always equal on both of them. Thus, we have again two flow components which sum up to $F_1 + F_2 + F_2 = 1$ so that we can express $F_1$ in terms of $F_2$:

$$F_1(F_2) = 1 - 2F_2 \,.$$

Since all edges have the same weight $t$, the time needed to travel along the branch of $F_1$ is $t_1 = t$ and for the branches of $F_2$ we find $t_2 = 3t$. The total travel time $\tau$ therefore reads:

$$\begin{aligned}
\tau &= t(F_1 + \eta F_1^2) + 2 \cdot 3t(F_2 + \eta F_2^2) \\
&= t\left(1 - 2F_2 + \eta(1 - 2F_2)^2\right) + 6t\left(F_2 + \eta F_2^2\right) \\
&= t + 4tF_2 + 6\eta t F_2^2 + \eta t(1 - 4F_2 + 4F_2^2) \\
&= (1 + \eta)t + 4(1 - \eta)tF_2 + 10\eta t F_2^2 \,.
\end{aligned}$$

The derivative w.r.t. $F_2$ is

$$\partial_{F_2}\tau = 4(1 - \eta)t + 20\eta t F_2$$

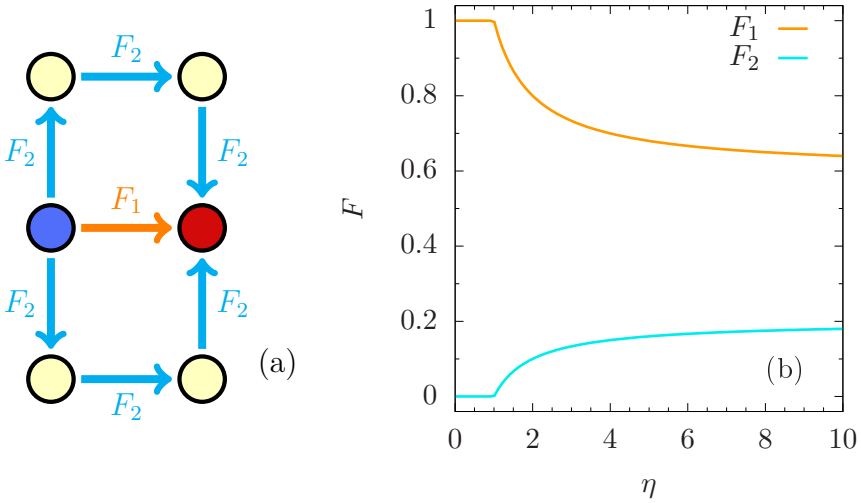



Figure III.3.: A $2 \times 3$ square lattice with a source (green) and a sink (red). The flow splits up into a part $F_1$ which takes the direct link and a part $F_2$ which take the upper and lower branch. (a) Sketch of the network, (b) the flow components in dependence of the congestion parameter $\eta$.

which vanishes for

$$\tilde{f}_2(\eta) = \frac{\eta - 1}{5\eta} = \frac{1 - \eta^{-1}}{5}.$$

Taking into account that both $F_1$ and $F_2$ must be non-negative, we find the expression for the flow along the long branches

$$F_2(\eta) = \begin{cases} 0 & , \eta < 1 \\ \tilde{f}_2(\eta) & , \eta \geq 1 \end{cases}. \tag{III-3}$$

The flows are plotted over $\eta$ in fig. III.3b. We find that the long branches only carry flow for $\eta > 1$, which can be generalised to larger squaregrids where source and sink are next to each other on the lattice which we will do in sec. III.3. Thus we can define a critical $\eta_c = 1$ for square grids, below which we observe a flow only on the shortest path while above this value, the flow splits up on more branches. Note that this result is independent of the lattice scale $t$.

## III.2. Branching

In the previous section, we found for $\eta$ larger than a critical $\eta_c$, which depends on the network geometry, that the flow is branching if there are several paths available. Studying this branching process in more detail plays an important role when considering networks with multiple source-sink pairs as this effect can lead to interference

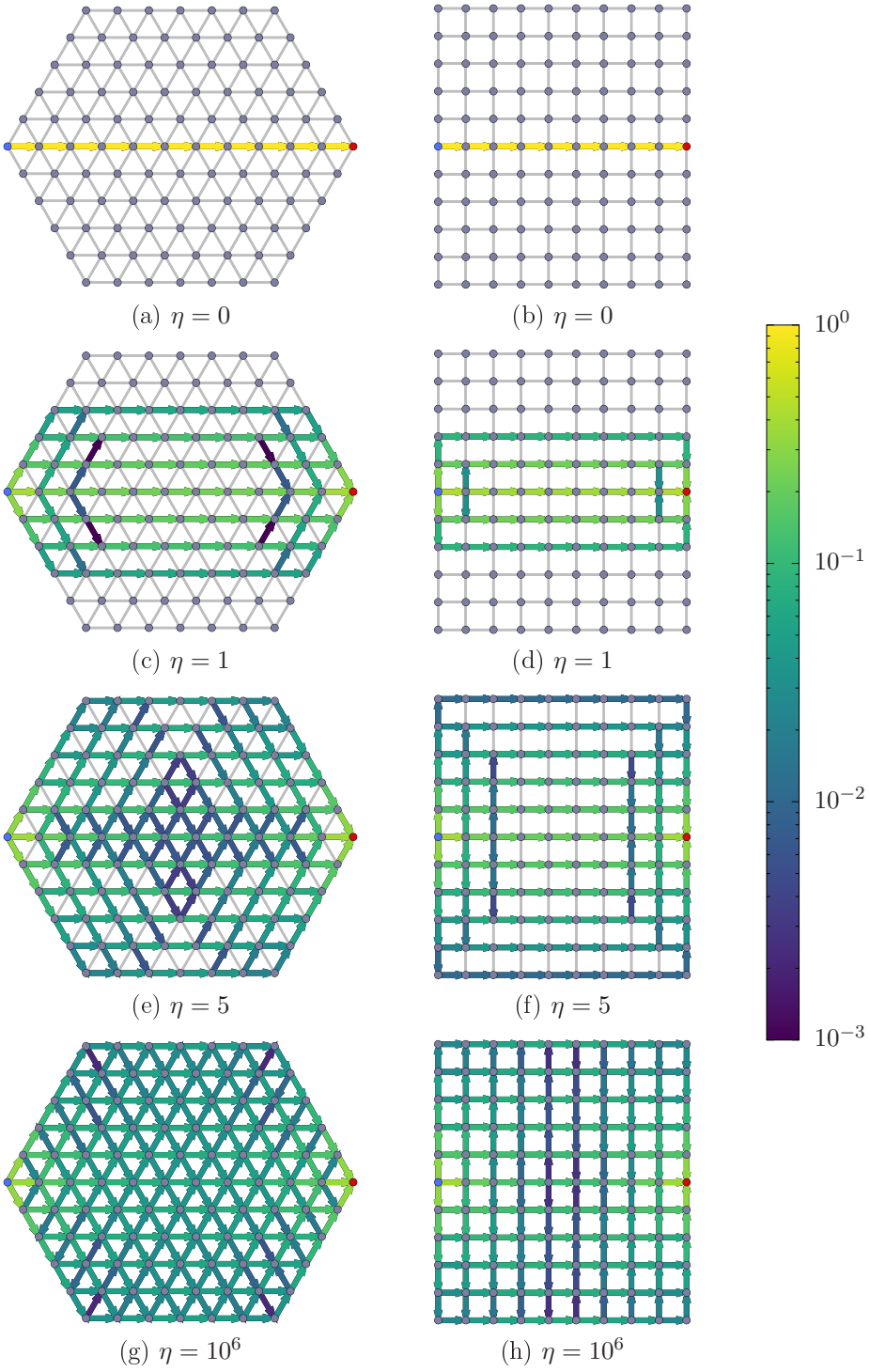



Figure III.4.: Flow pattern in dependence of the congestion parameter $\eta$ on both the triangular and square lattice which a source at the left end and a sink at the right end of the grid. Non-zero flows are illustrated by arrows pointing in the direction of flow and colored according to the flow strength.

between different flows although the shortest paths between their source-sink pairs do not intersect.

In sec. II.1, we found that using the linear congestion model, we get in the limit $\eta \to \infty$ the same pattern as in Ohmic flow networks. A general observation in this type of networks is a very broad branching covering all available paths. In the other limit, $\eta = 0$, the flow will only use the path with lowest cost as we have seen in the previous section. This prompts the question what the branching looks like for intermediate $\eta$.

We examine this question for both a triangular and a square lattice with equal weights $t_l = t$ for all edges. In both cases, the source is located on the left end of the lattice while the sink lays on the same height at the opposite end. In the vertical direction, the lattices are symmetric around the axis which connects the source and the sink. The strength of source and sink are again normalised to 1. In fig. III.4, the networks are illustrated with the flow pattern visualised for different values of $\eta$.

We finds that in the limit $\eta = 0$ (fig. III.4a-b), all the flow is only taking the shortest path from the source to the sink, as expected. When increasing the congestion parameter $\eta$, we observe that the flow expands to more and more alternative paths until it fills all links in the $\eta \to \infty$ limit (fig. III.4g-h). Comparing the pattern on the square grid and the triangular lattice, we observe a general difference. On the triangular lattice, the flow activity is higher in the centre of the lattice while the edges on the outer branches carry less flow. On the square lattice we observe the opposite effect: The links perpendicular to the direct source-sink connection carry less flow in the centre than on the left and right parts of the lattice.

From this observation we can conclude that the choice of the lattice type can have an important impact on the pattern.

## III.3. Rerouting pattern

In network science, the stability and resilience against damages is an important issue and subject to current research [38, 39, 40, 41]. There are different types of damages and ways, how it can impact the network such as reductions in capacity up to the complete failure of one or multiple links. In this thesis, we will focus on the complete failure of a single link that is mathematically equivalent to removing this link from the network. In transport networks, these outages frequently occur due to accidents, tempests or technical defects. When planning a transport network, it is therefore important to study the rerouting pattern, that is the difference in the flow caused by the removal of a link.

In the limit $\eta = 0$, i.e. when congestion is not coupled to the travel time, all the flow from the failing link will take the shortest rerouting path(s) – given the network

does not become disconnected by the failure. The Ohmic limit $\eta \to \infty$ is particularly important in the analysis of line outages in power grids [41] or pipe failures in hydraulic or vascular networks [7]. In this case, flow rerouting pattern can become very complex and include strongly non-local impacts. The blackout in wide parts of Europe on the 4th of November 2006 [42] is an example of the tremendous consequences which the removal of a single link can imply in such networks.

In the following we briefly review some results obtained in the Ohmic $\eta \to \infty$ and then numerically investigate flow rerouting as a function of the congestion parameter $\eta$.

## III.3.1. Flow rerouting in the Ohmic limit

The Ohmic limit $\eta \to \infty$ has been extensively discussed in power system security assessment. An explicit formula for the flow changes after a single line failure can be found in most textbooks on power system operation (see, e.g. [43]). So consider an edge $e$ that initially carries a flow $F_e$. If this edge is removed from the network, the flows on all other links change as

$$\Delta F_l = \mathrm{LODF}_{l,e}\, F_e\,, \tag{III-4}$$

where the prefactor is commonly referred to as the Line Outage Distribution Factor (LODF). These distribution factors reveal how strongly the flow on a link is affected by the failure of another link. A deeper analysis of these factors analysis, which is subject to current research [41], allows to identify critical links which can menace the stability of the network when they fail. Subsequently, one can optimise the network to mitigate the impact of failures of critical links and increase the resilience of the network.

Here, we briefly review the derivation of the LODFs following [44] in the Ohmic limit. The starting point is the equation (II-13) which is rewritten in matrix form as

$$A\,\vec{\lambda} = \vec{P}\,. \tag{III-5}$$

When an edge $l \mathrel{\hat{=}} (s,r)$ is removed from the network, the matrix $A$ changes as

$$\begin{aligned} A \to \hat{A} &= A + \Delta A \\ \Delta A &= A_{rs}\vec{\nu}_{rs}\vec{\nu}_{rs}^{\,\top}\,, \end{aligned} \tag{III-6}$$

where $\vec{\nu}_{rs}$ is a vector which is +1 at position $r$, −1 at position $s$ and zero otherwise. This causes a change of the Lagrangian multipliers

$$\vec{\lambda} \to \hat{\vec{\lambda}} = \vec{\lambda} + \vec{\psi}.$$

Subtracting Equation (III-5) for the perturbed and unperturbed grid, we obtain

$$\begin{aligned}\vec{\psi} &= -(A+\Delta A)^{-1}\Delta A\,\vec{\lambda} \\ &= (A+\Delta A)^{-1}\,\vec{\nu}_{rs}F_{rs}. \end{aligned} \tag{III-7}$$

The change of flows on another edge $(m,n)$ and thus the LODFs are calculated from equation (II-11) which yields

$$\begin{aligned}\Delta F_{mn} &= -A_{mn}(\psi_m - \psi_n) \\ &= -A_{mn}\vec{\nu}_{mn}^{\top}(A+\Delta A)^{-1}\vec{\nu}_{rs}F_{rs}. \end{aligned} \tag{III-8}$$

The matrix inverse can be evaluated explicitly with the help of the Woodbury matrix identity [45], which yields

$$(A + A_{rs}\vec{\nu}_{rs}\vec{\nu}_{rs}^{\top})^{\dagger}\vec{\nu}_{rs} = (1 + A_{rs}\vec{\nu}_{rs}^{\top}A^{-1}\vec{\nu}_{rs})^{-1}\,A^{\dagger}\vec{\nu}_{rs}, \tag{III-9}$$

such that the flow change (III-8) reads

$$\Delta F_{mn} = \frac{-A_{mn}\vec{\nu}_{mn}^{\top}A^{\dagger}\vec{\nu}_{rs}}{1 + A_{rs}\vec{\nu}_{rs}^{\top}A^{-1}\vec{\nu}_{rs}} \times F_{rs}\,,$$

which directly yields the Line Outage Distribution Factors.

Further analytic insight can be obtained by focusing on the change of the Lagrange multipliers [39]. Using again the Woodbury matrix identity, we can rewrite (III-7) as

$$A\vec{\psi} = \frac{F_{rs}}{(1 + A_{rs}\vec{\nu}_{rs}^{\top}A^{-1}\vec{\nu}_{rs})}\vec{\nu}_{rs}\,. \tag{III-10}$$

Now one can observe that $A$ is a Laplacian matrix, such that equation (III-10) can be interpreted as a discrete Poisson equation. The right-hand side includes only two non-zero entries of opposite sign, such that it can be interpreted as dipole source. On a regular lattice, the flow changes thus coincide with a dipole field of the respective dimension. This correspondence has been states in a more precise form in [39]. On a two-dimensional lattice one thus finds that the flow changes $|\Delta F_l|$ decay algebraically with the distance to the removed edge with the exponent 2,

$$|\Delta F_l| \sim \text{distance}^{-2}\,. \tag{III-11}$$

## III.3.2. Numerical Simulations

We now turn to the flow rerouting problem for arbitrary values of the congestion parameter $\eta$. In this thesis, we consider an $n \times (n+1)$ square lattice with a unit-strength source-sink pair located on neighbouring nodes in the centre of the grid as

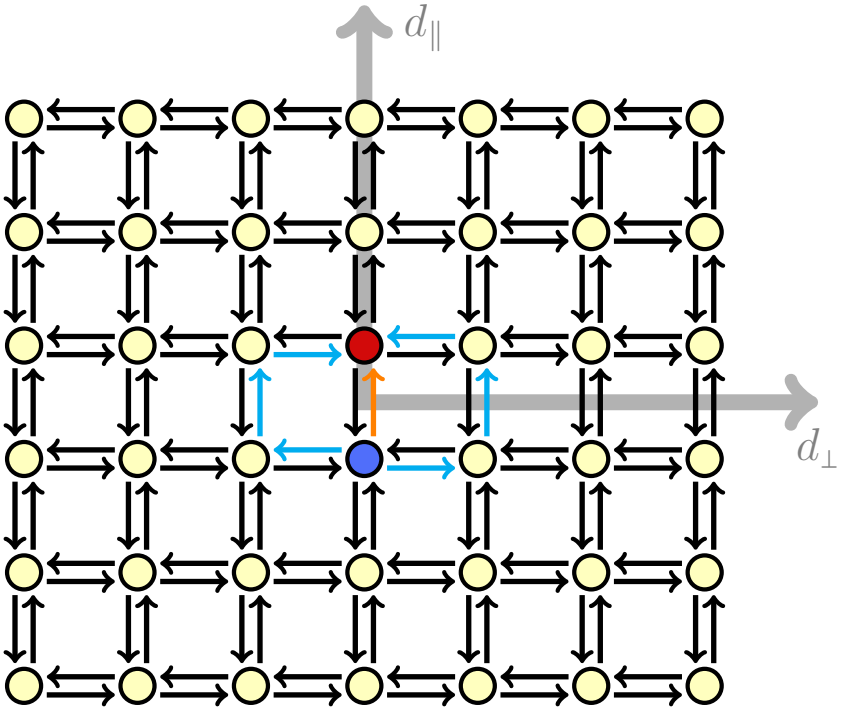



Figure III.5.: A square lattice with a source (blue) and a sink (red) in the centre next to each other. Each node on the lattice is connected in both directions with its closest neighbours. In this section we investigate how a failure of the direct connections between source and sink (orange) impacts the flow pattern in the lattice. Of particular interest are the flows on the edges along the axes parallel ($d_\parallel$) and perpendicular ($d_\perp$) to the failing link which are drawn grey. In the limit $\eta = 0$, only the edges on the shortest rerouting paths (cyan) are affected.

illustrated in Fig. III.5. To ensure mirror symmetry, we restrict the analysis in fact to even $n$. The basic question is: How does the flow in the network change when we remove the direct connection from the source to the sink?

We discuss the impact of the link failure first qualitatively for a small $10 \times 11$ lattice before we examine flow rerouting on a larger $120 \times 121$ lattice. While in the first case, the lattice size was chosen such that the network can still be visualised in a suitable way, the size in the latter case was chosen to keep the computation time feasible.

### III.3.3. Impact of the link failure on the flow range

In fig. III.6, the flows on a small square lattice are visualised for different values of the congestion parameter $\eta$. For each $\eta$, the flows on the full lattice are drawn next to the flows after removing the direct link between source and sink.

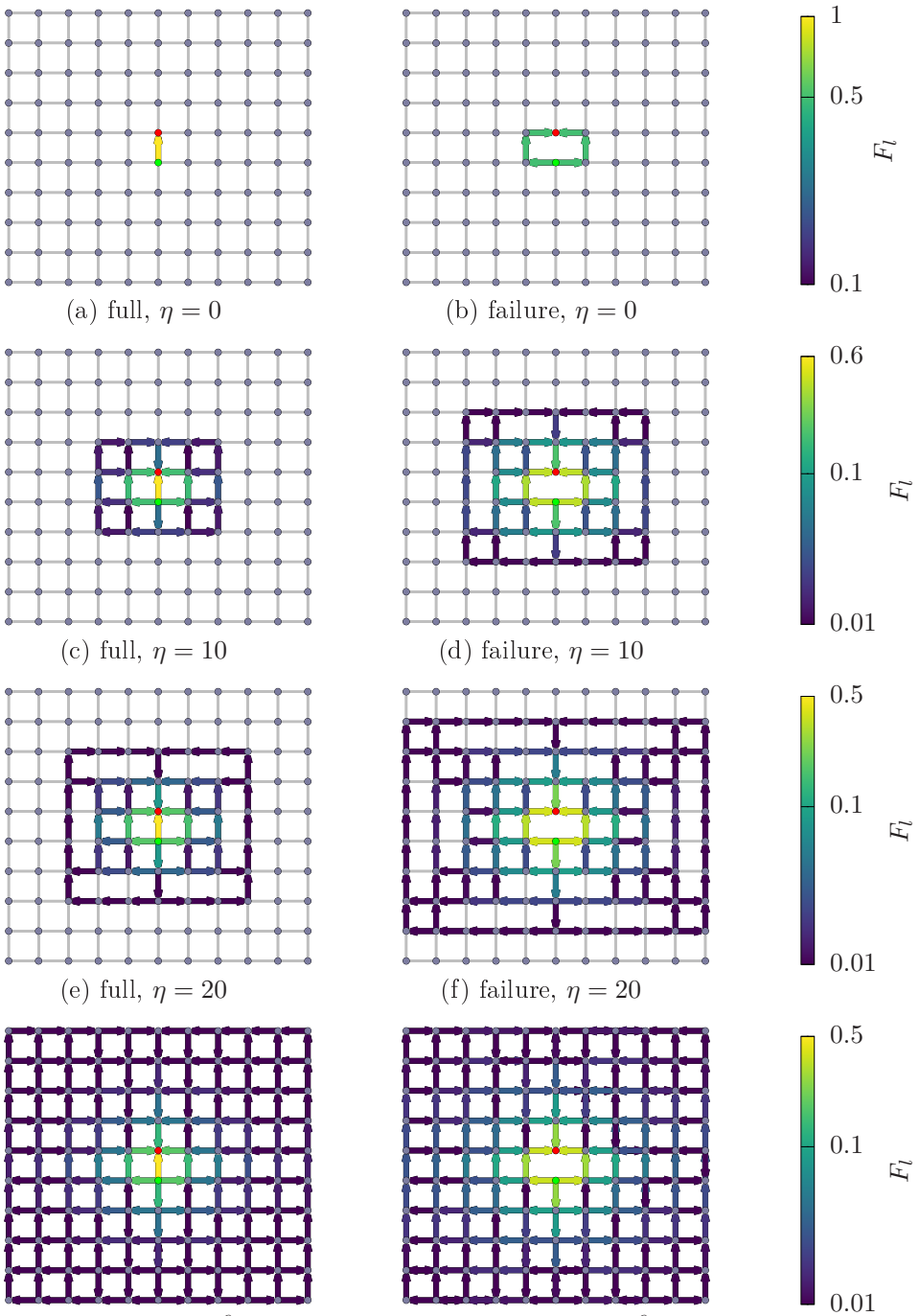



Figure III.6.: Flow pattern on the $10 \times 11$ square grid for different values of the congestion parameter $\eta$ for the full lattice (a,c,e,g) and with the central link between source and sink removed (b,d,f,h). Normal nodes and links without flow are drawn in grey. The source is marked green and the sink red. The flows are drawn as arrows with their head pointing in the direction of the flow and their colour indicating the flow strength.

For $\eta = 0$ (fig. III.6a-b), the flow only takes the shortest paths in both cases, as expected. In this case, there are two such paths (which are highlighted in cyan in Fig. III.5). Generalising the LODF to the case of an arbitrary $\eta$ via the defining equation (III-4), we thus find that LODF = 0.5 for the edges on these paths and zero for all edges laying not on these paths.

In the Ohmic limit $\eta \approx \infty$ (fig. III.6g-h), we observe a complex pattern which fills the entire lattice. For intermediate $\eta$ (fig. III.6c-f), we find flow patterns which cover an area which grows with increasing $\eta$. Close to the failure, the pattern look similar to the picture in the Ohmic limit. But beyond a certain range $r(\eta)$, the flows are exactly zero.

When comparing the flows with and without the central link for a fixed intermediate $\eta$, we observe that the cut-off distance $r(\eta)$ is larger than when the central link is removed. For $\eta = 20$ (fig. III.6e-f), we find the cut-off perpendicular to the failing link to be in lattice units $r_\perp(20) = 3$ while it is $r'_\perp(20) = 5$ with the central link removed.

Hence, we can conclude that the failure of the link may extend the range of flow. Considering a network with multiple flows between different source-sink pairs, this can evoke interference between different flow layers in case of a link failure even if the flows do not cross each other when the network is complete.

### III.3.4. Rerouting flow over distance from failure

In Fig. III.7a-b, the LODF on a $120 \times 121$ lattice are plotted over the distances perpendicular $d_\perp$ and parallel $d_\parallel$ to the failing link in the centre of the lattice. The axes $d_\perp$ and $d_\parallel$ are illustrated in Fig. III.5. We recognise in both directions cut-offs which we denote by $r_\perp$ and $r_\parallel$ respectively. Qualitatively, we observed these cut-offs already in Fig. III.6 and discussed them in the previous section.

In sec. III.3.1, we derived in the Ohmic limit that LODF $\propto d_{\perp/\parallel}^{-2}$. Close to the outer end of the lattice, the LODF vary from this relation due to finite size effects which were already observed and discussed in [39]. When decreasing $\eta$ from $\infty$ to finite values, we observe that the finite size features move closer to the centre. Furthermore, we find for distances sufficiently far below $r_{\perp/\parallel}$ that LODF $\propto d_{\perp/\parallel}^{-2}$. Hence, we can conclude that the flow pattern for finite values of $\eta$ is similar to the pattern on a lattice with a size proportional to $r_\perp$ and $r_\parallel$ in the Ohmic limit although we do not have an analytical proof for this yet.

In addition, we find in the $d_\perp$ direction a peak for $d_\perp$ closely below the cut-off $r_\perp$ which does not correspond to any feature observed in the Ohmic limit, where we only find a flattening of the LODF($d_\perp$) curve. This peak might arise from the difference in the flow range $r_\perp$ in the lattice before and after the link fails. We found already in Fig.

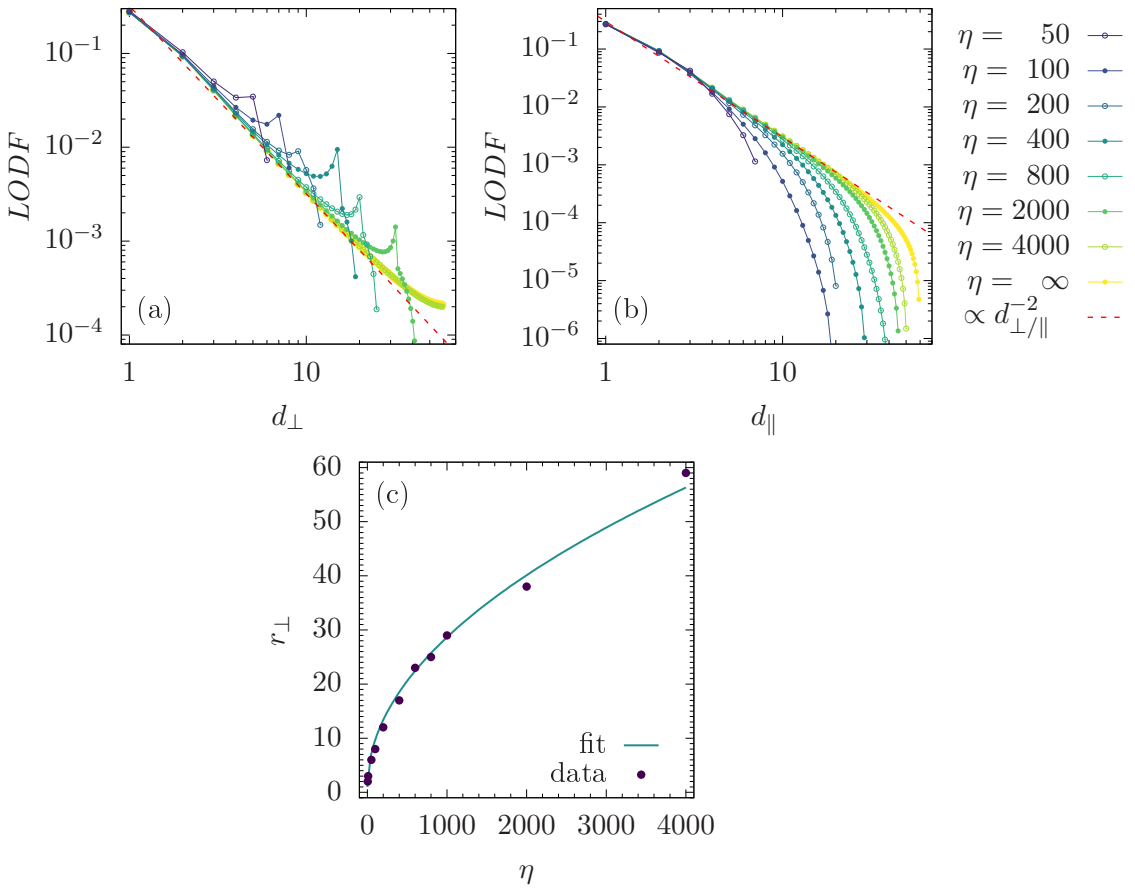



Figure III.7.: Line Outage Distribution Factors (LODF) for different values of the congestion parameter $\eta$ on the edges in (a) perpendicular and (b) parallel direction to the failing link along the axes drawn in fig. III.5. In the perpendicular direction, a sharp cutoff distance $r_\perp$ can be observed which is plotted in (c).

III.6 that the failure of the central link increases the flow range, i.e. there are links with flow $F_l = 0$ in the full lattice which carry a flow $F_l > 0$ after the outage. Thus, we can conclude that the link failure activates new paths in a distance $\approx r_\perp$ to the failing link disruptively, i.e. the LODF for these initially non-active links are significantly larger than expected at the corresponding distance.

In the next step, we quantify the relation between the flow range $r_\perp$ and the congestion parameter $\eta$. For this purpose, we plotted $r_\perp$ over $\eta$ in Fig. III.7c. We find a remarkably accurate $r_\perp(\eta) \propto \sqrt{\eta}$ relation which can be verified by the fit

$$r_\perp(\eta) = 1 + (0.875 \pm 0.0.016)\sqrt{\eta}\,,$$

which is plotted in the same figure. To verify this observation, more $\eta$ values should be evaluated.

Since the LODF diminish in $d_\parallel$ direction until they disappear within the numerical noise, the current data are not sufficiently precise to further quantify the properties of $r_\parallel$.

## III.4. Summary & Outlook

In this chapter, we first derived analytical solutions for two fundamental types of networks which can be used as benchmarks for numerical solvers. We found that the range of weight ratios of two concurrent paths for which a unit-strength flow split up on both paths grows linearly with $\eta$. We then generalised the analysis for flow branching to triangular and square lattices where we found increasingly complex branching pattern.

Last, we studied the rerouting pattern on a square lattice with congestion. We found a pattern similar to that of a dipole, as it was already observed in a previous work [39] for the Ohmic limit. For finite congestion, we found a flow cut-off which is proportional to $\sqrt{\eta}$ in the direction perpendicular to the failing link. Further, we found disruptive activation of links with zero flow before the outage.

For both the dipole pattern and the relation between cut-off range $r_\perp$ and congestion parameter $\eta$, analytical evidence should be found to support the empirical findings.

Based on these results, in a next step the impact of $\eta$ on the rerouting pattern should be studied in irregular lattices. In particular, the question of how to decouple certain regions of a network from rerouting flows caused by outages in other regions is very central to mitigate non-local impacts of a disturbance in the network. In Ohmic networks, this can be achieved for either weak coupling of these regions or by introducing network isolators, which were recently discovered [41]. This prompts the question whether one can observe similar effects also for finite congestion $\eta < \infty$.

# IV. Loop formation in a single-source network

As mentioned before, the proper functioning of supply networks play a crucial role in a variety of systems. However, most of the networks are facing limited resources such as available material, space or financial budgets. On the other hand, it is necessary to invest into redundant network components to keep it stable in case of extraordinary loads. Therefore, the resource constrains urge the systems to find a balance between minimal flow cost and maximal resilience.

In this chapter we focus on single-source networks where one provider supplies multiple users. Such networks can widely be observed: in natural systems like plant leaves but also in human-build networks like local electrical power grids. Public transport systems can also be studied as a single-sink network, which we will study in chapter V. The trade-off between efficiency and resilience was widely analysed for this class of networks [46, 47, 48]: if the demand is reliable, the optimal topology is graph-theoretically a tree, i.e. without any loop. In real-world networks, however, the demand is often not reliable which leads to a more complex topology and loops are observed.

In [40], the occurrence of a loop is analysed for a fundamental network in dependence of the source fluctuations and the budget constrains. The key observation is a discontinuous transition between a non-loopy and loopy topology. While the Ohmic dissipation was used as cost function in [40], we are generalising this analysis in this chapter for the cost function of the linear congestion model introduced in chapter II.

## IV.1. The scenario

Consider a network with one source which supplies 4 sinks arranged in 2 branches as sketched in fig. IV.1, where the source is red (node 1) and the sinks are blue (nodes 2-5). Assume we always have the black links 1-4 connecting all nodes with a tree network which could be interpreted as a star network with two branches. The cost, i.e. the time needed to travel along a link $l$, depends on the link capacity $\kappa_l$ and the local flow $F_l$. We are using again undirected networks, i.e. we consider one single

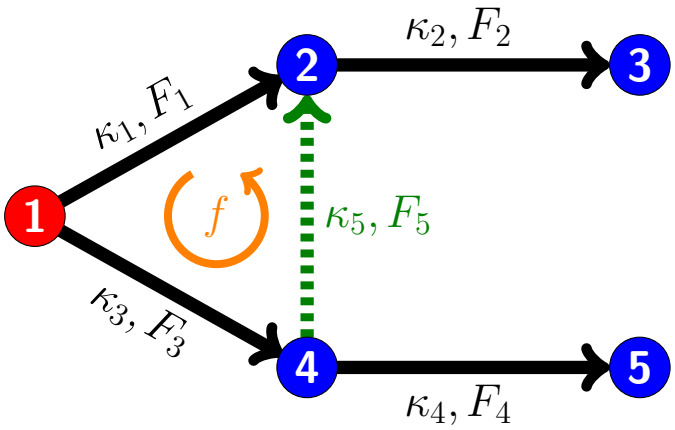



Figure IV.1.: The network scheme which is investigated in this chapter. The source (1) supplies the other nodes with variable flow via the solid black links with each having an individual capacity $\kappa_i$. In this chapter we analyse when it is useful to add the green dashed link which implements a loop.

edge which can carry flows in both directions.

Considering economical constrains, the capacities which can be built into the network are limited by

$$\sum_l \kappa_l^\gamma = K^\gamma \,, \tag{IV-1}$$

where $K^\gamma$ can be interpreted as the total budget and $\gamma$ states how the cost for adding capacity to a link scales. Typical values of $\gamma$ can be estimated by comparing the scaling of the material costs: In electrical networks, $\kappa_l$ scales with the area of the cable. Assuming the costs to be dominated by the material prize, which scales also with the area, we find $\gamma \approx 1$ for electrical power networks. In pipe flow networks, the capacity scales with the radius to forth power [49] while the material cost scales linearly with the radius. We would therefore expect $\gamma \approx 1/4$. Note that we considered so far only the material costs. In real life, however, there are additional costs for planning and installation which might scale differently so that overall, $\gamma$ can take different values. It is therefore convenient to study various $\gamma$ values. In the case $\gamma = 1$, the cost is linear to the capacity gain while for $\gamma < 1$, the gain-invest ratio increases with $\kappa_i$, i.e. it becomes cheaper to build one link with a certain capacity instead of building two links with each having half of this capacity. Vice versa, it becomes cheaper to build two smaller links instead of one large link if $\gamma > 1$. As this is scaling is not very realistic in most use-cases, we restrict this analysis to $\gamma \in [0, 1]$.

In the linear congestion model introduced in chapter II, the time $t_l(F_l)$ needed to travel along a link $l$, depends linearly on the local flow $F_l$. In chapter II we introduced the congestion parameter $\eta$ as coupling of the local flow to the travel time. In this notation, starting at $\eta = 0$ and driving $\eta \to \infty$ corresponds to the transition from the

shortest path to the Ohmic limit. In this chapter, however, we start in the Ohmic limit as this scenario was already studied for this case [40]. We introduce therefore the inverse of the congestion parameter

$$\lambda := \frac{1}{\eta} \tag{IV-2}$$

so that the time it takes to travel along an edge $l$ can be written as

$$\tau_l = t_l(\lambda + |F_l|)\,, \tag{IV-3}$$

where we divided the $\tau_l$ by the global factor $\eta$ compared to (II-1). The Ohmic limit is then reached for $\lambda = 0$. The budget constrain (IV-1) is included by setting $t_l = 1/\kappa_l$ in analogy to [40], i.e. we define the capacity as the inverse of the travel time on the empty link. When considering $\lambda > 0$, there are two convenient ways how to generalise the analysis from [40]:

In the first case, we couple the capacity $\kappa_l$ to both terms in (IV-3):

$$t_l^A = \frac{1}{\kappa_l}\left(|F_l| + \lambda\right) \tag{IV-4}$$

while in the second case, the $\kappa_l$ only affects the $F_l$ term, i.e.

$$t_l^B = \frac{1}{\kappa_l}|F_l| + \lambda\,. \tag{IV-5}$$

In the following, the analysis is done for the first case. Thus we use the short notation

$$t_l \equiv t_l^A$$

The cost function[1] $\tau$, i.e. the total travel time, is the sum over each $t_l$ weighted with the local flow strength $|F_l|$:

$$\tau = \sum_l t_l|F_l| = \sum_l \frac{1}{\kappa_l}\left(F_l^2 + \lambda|F_l|\right)\,, \tag{IV-6}$$

where the flows are arranged such that they minimise $\tau$ for a given set of $\kappa_l$. Note that a non-existing link $e$ is equivalent to $\kappa_e = 0$, as long as the network is fully connected by links with finite capacities, since it would be infinitely expensive if any flow took this link instead of an alternative path.

Furthermore, Kirchhoffs current law (II-7b) applies. In this analysis, we assume the inflows $P_i$ for each sink to be iid Gaussian random variables with mean $-\mu$ and variance $\sigma$. The source always supplies the outflows of all sinks so that the sum of all

[1] Note that the cost function is called dissipation and denoted by $D$ in [40].

inflows is always $\sum_i P_i = 0$. The inflow at the source $P_1$ can therefore be written in terms of the $P_i$ of the sinks:

$$P_1 = -\sum_{i=2}^{5} P_i \,. \tag{IV-7}$$

As the $P_i$ are fluctuating, we are interested in minimising the average cost function $\langle \tau \rangle$ over all configurations of the inflow vector.

To sum up, we seek the optimal set of non-negative $\kappa_l$ which is the solution of the following optimisation problem:

$$\min_{\kappa_l, F_l} \qquad \langle \tau \rangle = \sum_l \frac{1}{\kappa_l} \left( \langle F_l^2 \rangle + \lambda \langle |F_l| \rangle \right) \tag{IV-8a}$$

$$\text{s.t.} \quad \sum_i I_{il} \langle F_l \rangle = \langle P_i \rangle \tag{IV-8b}$$

$$\sum_l \kappa_l^\gamma = K^\gamma \,. \tag{IV-8c}$$

We want to address the following question: In which region of the parameter space over $(\gamma, \sigma, \lambda)$ is the tree network without the dashed green link 5 in fig. IV.1 optimal and where is it better to include a loop? Mathematically these phases are characterised by the optimal $\kappa_5$: if $\kappa_5 = 0$, the optimal network is a tree and if $\kappa_5 > 0$, it becomes loopy.

Note that the mean of the sink strengths $\mu$ fixes the scale and can therefore be set to $\mu = 1$ in appropriate units.

## IV.2. Optimal capacities

To determine the optimal set of capacities $\kappa_l$, we use the Lagrange multiplier method [33] to solve (IV-8). The Lagrange function then reads:

$$\mathcal{L} = \langle \tau \rangle - \xi \left( \sum_l \kappa_l^\gamma - K^\gamma \right)$$

with the Lagrange multiplier $\xi$. The optimum is found by minimising $\mathcal{L}$:

$$\partial_{\kappa_l} \mathcal{L} = \partial_{\kappa_l} \langle \tau \rangle - \xi \gamma \kappa_l^{\gamma - 1} \stackrel{!}{=} 0 \,.$$

With

$$\partial_{\kappa_l} \langle \tau \rangle = -\frac{1}{\kappa_l^2} \left( \langle F_l^2 \rangle + \lambda \langle |F_l| \rangle \right)$$

this reads

$$\frac{\langle F_l^2\rangle + \lambda\langle|F_l|\rangle}{\kappa_l^2} + \xi\gamma\kappa_l^{\gamma-1} = 0\,.$$

Solving this equation for $\kappa_l$ yields:

$$\kappa_l = \left(-\frac{\langle F_l^2\rangle + \lambda\langle|F_l|\rangle}{\xi\gamma}\right)^{\frac{1}{1+\gamma}}\,.$$

Inserting this expression for $\kappa_l$ into (IV-8c) allows us to eliminate the Lagrange multiplier $\xi$:

$$\sum_l\left(-\frac{\langle F_l^2\rangle + \lambda\langle|F_l|\rangle}{\xi\gamma}\right)^{\frac{\gamma}{1+\gamma}} = K^\gamma \quad\Leftrightarrow\quad -\xi\gamma = \left(K^{-\gamma}\sum_l\left(\langle F_l^2\rangle + \lambda\langle|F_l|\rangle\right)^{\gamma/(1+\gamma)}\right)^{\frac{1+\gamma}{\gamma}}\,.$$

Thus

$$\kappa_l = \frac{\left(\langle F_l^2\rangle + \lambda\langle|F_l|\rangle\right)^{\frac{1}{1+\gamma}}}{\left(\sum_l\left(\langle F_l^2\rangle + \lambda\langle|F_l|\rangle\right)^{\frac{\gamma}{1+\gamma}}\right)^{\frac{1}{\gamma}}}K\,. \tag{IV-9}$$

It remains therefore to find the $\langle F_l^2\rangle$ and $\langle|F_l|\rangle$.

## IV.3. Tree

For the tree, i.e. $\kappa_5 = 0$, the flows are fully determined by flow conservation (IV-8b):

$$F_1 = -(P_2 + P_3)\,, \tag{IV-10a}$$
$$F_3 = -(P_4 + P_5)\,, \tag{IV-10b}$$
$$F_2 = -P_3\,, \tag{IV-10c}$$
$$F_4 = -P_5\,. \tag{IV-10d}$$

According to appendix A.1, $F_1$ and $F_3$ are iid Gaussian random variables with mean $2\mu$ and variance $\sqrt{2}\sigma$ while $F_2$ and $F_4$ have mean $\mu$ and variance $\sigma$. Using appendix A.2, one finds:

$$\langle|F_1|\rangle = \langle|F_3|\rangle = \langle|P_2 + P_3|\rangle = \frac{2\sigma}{\sqrt{\pi}}\exp\left(-\frac{\mu^2}{\sigma^2}\right) + 2\mu\,\mathrm{erf}\left(\frac{\mu}{\sigma}\right)\,, \tag{IV-11a}$$

$$\langle|F_2|\rangle = \langle|F_4|\rangle = \langle|P_3|\rangle = \frac{2\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \mu\,\mathrm{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right)\,. \tag{IV-11b}$$

The second moments of the flows are according to appendix A.3:

$$\langle F_1^2 \rangle = \langle F_3^2 \rangle = 4\mu^2 + 2\sigma^2 \,, \tag{IV-12a}$$
$$\langle F_2^2 \rangle = \langle F_4^2 \rangle = \mu^2 + \sigma^2 \,. \tag{IV-12b}$$

Inserting (IV-11) and (IV-12) into (IV-9) gives an analytical expression for the optimal set of capacities for each parameter set $(\gamma, \sigma, \lambda)$. This lengthy but straight forward calculations are provided in appendix B.1.1.

## IV.4. Loopy network

When adding the dashed green link in fig. IV.1 to the network, an additional degree of freedom arises for the flows. This degree of freedom can be characterised by a circular flow $f$ which goes along the loop counter-clockwise and is added to the flows on the tree (IV-10). The cycle flow $f$ can take any value without violating current conservation (II-7b) as the total inflow on each node remain unchanged. The flows $\tilde{F}_l$ in the loopy network therefore read:

$$\tilde{F}_1 = F_1 - f \tag{IV-13a}$$
$$\tilde{F}_3 = F_3 + f \tag{IV-13b}$$
$$\tilde{F}_2 = F_2 \tag{IV-13c}$$
$$\tilde{F}_4 = F_4 \tag{IV-13d}$$
$$\tilde{F}_5 = f \,, \tag{IV-13e}$$

where the $F_l$ are the flows on the tree which are given by (IV-10). The detailed calculations to obtain an expression for the cycle flow $f$ are done in appendix B.1.2. We find

$$f(F_1, F_3) = \begin{cases} f_a(F_1 - F_3) & , \; F_1, F_3 \geq 0 \\ f_b(F_3, F_1 - F_3) & , \; F_3 \geq 0 > F_1 \\ -f_b(F_3, F_1 - F_3) & , \; F_1 \geq 0 > F_3 \\ -f_a(F_1 - F_3) & , \; 0 > F_1, F_3 \end{cases} \tag{IV-14}$$

with

$$f_a(\Delta F) = \frac{\Delta F - \operatorname{sgn}(\Delta F)\, \Delta F_c}{\rho} \Theta\left(|\Delta F| - \Delta F_c\right)$$

and

$$f_b(F_3, \Delta F) = \begin{cases} f_b^{\alpha}(F_3, \Delta F) & , \; \Delta F < -2F_3 \\ f_b^{\beta}(F_3, \Delta F) & , \; \Delta F \in [-2F_3, -F_3] \end{cases}$$

$$f_b^{\alpha}(F_3,\Delta F)=\frac{1}{\rho}\begin{cases}\Delta F+\Delta F_c & ,\ \Delta F\in(-\infty\ ,\ -\rho F_3-\Delta F_c]\\ -\rho F_3 & ,\ \Delta F\in[-\rho F_3-\Delta F_c\ ,\ -\rho F_3-\Delta F_c+\lambda]\\ \Delta F+\Delta F_c-\lambda & ,\ \Delta F\in[-\rho F_3-\Delta F_c+\lambda\ ,\ -2F_3]\end{cases}$$

$$f_b^{\beta}(F_3,\Delta F)=\frac{1}{\rho}\begin{cases}\Delta F+\Delta F_c & ,\ \Delta F\in[-2F_3\ ,\ \frac{\Delta F_c-\rho F_3}{\rho-1}]\\ \rho(\Delta F+F_3) & ,\ \Delta F\in[\frac{\Delta F_c-\rho F_3}{\rho-1}\ ,\ \frac{\Delta F_c-\rho F_3+\lambda}{\rho-1}]\\ \Delta F+\Delta F_c-\lambda & ,\ \Delta F\in[\frac{\Delta F_c-\rho F_3+\lambda}{\rho-1}\ ,\ -F_3]\end{cases}$$

To determine the optimal capacities according to (IV-9), it remains to find the mean values of the absolutes and squares of the flows. They can be partially evaluated like in the tree case:

$$\langle|\tilde{F}_1|\rangle=\langle|F_1-f|\rangle \tag{IV-15a}$$
$$\langle|\tilde{F}_3|\rangle=\langle|F_3+f|\rangle \tag{IV-15b}$$
$$\langle|\tilde{F}_2|\rangle=\langle|F_2|\rangle=\frac{2\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\mu^2}{2\sigma^2}\right)+\mu\,\mathrm{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right) \tag{IV-15c}$$
$$\langle|\tilde{F}_4|\rangle=\langle|F_4|\rangle=\frac{2\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\mu^2}{2\sigma^2}\right)+\mu\,\mathrm{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right) \tag{IV-15d}$$
$$\langle|\tilde{F}_5|\rangle=\langle|f|\rangle \tag{IV-15e}$$

$$\langle\tilde{F}_1^2\rangle=\langle(F_1-f)^2\rangle=\langle F_1^2\rangle+\langle f^2\rangle-2\langle fF_1\rangle \tag{IV-15f}$$
$$\langle\tilde{F}_3^2\rangle=\langle(F_3+f)^2\rangle=\langle F_3^2\rangle+\langle f^2\rangle+2\langle fF_3\rangle \tag{IV-15g}$$
$$\langle\tilde{F}_2^2\rangle=\langle F_2^2\rangle=\mu^2+\sigma^2 \tag{IV-15h}$$
$$\langle\tilde{F}_4^2\rangle=\langle F_4^2\rangle=\mu^2+\sigma^2 \tag{IV-15i}$$
$$\langle\tilde{F}_5^2\rangle=\langle f^2\rangle \tag{IV-15j}$$

The values of $\langle|F_1-f|\rangle$, $\langle|F_3+f|\rangle$, $\langle|f|\rangle$ and $\langle f^2\rangle$ are still to be found. In principle, it would be possible to find analytical expressions for them by straight forward integration. However, there are a lot of lengthy terms and many cases which need to be considered separately. It would therefore be very lengthy calculations with high chance for arithmetical errors. It is therefore more convenient to use a numerical solver which computes these averages over a large set of random $F_1$ and $F_3$ according to the sink properties.

Note that for symmetry reasons, $\langle|F_1-f|\rangle=\langle|F_3+f|\rangle$ as we obtain the same graph when exchanging the upper and lower branch in fig. IV.1 and replacing $(F_1,F_3,f)\to(F_3,F_1,-f)$.

## IV.5. Numerical solver

With the results from the previous section, we can now implement a numerical solver to find the optimal capacities. Our aim is to find the optimal set of $\kappa_l$ which are minimising the average travel time

$$\langle \tau \rangle = \sum_l \frac{1}{\kappa_l} \left( \langle F_l^2 \rangle + \lambda \langle |F_l| \rangle \right)$$

with mean values given by (IV-15).

From the symmetry of the graph we can conclude that $\kappa_1 = \kappa_3$ and $\kappa_2 = \kappa_4$. From the budget constrain

$$\sum_{l=1}^{5} \kappa_l^\gamma = K^\gamma$$

we can derive an expression for $\kappa_2$ in terms of the other capacities:

$$\kappa_2(\kappa_1, \kappa_5) = \left( \frac{K^\gamma - 2\kappa_1^\gamma - \kappa_5^\gamma}{2} \right)^{\frac{1}{\gamma}} . \tag{IV-16}$$

As remaining degrees of freedom it remains to find the optimal $\kappa_1$ and $\kappa_5$. From the boundary condition we can also directly derive the upper limit of $\kappa_5$ for a given $\kappa_1$:

$$\kappa_{5,max}(\kappa_1) = \left( K^\gamma - 2\kappa_1^\gamma \right)^{\frac{1}{\gamma}} . \tag{IV-17}$$

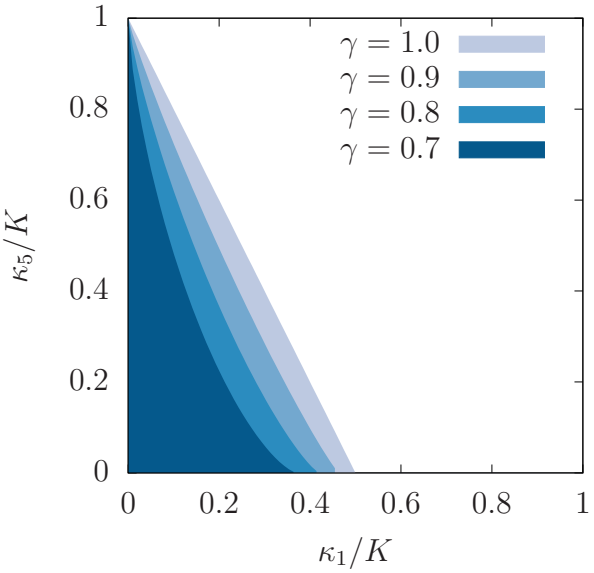



Figure IV.2.: The valid parameter space of $\kappa_1$ and $\kappa_5$ for different $\gamma$ values

As the cycle flow $f$ depends explicitly on $\kappa_1$ and $\kappa_5$, we have to solve a system of coupled non-linear equations. To solve this problem, the easiest way is to scan the $\kappa_1$-$\kappa_5$ plane. In Fig. IV.2, the coloured areas mark the valid parameter range for different values of $\gamma$. An appropriate strategy is to scan in a first run the entire parameter range. From this scan with low resolution we can estimate already the approximate area of the minimum so that we can scan around this area with a much higher $\kappa$ resolution to precisely determine the position of the minimum.

For each point, we need to compute the averages $\langle|F_1 - f|\rangle$, $\langle|f|\rangle$ and $\langle f^2\rangle$ by taking a set of random values of $F_1$ and $F_3$, compute $|f|$, $f^2$ and $|F_1 - f|$ for each of them and average in the end. With these averages, we can calculate the average travel time $\langle\tau\rangle$ for this choice of $\kappa_1$ and $\kappa_5$. Note that $F_1$ and $F_3$ are iid Gaussian random variables with mean $2\mu$ and variance $\sqrt{2}\sigma$ according to appendix A.1.

To make the result easier to interpret, we divide $\langle\tau\rangle$ by the average travel time of the optimal tree network $\langle\tau_{tree}\rangle$, which we can directly compute using the anayltical expression (IV-9) and the results of sec. IV.3. Defining the ratio

$$r := \frac{\langle\tau\rangle}{\langle\tau_{tree}\rangle}\,, \tag{IV-18}$$

we can easily read off whether the optimal network is loopy or not: if $r < 1$, inserting a loop reduces the average travel time while $r > 1$ means that the optimal tree would be better. For numerical reasons, the case $\kappa_5 = 0$ can't be computed with the above method as $\rho$ and $\Delta F_c$, which are used to compute $f$, diverge in this case. However, as the set of networks with $\kappa_5 = 0$ is the subset of tree-like networks, we obtain the optimal parameters on this subset by exploiting the analytical solving method for the tree.

We have the freedom of choosing natural units for flow intensities and the budget. It is convenient to choose $\mu = 1$ as natural flow intensity and $K = 1$ as unit of the budget. Therefore the remaining parameters which need to be analysed are the fluctuation strength $\sigma$, the budget parameter $\gamma$ and the congestion parameter $\lambda$. Note that for $\lambda = 0$, these parameters are already discussed in [40] with another convention of $\sigma' = \sqrt{2}\sigma$.

For each point in the $(\sigma, \gamma, \lambda)$-plane, a scan over the $\kappa_1$-$\kappa_5$-plane must be done and the minimal value of $r(\kappa_1, \kappa_5)$ indicates the optimal capacities.

To get a measure of the numerical fluctuations of the results, several scans with different sets of random numbers $F_1$ and $F_3$ are performed for each point $(\sigma, \gamma, \lambda)$. For each scan, the minimal $r$ and the corresponding optimal parameters $\kappa_1$, $\kappa_5$ and $\kappa_2$ are computed. By averaging over the results of all scans, we obtain a mean value and a variance for each quantity.

# IV.6. Results

The numerical solver introduced in sec. IV.5 enables us to investigate the optimal network capacities for in dependency of the sink fluctuation strength expressed by $\sigma$, the cost parameter $\gamma$ and the congestion parameter $\lambda$. The analysis focus on the following points:

1. comparison of numerical results for $\lambda = 0$ with analytical results from [40] to check the precision,

2. evolution of the optimal ratio $r(\gamma)$ and capacities $\kappa_1(\gamma)$, $\kappa_5(\gamma)$ and $\kappa_2(\gamma)$ for different $\lambda$ and fixed $\sigma$

3. derivation of the phase diagrams for different $\lambda$ indicating the regions in the $(\sigma, \gamma)$-plane where the optimal network is loopy or a tree. In general, loops emerge for large values of $\gamma$ and $\sigma$ [46]. We will thus focus on the critical value of the budget parameter $\gamma_c$, marking the transition between the tree-like and the loopy phase, in dependence of $\sigma$ and $\lambda$.

## IV.6.1. Comparison with analytical results

Before analysing the data generated by the numerical solver, we compare the results for $\gamma_c$ at $\lambda = 0$ with the results of the analytical solution used for the Ohmic case in

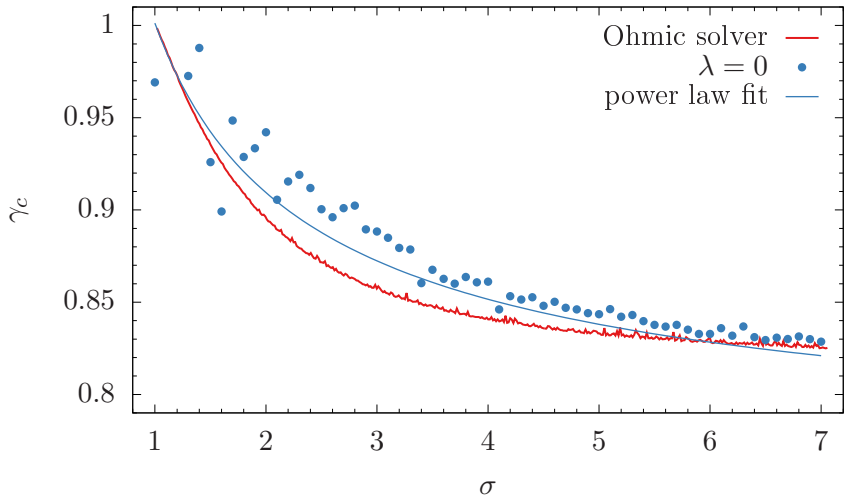



Figure IV.3.: Decay of critical cost parameter $\gamma_c$ where a single loop starts to become beneficial with fluctuation strength $\sigma$ in the Ohmic limit. Comparison of numerical results with the analytical solution from [40]

[40]. Mathematically, these cases are equivalent so that we should get identical results up to numerical errors. In fig. IV.3, these results are compared[2].

We find qualitatively similar results. In both cases, the $\gamma_c$ follow a power law like curve with seem to saturate for large $\sigma$ somewhere around $\gamma_c \approx 0.8$. However, we can observe that the $\gamma_c$ from the Ohmic solver saturate faster then the numerical results. We can fit a power law to our results and find the optimal fit parameter

$$\gamma_c(\sigma) = (0.2476 \pm 0.0064)\sigma^{-(0.669\pm0.038)} + (0.7537 \pm 0.0074)$$

which are plotted in fig. IV.3. In our data set, the $\gamma_c$ converge therefore to $\approx 0.75$.

An explanation for the mismatch could be the different methods used to determine $\gamma_c$ in both scenarios. Inaccuracies in the numerical method of finding the optimal capacities might also cause this gap as it is hard to mimic accurately Gaussian random distributions with large standard deviation $\sigma$.

A possibility to wipe out this discrepancy could be to determine $\gamma_c$ not by making a linear fit to the $r(\gamma)$ data points in the loopy regime but to determine the $\gamma$ value where the optimal $\kappa_5$ is jumping from a finite value to a value close to zero. This could also increase the accuracy where $\gamma_c \approx 1$ as the linear fit method is more susceptible to numerical errors as the fit range becomes very small so that small numerical fluctuations can cause large errors in the fit parameters.

### IV.6.2. Optimal capacities in dependence of $\gamma$ and $\lambda$

How are the optimal capacities influenced when driving the budget parameter $\gamma$ or the congestion parameter $\lambda$? We will discuss this question in the following for fixed value of the standard deviation $\sigma = 3$.

The optimal parameter for each $(\sigma, \lambda, \gamma)$ point are plotted in fig. IV.4. The parameter $\kappa_1$ and $\kappa_5$ are read off the point with minimal $r$ out of all scanned points. The corresponding $\kappa_2$ value was calculated by inserting the means of $\kappa_1$ and $\kappa_5$ into (IV-16). The errorbars for $r$, $\kappa_1$ and $\kappa_5$ in fig. IV.4a-c are the variances of the corresponding optimal values over 10 runs. For each run, $10^6$ random values for $F_1$ and $F_3$ were generated and averaged over. The resolution of the scan in $\kappa_1$ and $\kappa_5$ is $\Delta\kappa = 0.001$.

Considering the curves for $r(\gamma)$ in fig. IV.4a, we find a linear relation in the loopy phase, i.e. the region of $r(\gamma) < 0$. To find the critical budget parameter $\gamma_c$ we therefore fit a linear curve

$$r_\lambda(\gamma) = a_\lambda + b_\lambda \gamma$$

[2]Note that we need to redefine the $\sigma$ from [40] by $\sigma \mapsto \sigma/\sqrt{2}$ to match with our definition of $\sigma$.

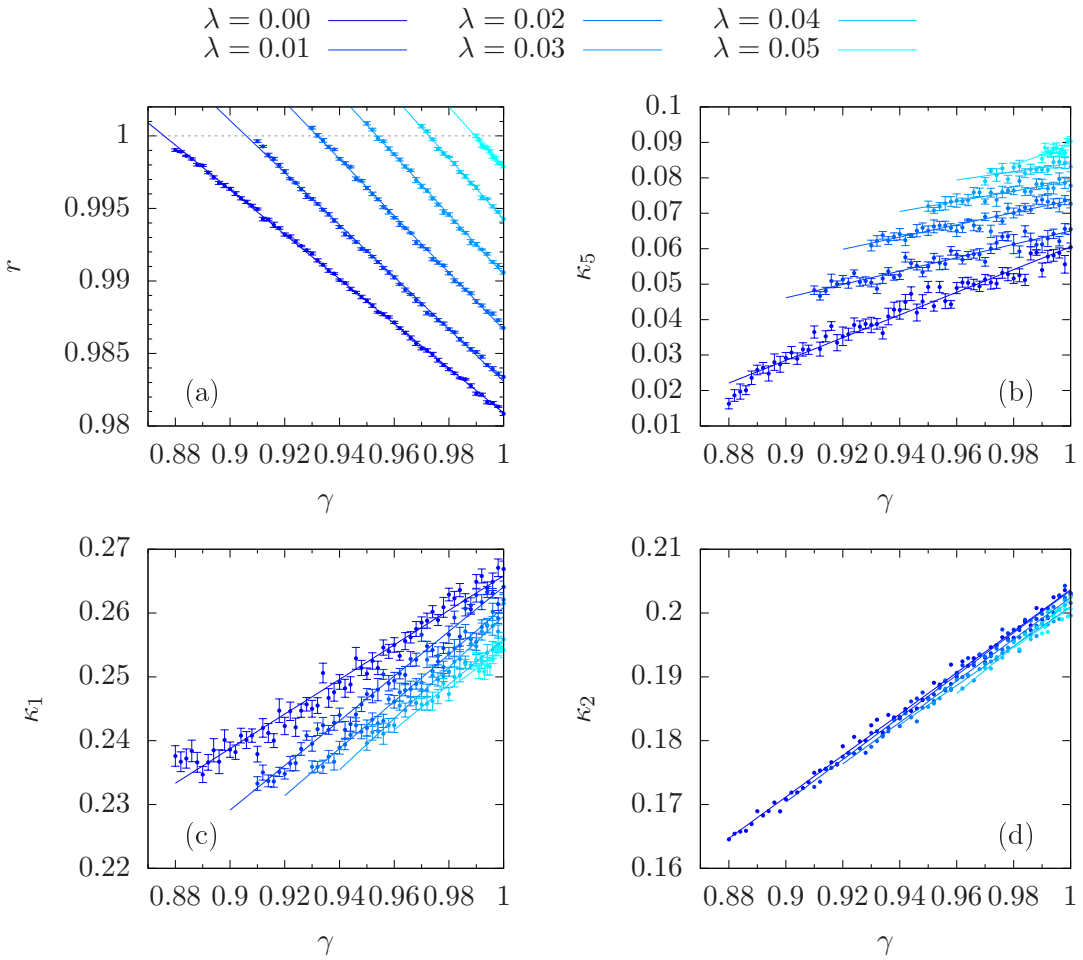



Figure IV.4.: Optimal parameters for fixed fluctuation strength $\sigma = 3$ in dependence of the budget parameter $\gamma$ for different values of the congestion parameter $\lambda$.

for each $\lambda$ value and estimate $\gamma_c$ as the crossing point of the fit with the $r = 1$ axis:

$$\gamma_c = \frac{1 - a_\lambda}{b_\lambda} . \tag{IV-19}$$

Note that this method of determining $\gamma_c$ differs from [40].

We find that $\gamma_c$ increases with increasing $\lambda$, hence it becomes attractive to build a loop only for very high $\gamma$. This observation is also what we would expect as the limit $\lambda \to \infty$ corresponds to the shortest path dynamic, where redundancies like a loop are not useful.

When considering the optimal capacity of the loop, $\kappa_5$, plotted in fig. IV.4b, we again find linear $\kappa_5(\gamma)$ curves for $\gamma > \gamma_c$. At $\gamma = \gamma_c$, we observe a discontinuous phase transition meaning that the optimal $\kappa_5$ makes a jump from $\kappa_5 = 0$ in the

non-loopy phase to a finite $\kappa_5$ value in the loopy phase. This phase transition for $\lambda = 0$ was already observed in [40]. In fig. IV.4b we find that the transition is also discontinuous for $\lambda > 0$. The jump size even grows with $\lambda$, i.e. the minimum non-zero $\kappa_5$ value becomes larger for increasing importance of congestion.

The budget, which is invested into additional capacities of the loop, is mainly taken from $\kappa_1$ while $\kappa_2$ only slightly decreases with $\lambda$, as fig. IV.4c-d show.

# IV.7. Optimal capacities for different $\sigma$

## IV.7.1. Phase diagram

It remains to discuss the phase diagram which shows in which part of the parameter space spanned by $\sigma$, $\lambda$ and $\gamma$ the loop occurs and where the optimal network is a tree. Instead of drawing a 3D plot, we will split the discussion in two parts: First we discuss the critical curve $\gamma_c(\lambda)$ for different $\sigma$ values before we investigate how the curve $\gamma_c(\sigma)$ changes with $\lambda$.

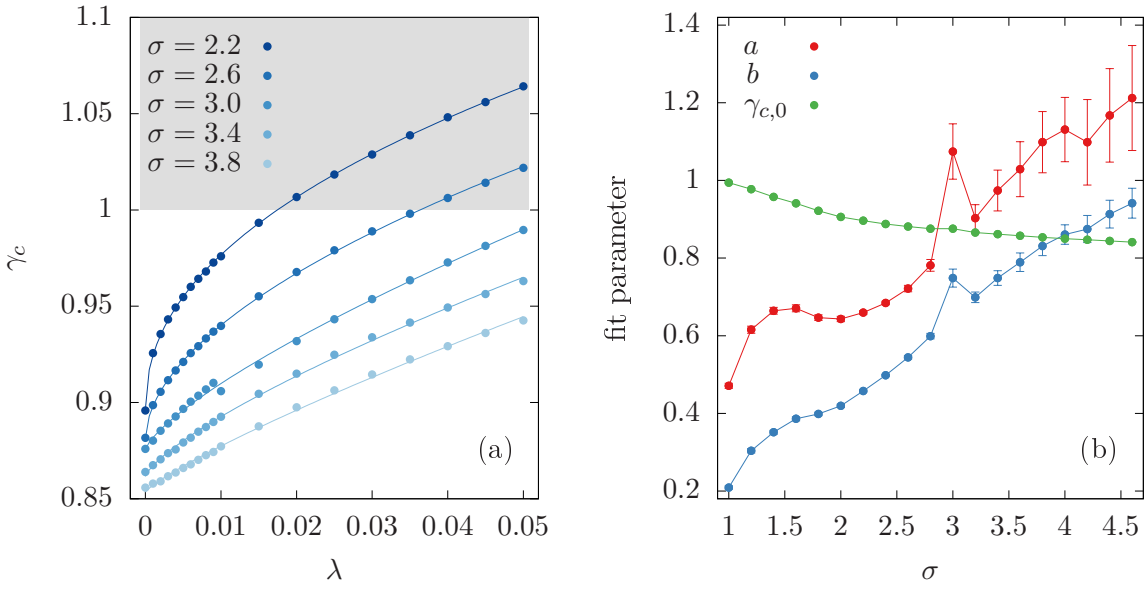



Figure IV.5.: Dependency of the critical $\gamma_c$ on congestion, plotted in (a). The grey area marks $\gamma > 1$ which has no physical meaning. The $\gamma_c(\lambda)$ are fitted for each $\sigma$ separately to a power law (IV-20). The corresponding fit parameters and their uncertainties are plotted in (b).

#### IV.7.1.1. Critical value $\gamma_c(\lambda)$ for different $\sigma$

In fig. IV.4 we found that the critical value $\gamma_c$, where the network topology changes from tree-like to loopy, increases with $\lambda$. We will now discuss in more details the relation between those parameters.

In fig. IV.5a, the critical value $\gamma_c$ is plotted over $\lambda$ for different values of $\sigma$. The slope of the lines diverges for small $\lambda$ and saturate for large $\lambda$. This is typical for a power law

$$\gamma_c(\lambda) = a\lambda^b + \gamma_{c,0} \tag{IV-20}$$

with an exponent $b < 1$. When fitting (IV-20) to each $\gamma_c(\lambda)$ data set, we get very accurate fits to the data as we can see in fig. IV.5a, where the dots are the observed data points and the lines the corresponding power law fits. Note that $\gamma_{c,0}$ is the critical value $\gamma_c$ in the Ohmic case.

In fig. IV.5b, the corresponding fit parameters are plotted as a function of $\sigma$ with their uncertainties[3]. We find an almost linear increase with $\sigma$ for the parameters $a$ and $b$. An outlier is found at $\sigma = 3$, which is attributed to a small step in the curve of $\gamma_c(\gamma)$ between $\lambda = 0.009$ and $\lambda = 0.01$. The reason for this is unclear.

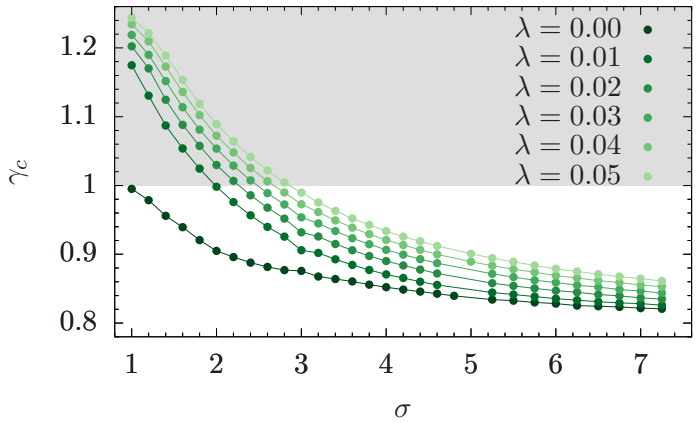



Figure IV.6.: Phase diagram with the critical value $\gamma_c(\sigma)$ for different $\lambda$ values. The area $\gamma_c > 1$ is highlighted grey as this parameter range has no physical relevance.

#### IV.7.1.2. Critical value $\gamma_c(\sigma)$ for different $\lambda$

In fig. IV.6 the critical value $\gamma_c(\sigma)$ is plotted for different values of $\lambda$. We see that the curve is shifted upwards with increasing $\lambda$ with a stronger shift for small $\sigma$. We find also a stronger impact for smaller values of $\lambda$ which we would expect due to the power laws observed in the previous section.

We can conclude that the size of the region in the parameter plane $(\sigma, \gamma)$ were the optimal network is loopy shrinks with increasing importance of congestion. In particular, when restricting to physically relevant values of the budget parameters $\gamma \leq 1$, we observe a significant increase of the value of $\sigma$ for which $\gamma_c = 1$. That means when increasing $\lambda$, we can observe that the minimal fluctuation strength $\sigma$, for which the optimal network becomes loopy, grows strongly.

## IV.8. Summary & Outlook

We investigated a simple network with two branches supplied by a central node and scanned the parameter space to determine in which cases it becomes useful to implement a loop by adding an additional link between node 2 and 4 in fig. IV.1. We generalised the discussion from [40] by adding linear terms to the cost function. This modification of the model significantly increases the complexity of solving the problem analytically so that we computed the average values of the flows numerically. In sec. IV.6.1 we found that the numerical results reproduce the analytical findings in the Ohmic case in general. Deviations between analytical and numerical results can be explained by the different underlying methods to determine the critical value $\gamma_c(\sigma)$.

We found the general trend that for increasing importance of the linear term in the cost function, i.e. the further we are going towards the shortest path limit, the critical curve $\gamma_c$ is shifted upwards. We observed a power law for $\gamma_c(\lambda)$ which causes stronger impacts for small $\lambda$. Furthermore we found that the impact of $\lambda$ on the critical value $\gamma_c$ decreases for stronger fluctuations $\sigma$.

Considering the optimal capacities, we observed that the capacity on the additional link $\kappa_5(\gamma)$ gets larger with increasing $\lambda$ leading to a larger jump at the phase transition. The extra budget invested in $\kappa_5$ is mainly taken from $\kappa_1$ leaving the capacities of the outer branches $\kappa_2$ almost unaffected by $\lambda$.

The analysis done in this chapter can easily be adapted to a class of more general networks by replacing the nodes 3 and 5 by arbitrary, identical subgraphs which are only connected with each other via node 1 and the optional loop.

[3] which correspond to the fit parameter errors given by the *GNUplot* fit function

# V. Optimal network structures for multimodal transport with a single destination

The design of public transport systems plays an important role for their efficiency which strongly affects the acceptance and usage of the service. Thus the geometry of the network should be chosen such that the average travel time for individual users to get from their starting position to their destination should be minimised. Facing finite budgets which usually limit the network size, the optimisation of the network needs to be carried out considering economical boundary conditions. Furthermore, the flow in urban transport networks are prone to congestion.

In this chapter, we thus examine a multimodal model which generalises the linear congestion model introduced in chapter II by adding a second mode of transport, namely by allowing the traveller also to walk all over the plane. We include the economical boundary conditions by assuming the network length $L$ to be fix by the budget.

We will examine in this thesis the distance-to-centre problem, that is we assume the destinations for all travellers to be the city centre. Hence, we seek the geometry with minimises the total travel time for all inhabitants to reach the city centre from their individual starting positions. This assumption simplifies the analysis significantly as we do not have to deal with multiple flow layers here (cf. discussion in section II.1). Nevertheless, we can already draw general conclusions on optimal geometries with this case study. Thus we ask what is the optimal geometry of the network for a given network length $L$ and with congestion?

In this thesis, we restrict to radial symmetric cities and compare different fundamental network geometries which are frequently observed in real world subway networks. In sec. V.1, we introduce the multimodal congested flow model which allows us to find the optimal shape for a parameterised geometry. For a simple star network, we find closed form solutions in sec. V.2 before we discuss the appearance of a loop in sec. V.3 and the branching of lines in sec. V.4.

# V.1. A model for congested flow networks

The central objective of this chapter is the structure of optimal transportation networks in the presence of congestion: Given a limited budget, what is the optimal shape of the network such that the overall travel time is minimised? In the following, we will formalise this optimisation problem and introduce several key methods to solve it.

## V.1.1. The objective function

Consider a city that is modelled as a two dimensional area, such that each point in the city can be described by a vector in the plane $\mathbf{x} = (x_1, x_2)^\intercal \in \mathbb{R}^2$. Assume that the population density in the city is described by the function $\rho(\mathbf{x})$. Now assume that the inhabitants want to travel from their home place to other places, which are described by a distribution of destinations $\rho_d(\mathbf{y}) = \rho_d(\mathbf{y}|\mathbf{x})$ which might differ for each starting position $\mathbf{x}$. We may then calculate the average travel time $\tau$ for all journeys in the whole city by integrating over the destinations and population density

$$\tau = \int d\mathbf{x}\, d\mathbf{y}\, \tau(\mathbf{x}, \mathbf{y})\, \rho(\mathbf{x})\, \rho_d(\mathbf{y}|\mathbf{x}) \tag{V-1}$$

where $\tau(\mathbf{x}, \mathbf{y})$ is the travel time between two points $\mathbf{x}$ and $\mathbf{y}$. Throughout this thesis, we assume that all travellers have the same destination, say they want to go to the city centre, which is described by the density function $\rho_d(\mathbf{y}|\mathbf{x}) = \delta(\mathbf{y})$. The travel time is thus given by

$$\tau = \int d\mathbf{x}\, \tau(\mathbf{x}, \mathbf{0})\, \rho(\mathbf{x}) \tag{V-2}$$

This expression for the total travel time $\tau$ is the central objective to be minimised. The solution of this optimisation problem depends on the properties of the city (via the function $\rho(\mathbf{x})$), the methods of transportation and the budget constraints which are discussed in the following.

## V.1.2. Travel time in multimodal traffic networks

The travel time $\tau$ is essentially determined by the available modes of transport and their velocities. In this chapter, we assume that there are two modes of transports: People can either walk or use a transportation network such as a subway network. Typically, people will have to use both modes of transport, first walking to the network and then travelling along the network. Hence, the travelling time to the centre $\tau(\vec{x}, \vec{0})$ is the sum of the travelling time along both modes which we will now discuss in detail.

First, people can walk in the plane between any two points with a constant velocity $v_w$, which we set to 1 in appropriate units. Assuming that people can walk directly and that no congestion applies here, the walking time between two points $\vec{x}$ and $\vec{y}$ is simply given by $\|\vec{x} - \vec{y}\|/v_w$, where $\|\cdot\|$ denotes the euclidean distance.

Second, people may choose a transportation network, for example a subway network, with a velocity $v_s$ which is typically much faster than walking $v_s \gg v_w = 1$. In this thesis, we are especially interested in the impact of congestion on the total travelling time and the structure of optimal transportation network. Congestion is taken into account up to linear order like in chapter II. That is, the travelling time or equivalently the inverse velocity is assumed to be an affine linear function of the flow $F$,

$$\frac{1}{v_s(F)} = a + bF\,. \tag{V-3}$$

Hence, the velocity is not constant, but can be different for every part on the network. The total travelling time of persons along the network is then obtained by integrating $1/v_s$ along the path on the network.

The model introduces two abstract parameters $a$ and $b$, where $a$ is the inverse of the average velocity when the network is empty, while $b$ is the congestion parameter which determines the coupling strength between local flow and velocity[1]. For a subway network, a typical value for $a$ would be $a \approx 1/8$, when we assume $v_w \approx 5\,\mathrm{km/h}$ and $v_a \approx 40\,\mathrm{km/h}$. For the congestion parameter $b$ it is not straightforward to define typical values. However, one can find constraints: Firstly, it needs to be positive, $b \geq 0$, since the subway would be faster with higher load otherwise. On the other hand, one can consider the critical flow $F_{\mathrm{crit}}$ at which the subway's average velocity becomes as slow as the walking speed (again with $v_w = 1$):

$$v_s(F_{\mathrm{crit}}) = \frac{1}{a + bF_{\mathrm{crit}}} = 1 \quad \Leftrightarrow \quad F_{\mathrm{crit}} = \frac{1-a}{b}$$

Since the flow is bounded, $F \in [0,1]$, the subway can slow down to the walking speed for very strong flows if $b \approx 1 - a$. Since such a scenario is hardly imaginable to lay in the linear range, we will mainly focus on $b \ll 1 - a \approx 1$ in the following.

The flow $F$ inside the subway network is arranged such that the average travel time is minimal. In the limit $b = 0$, the velocity in the network is constant and independent of the flow, thus all travellers are driving along the shortest path from their access point to the centre. In the limit $b \to \infty$, the flow pattern is identical to Ohmic flow networks such as in electrical power grids (cf. II.3). For intermediate congestion, the flow inside the network is governed by the linear congestion model introduced in sec. II.1. Note that for treelike networks there is only one way travellers can take.

[1] Note that in the notation used in chapter II, $a \mathrel{\hat{=}} t$ and $b \mathrel{\hat{=}} \eta t$.

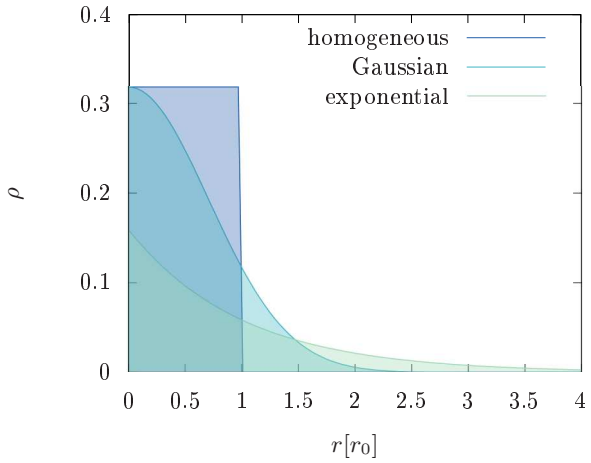



Figure V.1.: Comparison of the radial density functions considered in this chapter.

### V.1.3. The population density of cities

In this article we focus on radially symmetric population densities, in particular we analyse

1. a homogeneous disk $\rho_1(r) = \rho_0\Theta(r_0 - r)$ with constant density up to a radius of $r_0$,
2. a Gaussian density $\rho_2(r) = \rho_0 \exp(-r^2/r_0^2)$ and
3. an exponential density $\rho_3(r) = \rho_0 \exp(-r/r_0)$.

Using appropriate units, the factors $\rho_0$ are chosen such that the overall population is normalised to $\int d\mathbf{x}\, \rho(\mathbf{x}) = 1$. In Fig. V.1, these functions are plotted. The factor $r_0$ classify the typical scale of the city. As we do not have any other spacial reference points we can set $r_0 = 1$ and express all lengths in units of this parameter.

For all three cases we can easily evaluate the travelling times in the absence of a public transportation network

$$\tau_0 = 2\pi \int_0^\infty dr\, r\, \rho(r) = \begin{cases} \frac{2}{3} r_0 & \text{homogeneous disk} \\ 2r_0 & \text{Gaussian} \\ \frac{\pi}{2} r_0 & \text{exponential} \end{cases} . \tag{V-4}$$

The ratio $\hat{\tau} = \tau/\tau_0 \in [0, 1]$ quantifies the reduction of the total travelling time by the network and thus quantifies its effectiveness. This measure will be used as a measure of network performance in all following discussions and plots.

## V.1.4. Parametrising and optimising transportation networks

To optimise the travel times in multimodal flows, we have to make some assumptions about the transportation network to keep the problem feasible. In particular, we consider a set of three different rotational symmetric geometries which are frequently observed as pattern in urban transport networks. In each case we introduce the budget constraint by fixing the total length $L$ of the network. We then determine the optimal network structure as a function of $L$.

We first consider a star network with $n$ branches as illustrated in Fig. V.3. Each branch has the same length $l_b = L/n$ and the angle between neighbouring branches is always $\theta_n = 360°/n$. The high symmetry of this geometry allows to compute the travelling time and the optimal number of branches $n^*$ analytically. The computation will be carried out in section V.2 and will serve as a reference for different network topologies.

We then consider fundamental modifications of this elementary network structure, which are frequently present in actual public transportation networks. In many cities, one observes the addition of a loop around the city centre and the branching of tracks. In both cases we have to introduce further parameters, such as the radius of the loop, in addition to the number of branches $n$ as illustrated Fig. V.2. In these cases we will employ a versatile numerical method to solve the optimisation problem which is described in the following.

In summary, we consider three topologies as illustrated in Figures V.3 and V.2:

1. A star network parameterised by the number of branches $n$
2. A star network with $n$ branches and a loop with radius $r_l$
3. A star network with $n$ branches which split up into two subbranches under an angle $\alpha$ at radius $l_1$ and both having the same length.

## V.1.5. Routing and numerical optimisation

The numerical optimisation of transportation networks includes two different, but strongly intertwined steps. (i) Optimisation: The parameters characterising the network structure are varied to find the optimal values which lead to a minimum total travel time $\tau$. (ii) Routing and evaluation: Given a certain network structure, individual travellers choose a path from their starting point to the centre which yields $\tau(\vec{x}, \vec{0})$. Aggregating over all starting points $\vec{x}$ then yields $\tau$.

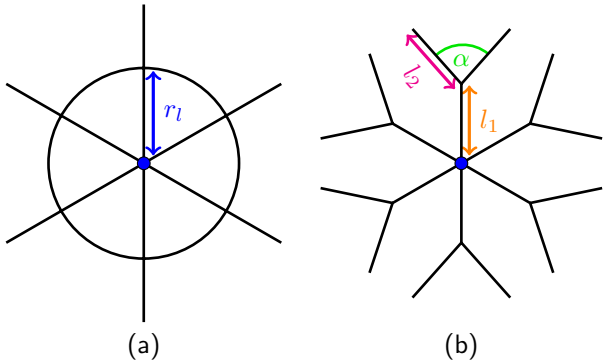



Figure V.2.: Sketch of the geometries investigated numerically:
(a) Star network with $n$ branches and a loop of radius $r_l$. The two parameters $n$ and $r_l$ are optimised numerically.
(b) star network with $n$ branches, which split at a radius $l_1$ into two branches with an angle $\alpha$. Numerical optimisation yields the optimal values of the parameters $n$, $l_1$ and $\alpha$ with $l_2 = (L - nl_1)/2n$.

We first discuss the optimisation step (i). The optimal parameters characterising the network structure, such as the number of branches $n$ or the radius of a loop $r_l$ are determined by a scan of the parameter space. We developed a versatile method to evaluate the average travel time $\tau$ using a discretisation of both the network and the starting points. In particular the solver proceeds as follows:

1. Draw $N$ starting points at random according to the population density $\rho(\vec{x})$.

2. Discretise the network: Generate the stations according to the given network geometry. Stations are created after each interval $\Delta l$ along the branches plus an additional station at the end to ensure the branch has the correct length.

3. Routing: Compute the optimal path to the centre for each starting point. This is discussed in detail below.

4. Compute the flow $F$ and the resulting velocity $v_s(F)$ for every segment of the network. The flow $F$ is directly proportional to the number of travellers using this segment of the network.

5. Sum up the travelling times for each starting point to obtain $\tau$.

Further details of the numerical solver are described in appendix C.

In the routing step (ii) we have to determine the path to the centre for every starting point. Generally, travellers aim to either optimise times or distances, but their objectives and preferences may differ in details. In this article we consider two different routing strategies

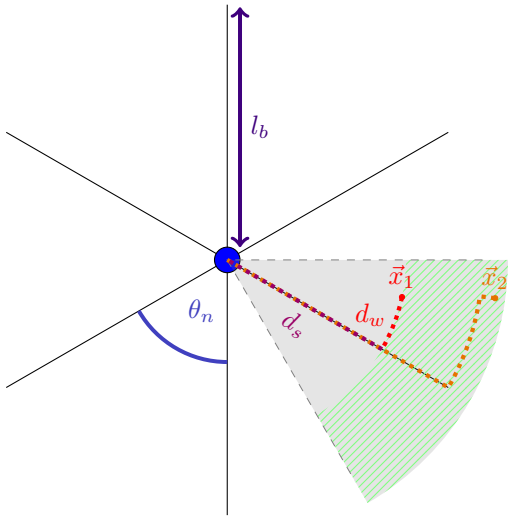



Figure V.3.: Sketch of a star network. The star consist of $n$ branches of length $l_b = L/n$ with an angular distance of $\theta_n = 2\pi/n$. To evaluate the total travel time $\tau$ we can exploit the symmetry and focus on one branch the surrounding segment (for instance the grey area). Travellers starting at a distance $|\vec{x}| \leq l_b$ (e.g. $\vec{x}_1$) walk spherically to the next branch and then travel via the network. Travellers starting at a distance $|\vec{x}| > l_b$ first walk radially inwards until $l_b$ and then spherically to the next branch of the network. All travellers living inside the green hatched area contribute to the flow on the transportation network at distance $d_s$.

(a) We first consider the case where travellers minimise their walking distance. That is, the traveller always walks to the closest station to access the network at this point. This will be referred to as the "lazy traveller model" in the following.

(b) Furthermore, we consider the case where a traveller minimises the overall travelling time to the centre. Such a traveller may accept a longer walking distance if this reduces his overall travel time. We assume that the traveller considers the uncongested velocity in the network to choose his path. This will be referred to as the "fast traveller model" in the following.

The essential step is the choice of the optimal entry station to the network, which may be different for the two cases. Further details are described in appendix C.

## V.2. Optimal star networks

We first analyse the travelling time and optimal geometry of a star network without branching and loops. This case admits a closed form solution for the case that

travellers always walk directly to the transportation network ("lazy traveller model", cf. step 3(a) in section V.1.5). Hence, the impact of congestion can be elucidated analytically. The geometry is optimised by choosing the number of branches $n$ appropriately for a fixed total length $L$ of the network.

So consider a star network with $n$ branches of length $l_b = L/n$ as depicted in Fig. V.3. The main step in the solution of the optimisation problem is the computation of $\tau(\mathbf{x}, \mathbf{0})$ given a radially symmetric population density $\rho(r)$. We assume that the network is accessible everywhere, i.e. we do not discretise this network as in the numerical approach described in Sec. V.1.5. Furthermore, we assume that each traveller takes the shortest possible way towards the network (the "lazy traveller model", step 3(a) in Sec. V.1.5) following a radial or spherical path (no diagonal walking paths).

In the following we denote the starting point as $\vec{x} = (x_1, x_2)$ in Cartesian coordinates or by the radius $r$ and the azimuth $\theta$ in polar coordinates. Furthermore, we can exploit the symmetry of the problem and consider only a single branch at $\theta = 0$ and travellers starting in the interval $\theta \in [-\theta_n/2, +\theta_n/2]$. Then the travel time of a single traveller is written as

$$\tau(\mathbf{x}, \mathbf{0}) = \frac{d_w(\mathbf{x})}{v_w} + \int_0^{d_s(\mathbf{x})} \frac{dr}{v_s(r)}$$

The first term is the time needed to go from $\mathbf{x}$ to the closest point on the network, where the traveller enters the subway. The second term gives the time spend within the public transportation network to drive from the access point at the radial coordinate $d_s(\mathbf{x})$ to the centre along the branch. Using $v_w = 1$ and $v_s(r) = 1/(a + bF(r))$, this becomes

$$\tau(\mathbf{x}, \mathbf{0}) = d_w(\mathbf{x}) + a\, d_s(\mathbf{x}) + b \int_0^{d_s(\mathbf{x})} dr\, F(r) \tag{V-5}$$

The flow $F(r)$ on a branch at radial position $r$ is here proportional to the population living in the considered sector further outside. In Fig. V.3, the green hatched area marks the region from where all travellers contribute to the flow at radius $d_s$. Hence we obtain

$$F(r) = \int_{-\theta_n/2}^{\theta_n/2} d\theta' \int_r^\infty dr' r' \rho(r') = \theta_n \int_r^\infty dr'\, r' \rho(r') \tag{V-6}$$

To proceed further, we have to separate the population into two parts starting at a radius $r$ smaller or larger than the branch length $l_b$. A traveller starting at a point with $r \le l_b$ will go spherically to the next branch of the transportation network such

that $d_w = |\theta| r$ and $d_s = r$. Integrating over all starting points with $r \leq l_b$ yields the contribution

$$\begin{aligned}
\tau_1 = n\tau_{b,1} &= 2n \int_0^{l_b} dr \int_0^{\pi/n} d\theta\, r\, \tau(x(r,\theta),0)\rho(r) \\
&= 2n \int_0^{l_b} dr \int_0^{\pi/n} d\theta\, r \Bigg[ \underbrace{r\theta}_{=\tau_w} + \underbrace{ar + \frac{2\pi b}{n} \int_0^r dr' \int_{r'}^{\infty} dr''\, r''\, \rho(r'')}_{=\tau_s} \Bigg] \rho(r) \\
&= 2\pi \int_0^{l_b} dr\, r\, \rho(r) \Bigg[ \left( \frac{\pi}{2n} + a \right) r + \frac{2\pi b}{n} \int_0^r dr' \int_{r'}^{\infty} dr''\, r''\, \rho(r'') \Bigg]
\end{aligned} \tag{V-7}$$

to the total travelling time. A traveller starting further outwards at a radius $r > l_b$ will first go inwards radially until he is on the same radial position like the outer end of the transportation network, and then proceed spherically to the branch. Hence, $d_w = (r - l_b) + |\theta| l_b$ and $d_s = l_b$ and we obtain the second contribution to the total travelling time,

$$\begin{aligned}
\tau_2 = n\tau_{b,2} &= 2n \int_{l_b}^{\infty} dr \int_0^{\pi/n} d\theta\, r\, \tau(x(r,\theta),0)\rho(r) \\
&= 2n \int_{l_b}^{\infty} dr \int_0^{\pi/n} d\theta\, r \Bigg[ \underbrace{r + (\theta - 1) l_b}_{=\tau_w} + \underbrace{a l_b + \frac{2\pi b}{n} \int_0^{l_b} dr' \int_{r'}^{\infty} dr''\, r''\, \rho(r'')}_{=\tau_s} \Bigg] \rho(r) \\
&= 2\pi \int_{l_b}^{\infty} dr\, r\, \rho(r) \Bigg[ r + \left( \frac{\pi}{2n} - 1 + a \right) l_b + \frac{2\pi b}{n} \int_0^{l_b} dr' \int_{r'}^{\infty} dr''\, r''\, \rho(r'') \Bigg].
\end{aligned} \tag{V-8}$$

The total travelling time is then obtained by summing both contributions

$$\tau = n(\tau_1 + \tau_2)\,.$$

Solving the integrals for all three population densities defined in Sec. V.1.3 finally yields

$$\hat{\tau}_{hom}(\hat{L}, n) = \begin{cases} 1 - \frac{3}{2}(1-a)\frac{\hat{L}}{n} + \frac{3}{2}\left(\frac{\pi}{2} + b\right)\frac{\hat{L}}{n^2} + \frac{1-a}{2}\frac{\hat{L}^3}{n^3} - \left(\frac{\pi}{4} + b\right)\frac{\hat{L}^3}{n^4} + \frac{3b}{10}\frac{\hat{L}^5}{n^6} & , \quad \hat{L} \leq n \\ a + \left(\frac{\pi}{2} + \frac{4b}{5}\right)\frac{1}{n} & , \quad \hat{L} > n \end{cases}$$

$$\hat{\tau}_{Gauss}(\hat{L}, n) = 1 + \left( \frac{\pi}{2n} + a - 1 \right) \mathrm{erf}\left( \frac{\hat{L}}{n} \right) + \frac{b}{\sqrt{2}n} \mathrm{erf}\left( \frac{\sqrt{2}\hat{L}}{n} \right)$$

$$\hat{\tau}_{exp}(\hat{L}, n) = \left( \frac{\pi}{2n} + a + \frac{5}{8}\frac{b}{n} \right) - \left( \frac{\pi}{2n} + a - 1 \right) \left( 1 + \frac{\hat{L}}{2n} \right) e^{-\hat{L}/n} - \frac{b}{n} \left( \frac{\hat{L}^2}{4n^2} + \frac{3}{4}\frac{\hat{L}}{n} - \frac{5}{8} \right) e^{-2\hat{L}/n}$$

where $\hat{L} = L/r_0$ is the network length in units of the typical city size $r_0$. The optimal value of the parameter $n^\star$ which minimises $\tau$ and the corresponding $l_b^\star = L/n^\star$ is then computed numerically and results are plotted in Fig. V.4a-f. Note that we plot $n^\star$ which minimise the expressions above as real numbers, whereas in reality $n$ must obviously be a natural number.

As expected both the optimal number $n^\star$ and length $l_b^\star$ of the network branches increase with the amount of available resources $L$, but not in the same way. The optimal length $l_b^\star$ increases rapidly at first and then slowly saturates. The optimal number of branches is $n^\star = 4$ in the limit of low resources $L \to 0$ and no congestion ($b = 0$). It increases slowly at first, and then becomes approximately proportional to $L$. In view of applications this implies that limited resources one should first invest in the elongation of existing lines until they reach the scale of a city $l_b \approx r_0$ and only then invest into the buildup of new lines.

For all three densities one finds that congestion leads to an increase of $n^\star$ and a decrease of $l_b^\star$ as one would expect since a growing congestion adds a penalty for strong flows on single branches. Therefore it is reasonable that with congestion it is better to build more shorter branches instead of fewer longer branches. Generally, the impact of congestion on the optimal geometry is stronger for less compact cities, in particular for the exponential population density (see fig. V.4c,f) compared to more compact cities like in the Gaussian case (see fig. V.4b,e). In cities with finite sizes like the homogeneous disk, the impact of congestion diminishes for very large networks (see fig. V.4a,d).

Furthermore, congestion affects $n^\star$ most strongly of small value of $L$. For all $L > 5r_0$ the impact of congestion saturates, i.e. the difference in $n^\star$ for different values of the congestion parameter $b$ becomes independent of $L$ such that

$$n^\star(b, L) - n^\star(0, L) \approx \alpha b$$

with a proportionality factor of $\alpha \approx 0.66$ (homogeneous), $\alpha \approx 1.0$ (Gaussian) and $\alpha \approx 1.3$ (exponential), respectively.

A surprising effect is found for an exponential population distribution, where $n^\star$ diverges for very small networks and $b > 0$ (fig. V.4c). In this region, the minimum of $\tau(n)$ vanishes leaving a monotonically decreasing function $\tau(n)$. This result is caused by the underlying assumption that travellers always use the public transportation network. However, this is no longer valid if the network is very small and congested, then situations arise where it is better to walk directly to the city centre. This divergence is cured in the numerical treatment (cf. sec. V.1.5), where the travellers can also access the network right at the centre.

Figure V.4g-l further shows the corresponding optimal travelling times $\tau^\star$ as well as the marginal times, i.e. the negatives of the derivatives $g(L) := -d\tau^\star/dL$. The

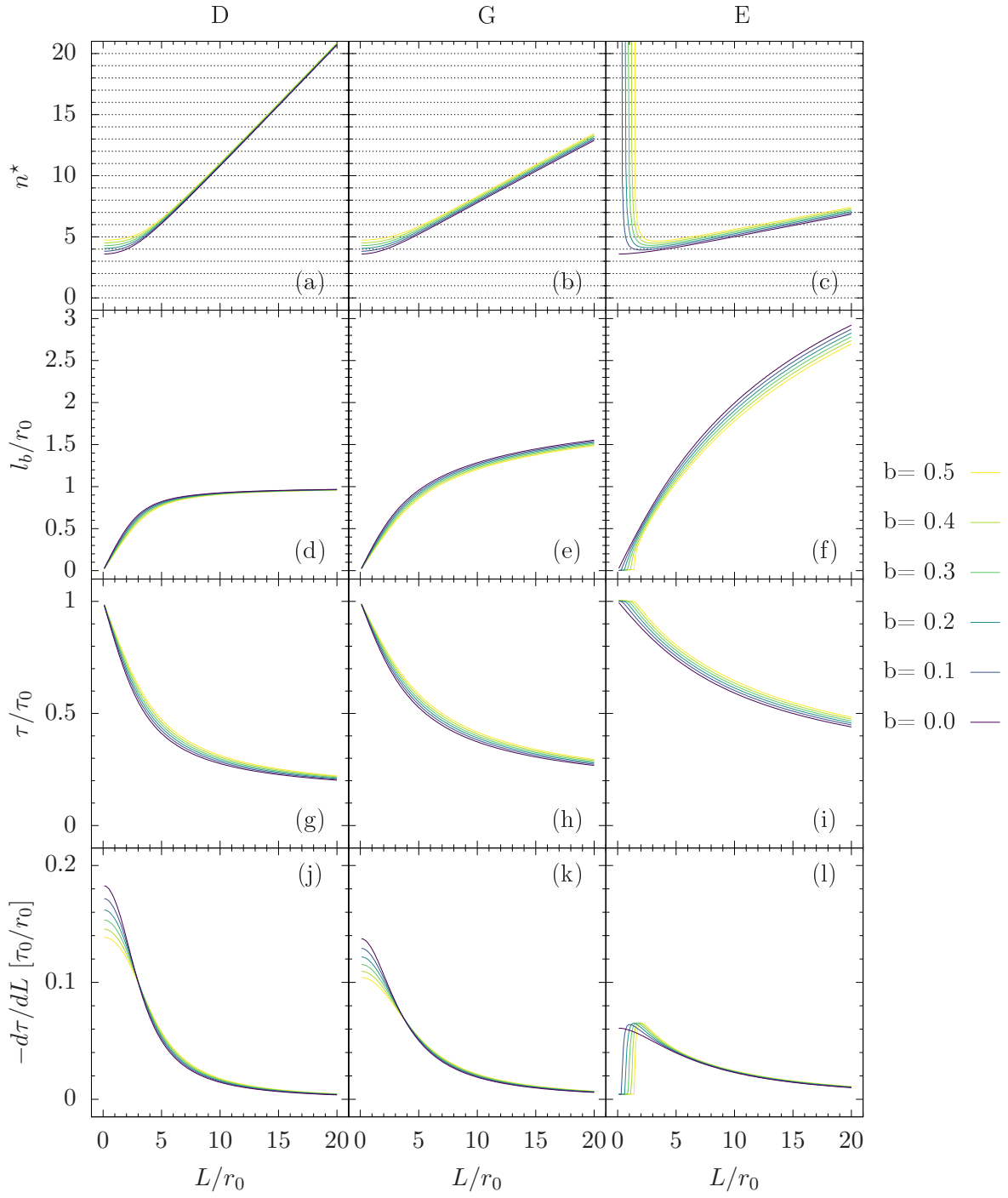



Figure V.4.: Analytical results for the star network: optimal number of branches $n$ (a-c), the corresponding branch lengths $l_b$ (d-f) and optimal tral times $\tau$ (g-i) as well as its derivative $d\tau/dL$ (j-l) for different radial symmetric densities: the homogeneous disk (left column), Gaussian density (middle column) and exponential density (right column).

effectiveness of the public transportation network is highest for the homogeneous population density representing a compact city, for which travelling time is reduced by a factor $\hat{\tau} \approx 0.2$ for $L \approx 20\,r_0$. Assuming that the construction costs of the transportation network are proportional to the length $L$, the marginal $g(L) = -d\tau^\star/dL$ gives the benefit-cost ratio for enlarging the network. For small networks, one find a high benefit-cost ratio in all cases up to the model breakdown in the exponential case. The more compact the city is, the higher is $g(L)$ which decays however rapidly and for $L > 5\,r_0$, the gain per unit length is almost equal in all cases. Notably, the benefit-cost ratio is lowered by congestion for small networks (small $L$), while it is increased for larger networks.

Considering economical constraints, this offers a possibility to determine an optimal network length by defining a minimal $g(L)$ value for which it is beneficial to extend the network. In very compact cities, one can easily find an value for $L$ were the benefit-cost ratio goes below a critical value. For less compact cities, this cutoff smears out. Adding congestion to the system, the slope of $g(L)$ decreases even more.

## V.3. Optimal star networks with a loop

In this section, we investigate a modification of the simple star network which is commonly observed in real public transportation networks: the addition of a loop track around the city centre as sketched in Fig. V.2a. In this case, the shape has an additional degree of freedom: the radius $r_l$ of the loop. The length of the branches $l_b$ is given by

$$L = n l_b + 2\pi r_l \quad \Leftrightarrow \quad l_b = \frac{L - 2\pi r_l}{n}\,.$$

To determine the optimal values of the geometrical parameters $n^\star$ and $r_l^\star$, we have scanned the parameter space spanned by $n$ and $r_l$ and choose the values which minimise $\tau$. As before, we fix $a = 1/8$ and analyse the results in dependence of the available resources $L$ and the congestion parameter $b$ for the different population density models from fig. V.1. We use the numerical approach introduced in Sec. V.1.5 to determine the total travel time and optimise the network structure.

The first important finding is that a pure star network is always superior to a loopy network for given resources $L$ in the "lazy traveller" model. In the numerical optimisation one thus always finds the optimal parameter $r_l^\star = 0$. This result is a direct consequence from the specific optimisation problem considered in this paper: All travellers want to to the city centre, such that loopy lines oriented orthogonal to this direction are not present in an optimal network. Many real public transportation networks do however feature loops to facilitate travel between different positions than the city centre.

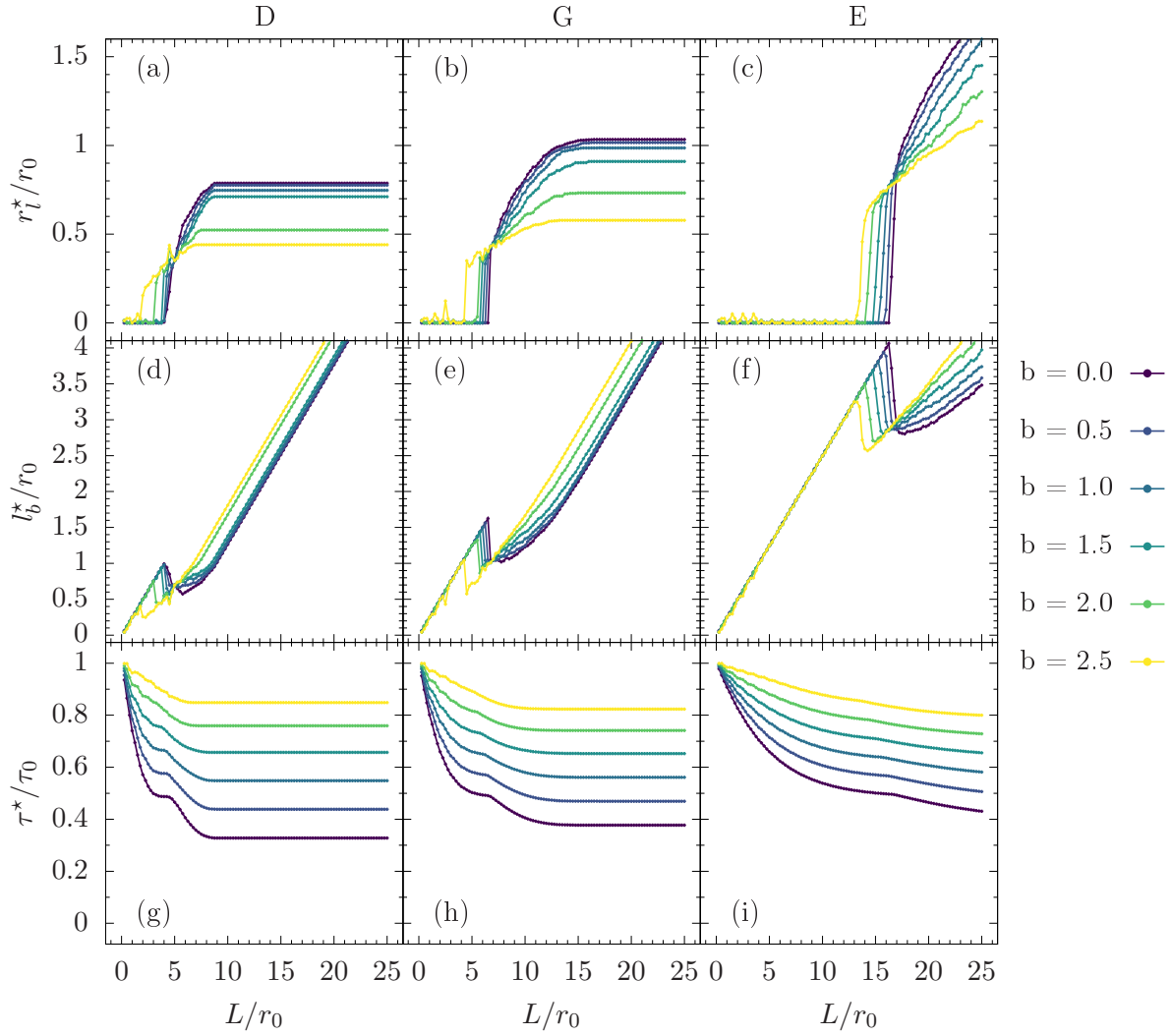



Figure V.5.: Optimal loop radius $r_l^\star$ for fixed number of branches $n = 4$ in the "lazy traveller" model

When using the "fast traveller" model we find, however, evidence that there are also parameters $L$ and $b$ such that a loop becomes beneficial, i.e. $r_l^\star > 0$ also in the distance-to-centre analysis. But unfortunately, the results are very noisy. We find a strong hint that some systematical error must be present in the data when plotting the optimal travel times $\tau^\star$ after optimising over both $L$ and $b$ as plotted in Fig. D.1j-l in the appendix. We find that the values of $\tau^\star$ are swapping back and forth between two different curves which are more and more separated for increasing $b$. While the variances over all scans for a fixed $L$ are negligible as illustrated by the almost point-like errorbars, we find strong variations between different values of $L$ which seem to be caused by an yet unknown systematical error in the simulation. Thus, further efforts are needed to find the cause of these unphysical findings.

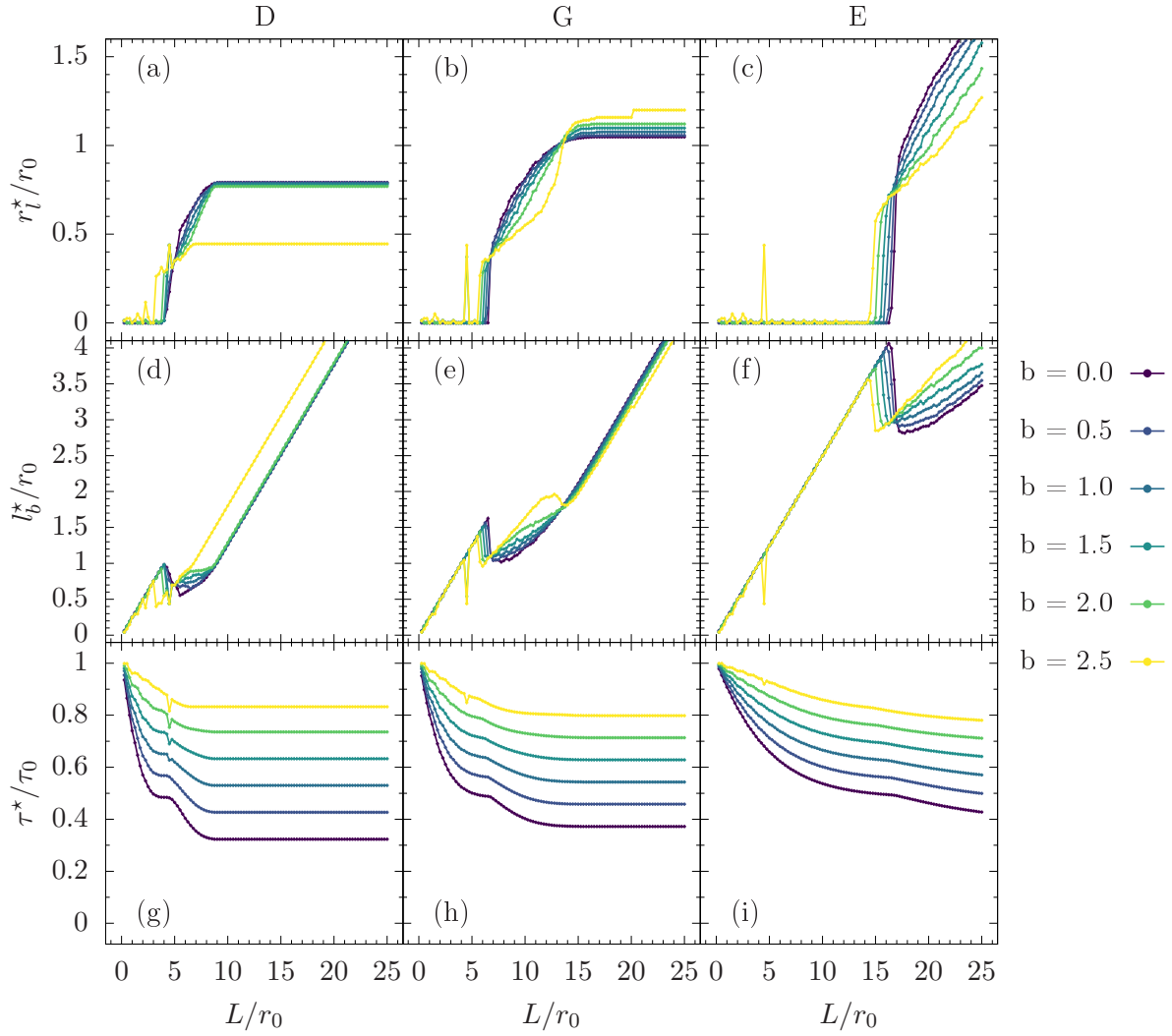



Figure V.6.: Optimal loop radius $r_l^\star$ for fixed number of branches $n = 4$ in the "fast traveller" model

### V.3.1. Optimal shape for fixed number of branches $n$

Loops can be beneficial for the travelling time to the centre also for the "lazy traveller" if the number of branches is constrained. Such a situation can indeed occur in practical applications where a star shaped network with a given number of branches $n$ is already present and space for further branches is not available. Then the question arises if existing branches should be extended further or if a loop should be established.

Figure V.5 show the results for this constrained optimisation problem in the "lazy traveller" model, fixing the number of branches at $n = 4$. Figure V.6 show the same data for the "fast traveller" model.

We find for all population density models studied in this chapter and for both the "lazy traveller" and the "fast traveller" qualitatively the same behaviour: For small values of $L$, the optimal structure is still a pure star with $r_l^\star = 0$. As $L$ increase beyond a critical value $L_{\text{crit}}$, the loopy network becomes superior and the optimal value $r_l^\star$ becomes non-zero. Remarkably, the transition is discontinuous in the sense that the optimal parameter $r_l^\star$ jumps at $L_{\text{crit}}$ – i.e. the loop comes into being with a non-zero radius. Correspondingly, the optimal branch length $l_b^\star$ jumps to a lower value as parts of the resources are now needed for the loop. After the phase transition, the optimal radius $r_l^\star$ increases with $L$ to some extent and then saturates. While the qualitative behaviour is the same, the actual values of $L_{\text{crit}}$ and $r_l^\star$ differ strongly for different density models: For more stretched population densities, it becomes more important to extent the branches to better reach outer parts of the city. As a consequence, the transition and saturation occur for much larger values of $L$ but when the loop occurs, its radius $r_l^\star$ is also larger.

The emergence of a discontinuity can be understood from the behaviour of the total travel time $\tau(L)$, which becomes mostly flat before the phase transition. Hence, the gain from extending the branches diminishes. At the phase transition, the gain from adding a loop with a finite radius exceeds the losses from decreasing $l_b$ by finite value. We note that such a discontinuity transition was already observed by Aldous & Barthélemy [12] and rigorously established for a different type of optimal networks in [40].

Comparing different numbers of branches $n$, we find no difference in the qualitative behaviour. The main quantitative difference is an increase of the critical value $L_{\text{crit}}$ with $n$.

Finally, congestion has only little effects on the optimal shape. In particular, the critical network length $L_{\text{crit}}$, where the phase transition takes place, is only weakly affected by the value of the congestion parameter $b$. On the other hand, we find a significant decrease of the optimal radius $r_l^\star$ with increasing congestion as it becomes less beneficial to build a long loop when the network velocity decreases due to congestion. We find a highly non-linear relation between the saturation value of $r_l^\star$ and $b$ when considering the homogeneous disk and Gaussian densities: While we observe only little impact on $r_l^\star$ for small $b$, the loop seem to collapse for very strong congestion in particular for the very compact city in the homogeneous disk model. Furthermore we find that congestion smooths out the phase transition, i.e. it reduces the height of the jumps in $r_l^\star$ and $l_b^\star$.

## V.4. Optimal star networks with branching lines

As a second extension to the simple star network we consider branching of the lines at a distance $l_1$ from the centre (cf. Fig. V.2b). The network is now characterised by three parameters: $n$, $l_1$ and the angle $\alpha$ between the outer branches, whose optimal values are determined using the numerical method introduced in sec. V.1.5 and a scan of the parameter space. Given a fixed number of resources $L$ we now have the decision whether to enlarge or split a branch to better cover the outskirts of a city.

For each value of the network length $L$ and the congestion parameter $b$, we performed 20 parameter scans with each having $N = 10^4$ starting points again for all three densities illustrated in Fig. V.1. We used the station distance $\Delta l = 0.05$. For each scan, we looked for the optimal parameter set $(n^\star, l_1^\star, \alpha^\star)$ which has minimal $\tau$. As we need to scan over a 3-dimensional parameter space, the procedure is computationally very expensive in particular for large networks, as there the number of possible parameter combinations grow with $L$. Thus, we focus in the following on networks with $L \leq 10$.

We found no significant differences between the "lazy traveller" and the "fast traveller" model. We could only observe a small upshift of the optimal travel times $\tau^\star$ for the lazy traveller compared to the fast traveller as expected. The following discussing thus uses only the results in the lazy traveller model which are plotted in Fig. V.7.

For $n^\star$, $l_1^\star$ and $\alpha^\star$ (Fig. V.7a-i) as well as $\tau^\star$ (Fig. V.7m-o), we plotted the optimal values of each parameter averaged over all scans. The errorbars illustrate their variances. The optimal outer branch length $l_2^\star$ (V.7j-l) is calculated using

$$l_2^\star(L, n^\star, l_1^\star) = \frac{L - n^\star l_1^\star}{2n^\star} = \frac{L}{2n^\star} - \frac{l_1^\star}{2}$$

and their errors are estimated from the variances $\Delta n^\star$ and $\Delta l_1^\star$ using a Taylor expansion

$$\Delta l_2^\star = \frac{L}{2(n^\star)^2}\Delta n^\star + \frac{1}{2}\Delta l_1^\star \,.$$

We notice that the values of the optimal parameters $n^\star$, $l_1^\star$, $\alpha^\star$ and $l_2^\star$ are very noisy in particular for the Gaussian density. The reason for this is that the $\tau$ values lay close together so that we observe different optimal points in parameter space which are alternating throughout the scan repetitions due to numerical fluctuations. Often we find two alternating points, one pure star with $l_2 = 0$ and one point with a finite length of the outer branches $l_2 > 0$. Further analysis with higher numerical precision should be carried out to determine which of the candidates corresponds to the real optimum.

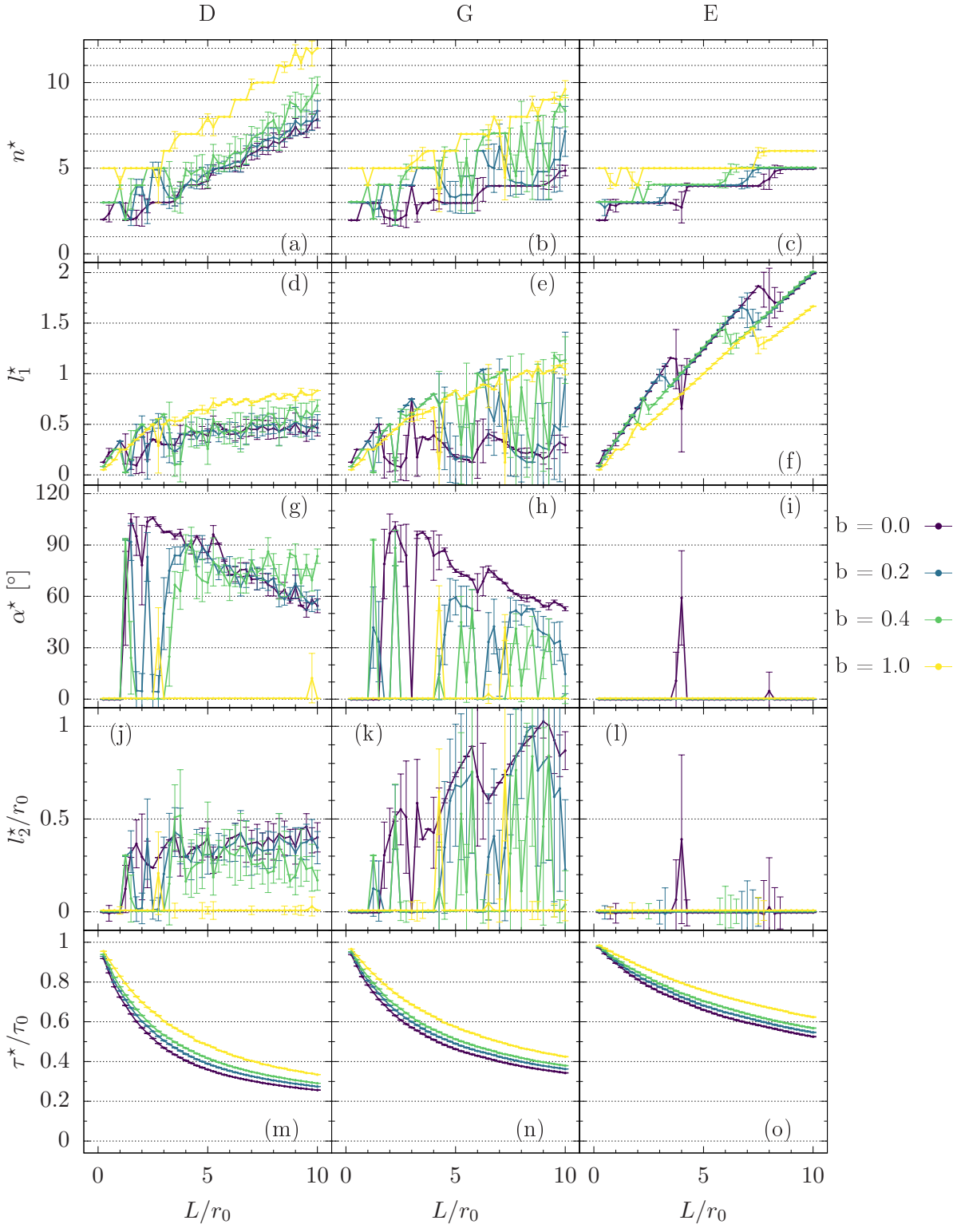



Figure V.7.: Optimal parameters for a star network with branching in the lazy traveller model.

Nevertheless, we can spot two different phases: In one region of the space spanned by $L$ and $b$, we observe a pure star network, i.e. $l_2^\star = 0$, which is also visualised with $\alpha^\star = 0°$ in Fig. V.7. On the other hand, we find regions where we observe branching, i.e. where $l_2^\star > 0$.

Comparing the different densities and values of $b$ we can conclude that the less compact the city, the less likely it is to find a branching of tracks, i.e. $l_2^\star > 0$. For the exponential density distribution, we find no branching at all in the considered region of parameters space with $L \leq 10$ (cf. Fig. V.7i,l). For the more compact densities we find that branching only occurs for weak congestion while for very strong congestion ($b = 1$) we observe only pure stars for all densities, i.e. $l_2^\star = 0$.

For the homogeneous disk and weak congestion ($b < 1$), we find a discontinuous transition from the pure star phase to the phase with branching, that means we find a jump from $l_2^\star = 0$ to a finite value $l_2^\star \in [0.3, 0.4]$ at a critical $L_{\text{crit}} \approx 1.5 r_0$ which seems to be independent of $b$, as long as the phase transition takes place at all. Beyond the critical network length $L_{\text{crit}}$, we observe a constant value of $l_2^\star$ while $n^\star$ and $l_1^\star$ are growing with $L$ and the angle $\alpha^\star$ decreases as expected. The data in the Gaussian density are too noisy to quantify the results on this data set. However, we find a similar pattern to the homogeneous disk with the major difference being a significant growing of $l_2^\star$ with $L$.

For both the homogeneous disk and the Gaussian density, we observe larger errorbars for larger $b$. This can be seen as a hint that the travel times $\tau$ of the optimal pure star and the optimal shape with branching are getting closer together when increasing the congestion parameter $b$ until the pure star finally becomes the global minimum for very strong congestion.

Hence, we can conclude that branching appears for networks larger than $L_{\text{crit}} \approx 1.5 r_0$ for compact cities while we do not observe branching in very spacious cities. The presence of congestion gives a further benefit of a pure star so that we find branching only in the large $L$ and small $b$ corner of the $L$-$b$ plane.

However, the data are very noisy so that quantitative conclusions are not possible at the moment. More precise simulations are needed to obtain more reliable results.

## V.5. Summary & Outlook

In this chapter, we discussed the optimal shape of public transport networks using different radial symmetry population densities by examining three fundamental geometries.

For the pure star network, we found a closed from solution. We found that in the presence of congestion it is more beneficial to build a larger number of shorter branches instead of fewer but longer lines. However, of reasonable values of the congestion parameter this shift in the optimal number of branches is only of order $\Delta n^\star \approx 1$.

Furthermore, we developed a numerical method that allows to examine more complex geometries by numerical optimisation. We focused on the appearance of a loop and the branching of lines in the outer part of the city separately.

When optimising over all free parameters of the respective geometry, we didn't observed the occurrence of a loop. If the number of branches is fixed, we found a discontinuous appearance of the loop when enlarging the network above a critical length. In presence of congestion, this phase transition is smoothed and the saturation radius of the loop decreases non linearly. For very strong congestion we observed jumps in the optimal travelling time $\tau^\star$ between neighbouring values of $L$ leading to fundamentally different optimal geometries. As these strong variations of $\tau^\star$ are hardly imaginable to be caused by physical effects, it looks like there is some systematic error in the implementation of the numerical solver for this geometry.

Considering branching of lines, we found for long networks and sufficiently weak congestion also a discontinuous appearance of branching. We observed that the branching is more likely to appear in more compact cities. However, the numerical results contain strong fluctuations. To further quantify the observations, the numerical precision needs to be improved.

In this work, we scanned over the entire parameter space with equidistant points to find the global minimum of the travel time. This procedure is numerical very expensive in particular for geometries with a large amount of free parameter, like the star network with branching. Finding a more efficient way to scan the parameter space for the optimum could significantly reduce the computation time and would allow to increase numerical precision by using a larger amount of starting points $N$ as well as to perform more runs on the same parameter point to get a larger statistics.

Once having more precise data, we can draw a phase diagram indicating in which part of the $L$-$b$-plane which of the considered geometries are optimal. Furthermore, it would allow us to quantify the impacts of congestion on the optimal shape such as the critical length where phase transitions occur.

# VI. Summary & Outlook

In this thesis, we introduced a fundamental model of linearly congested flow in transportation networks. In chapter II, we formulated the model as a quadratic programme (QP) which passes into regular shortest path flows for vanishing congestion and into current flows determined by Ohm's and Kirchhoff's law in the limit of very strong congestion. We generalised the formulation of Lagrange multipliers for finite congestion but didn't found an explicit solution to determine the flows directly, leading to a non-analytical dependence of the flows on the multipliers, impeding the use in numerical simulations. Hence, we implemented a numerical QP solver and discussed possibilities to reduce the numerical complexity of the problem using an adapted Hardy-Cross method.

We then investigated fundamental flow patterns in the presence of congestion in chapter III. In particular, we analysed the branching of flow and the rerouting pattern and found flow pattern for finite congestion to be similar to the pattern in the Ohmic limit up to a cut-off distance which grows with the square root of the congestion parameter $\eta$. The next step is to search for analytical evidence to support our numerical finding. Furthermore, we discussed the branching and rerouting patterns in this thesis only for regular lattices. In the next step, this analysis should be extended to real world networks to analyse the impact of congestion in irregular networks.

In chapter IV, we discussed the impact of congestion on loop formation in the optimisation of a general network where two branches are supported by a single fluctuating source. We generalised the discussion from [40] for finite congestion and found that the range of the budget parameter and the fluctuation strength for which it is optimal to include a loop shrinks significantly when approaching ordinary shortest path flows.

Finally, we discussed the optimal shape for multimodal transportation networks and developed a versatile method to find the optimal shape for any parameterised geometry and population density function for the distance-to-centre limit although the implementation requires further systematic testing to rule out systematic errors and more simulation runs to reduce numerical fluctuations. We applied this method to study three geometries which are frequently observed in real world subway systems and examined which geometry is optimal under which circumstances. We found the branching in a star network to occur for long networks and weak congestion but never observed the loop to be optimal. However, when fixing the number of branches we found a discontinuous appearance of the loop when enlarging the network above a

critical length. In presence of congestion, this phase transition is smoothed and the saturation radius of the loop decreases non linearly.

To conclude, we applied our linear congestion model in three different case-studies. We found empirical evidence that the pattern of congested flow is locally similar to the flow pattern in networks governed by Kirchhoff's and Ohm's law up to a certain distance which grows with the strength of congestion. Using this local similarity, we can profit from the enormous amount of research done for Ohmic flow networks, such as electrical power grids, to study congested flow networks, Analytical investigations should reveal more evidence for this empirical finding.

However, when considering multiple flow layers, the flows in transportation networks behave fundamentally different to current flows. To generalise our model to more complex flow scenarios, a multi-flow model should be implemented.

# References


[1] Lawren Sack and Christine Scoffoni. "Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future". In: *New Phytologist* 198.4 (2013), pp. 983–1000.

[2] Roland N Pittman. "Regulation of Tissue Oxygenation". In: *Integrated Systems Physiology: From Molecule to Function to Disease* (2011).

[3] Debashish Chowdhury, Andreas Schadschneider, and Katsuhiro Nishinari. "Physics of transport and traffic phenomena in biology: from molecular motors and cells to organisms". In: *Physics of Life reviews* 2.4 (2005), pp. 318–352.

[4] Dirk Brockmann and Dirk Helbing. "The Hidden Geometry of Complex, Network-Driven Contagion Phenomena". In: *Science* 342.6164 (2013), pp. 1337–1342.

[5] Joel Cohen, Frederic Briand, and Charles Newman. *Community Food Webs: Data and Theory*. Springer-Verlag Berlin Heidelberg, 1990.

[6] Charles D Brummitt et al. "Transdisciplinary electric power grid science". In: *Proceedings of the National Academy of Sciences* 110.30 (2013), pp. 12159–12159.

[7] Robert J Houghtalen, A Osman, and Ned HC Hwang. *Fundamentals of hydraulic engineering systems*. Prentice Hall New York, 2016.

[8] Rui Carvalho et al. "Resilience of natural gas networks during conflicts, crises and disruptions". In: *PloS one* 9.3 (2014), e90265.

[9] Kai Nagel and Michael Schreckenberg. "A cellular automaton model for freeway traffic". In: *Journal de Physique I* 2.12 (1992), pp. 2221–2229.

[10] Debashish Chowdhury, Ludger Santen, and Andreas Schadschneider. "Statistical physics of vehicular traffic and some related systems". In: *Physics Reports* 329.4-6 (2000), pp. 199–329.

[11] Camille Roth et al. "Structure of urban movements: polycentric activity and entangled hierarchical flows". In: *PloS one* 6.1 (2011), e15923.

[12] David Aldous and Marc Barthelemy. "Optimal geometry of transportation networks". In: *Phys. Rev. E* 99 (5 2019), p. 052303.

[13] C. von Ferber et al. "Public transport networks: empirical analysis and modeling". In: *The European Physical Journal B* 68.2 (2009), pp. 261–275.

[14] Albert-László Barabási and Réka Albert. "Emergence of Scaling in Random Networks". In: *Science* 286.5439 (1999), pp. 509–512.

[15] Pablo Kaluza et al. “The complex network of global cargo ship movements”. In: *Journal of the Royal Society Interface* 7.48 (2010), pp. 1093–1103.

[16] Tanya Latty et al. “Structure and formation of ant transportation networks”. In: *Journal of The Royal Society Interface* 8.62 (2011), pp. 1298–1306.

[17] Frank Schweitzer, Giorgio Fagiolo, and Douglas White. “Economic Networks: What do we know and what do we need to know?” In: *Advances in Complex Systems* 12 (2009), pp. 407–422.

[18] Jean-Philippe Bouchaud and Marc Mézard. “Wealth condensation in a simple model of economy”. In: *Physica A: Statistical Mechanics and its Applications* 282.3 (2000), pp. 536–545.

[19] Frank Schweitzer et al. “Economic Networks: The New Challenges”. In: *Science* 325.5939 (2009), pp. 422–425.

[20] W. Klingsch et al. *Pedestrian and Evacuation Dynamics 2008*. 2010.

[21] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. USA: Cambridge University Press, 2004.

[22] Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. “Network flows”. In: (1988).

[23] Béla Bollobás. *Modern graph theory*. en. Graduate texts in mathematics 184. New York: Springer, 1998.

[24] Gustav Kirchhoff. “Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird”. In: *Annalen der Physik* 148.12 (1847), pp. 497–508.

[25] ALGLIB Project. *ALGLIB*. `https://www.alglib.net`. Version 3.11.0-3. 2017.

[26] United States. Bureau of Public Roads. *Traffic assignment manual for application with a large, high speed computer*. Vol. 2. US Department of Commerce, 1964.

[27] Anna Nagurney and Qiang Qiang. “Robustness of transportation networks subject to degradable links”. In: *EPL (Europhysics Letters)* 80.6 (2007), p. 68001.

[28] Yosef Sheffi. *Urban transportation networks*. Vol. 6. Prentice-Hall, Englewood Cliffs, NJ, 1985.

[29] Mark Newman. *Networks: An Introduction*. Oxford University Press, 2010, p. 784.

[30] Panos M. Pardalos and Stephen A. Vavasis. “Quadratic programming with one negative eigenvalue is NP-hard”. In: *Journal of Global Optimization* (1991).

[31] M.K. Kozlov, S.P. Tarasov, and L.G. Khachiyan. “The polynomial solvability of convex quadratic programming”. In: *USSR Computational Mathematics and Mathematical Physics* 20.5 (1980), pp. 223–228.

[32] Thomas H. Cormen et al. *Introduction to Algorithms*. MIT Press.

[33] Dimitri P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. en. Athena Scientific, 1982.

[34] Hardy Cross. "Analysis of flow in networks of conduits or conductors". In: *Engineering Experiment Station Bulletin* 286 (1936).

[35] KVB. *Bahnen- & Schienennetz 2020*. `https://www.kvb.koeln/fahrtinfo/liniennetzplaene.html#lightbox/0/` Accessed: 27/08/2020.

[36] Luca Di Gaspero. *QuadProg++*. `https://github.com/liuq/QuadProgpp`. 2020.

[37] D. Goldfarb and A. Idnani. "A Numerically Stable Dual Method for Solving Strictly Convex Quadratic Programs". In: *Math. Program.* 27.1 (1983), pp. 1–33.

[38] Franz Kaiser, Julius Strake, and Dirk Witthaut. "Collective effects of link failures in linear flow networks". In: *New Journal of Physics* 22.1 (2020), p. 013053.

[39] Julius Strake et al. "Non-local impact of link failures in linear flow networks". In: *New Journal of Physics* 21.5 (2019), p. 053009.

[40] Franz Kaiser, Henrik Ronellenfitsch, and Dirk Witthaut. "Discontinuous transition to loop formation in optimal supply networks". In: (2020). in preparation.

[41] Franz Kaiser, Vito Latora, and Dirk Witthaut. "Inhibiting failure spreading in complex networks". In: *arXiv:2009.02910* (2020).

[42] Chunyan Li, Yuanzhang Sun, and Xiangyi Chen. "Analysis of the blackout in Europe on November 4, 2006". In: *2007 International Power Engineering Conference (IPEC 2007)*. 2007, pp. 939–944.

[43] Allen J. Wood, Bruce F. Wollenberg, and Gerald B. Sheblé. *Power Generation, Operation and Control*. New York: John Wiley & Sons, 2014.

[44] J. Guo, M. L. Crow, and J. Sarangapani. "An Improved UPFC Control for Oscillation Damping". In: *IEEE Transactions on Power Systems* 24.1 (2009), pp. 288–296.

[45] Max A. Woodbury. *Inverting modified matrices*. Statistical Research Group, Memo. Rep. no. 42. Princeton University, Princeton, N. J., 1950, p. 4.

[46] Francis Corson. "Fluctuations and Redundancy in Optimal Transport Networks". In: *Phys. Rev. Lett.* 104 (4 2010), p. 048703.

[47] Steffen Bohn and Marcelo O. Magnasco. "Structure, Scaling, and Phase Transition in the Optimal Transport Network". In: *Phys. Rev. Lett.* 98 (8 2007), p. 088702.

[48] Eleni Katifori, Gergely J. Szöllősi, and Marcelo O. Magnasco. "Damage and Fluctuations Induce Loops in Optimal Transport Networks". In: *Phys. Rev. Lett.* 104 (4 2010), p. 048704.

[49] Ned Hwang and Robert Houghtalen. *Fundamentals of Hydraulic Engineering Systems*. Upper Saddle River, NJ: Prentice Hall, 1996.

[50] Tatyana Gavrilchenko and Eleni Katifori. “Resilience in hierarchical fluid flow networks”. In: *Physical Review E* 99.1 (2019), p. 012321.

[51] Petros Polichronidis and Michael Schreckenberg. “Movement Characteristics of Processions”. In: *Collective Dynamics* 5 (2020), pp. 348–355.

[52] Lars Hufnagel, Dirk Brockmann, and Theo Geisel. “Forecast and control of epidemics in a globalized world”. In: *Proceedings of the National Academy of Sciences* 101.42 (2004), pp. 15124–15129.

[53] Rui Carvalho et al. “Robustness of trans-European gas networks”. In: *Physical review E* 80.1 (2009), p. 016106.

# A. Useful Integrals

In the following, some integrals which are important for this thesis are evaluated.

## A.1. Differences and sums of Gaussian random variables

Let $X_1$ and $X_2$ iid Gaussian random variables with both mean $\mu$ and variance $\sigma$. Their PDF reads then

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

For the difference $\Delta X = X_1 - X_2$ one finds:

$$\begin{aligned}
p(\Delta x) &= \int_{\mathcal{R}} dx p(x) p(x+\Delta x) \\
&= \frac{1}{2\pi\sigma^2} \int_{\mathcal{R}} dx \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} - \frac{(x+\Delta x-\mu)^2}{2\sigma^2}\right) \quad \Big| \; x \to x-\mu \\
&= \frac{1}{2\pi\sigma^2} \int_{\mathcal{R}} dx \exp\left(-\frac{x^2 + (x^2 + 2x\,\Delta x + \Delta x^2)}{2\sigma^2}\right) \\
&= \frac{1}{2\pi\sigma^2} \int_{\mathcal{R}} dx \exp\left(-\frac{2(x^2 + x\Delta x + \frac{\Delta x^2}{4}) + \frac{\Delta x^2}{2}}{2\sigma^2}\right) \\
&= \frac{1}{2\pi\sigma^2} \int_{\mathcal{R}} dx \exp\left(-\frac{2(x+\frac{\Delta x}{2})^2 + \frac{\Delta x^2}{2}}{2\sigma^2}\right) \quad \Big| \; x \to x + \frac{\Delta x}{2} \\
&= \frac{\exp\left(-\frac{\Delta x^2}{4\sigma^2}\right)}{2\pi\sigma^2} \underbrace{\int_{\mathcal{R}} dx \exp\left(-\frac{2x^2}{2\sigma^2}\right)}_{=\sqrt{2\pi\frac{\sigma^2}{2}}}
\end{aligned}$$

$$= \frac{1}{\sqrt{4\pi\sigma^2}} \exp\left(-\frac{\Delta x^2}{4\sigma^2}\right) \quad | \quad \sigma' = \sqrt{2}\sigma$$
$$= \frac{1}{\sqrt{2\pi\sigma'^2}} \exp\left(-\frac{\Delta x^2}{2\sigma'^2}\right)$$

Hence $\Delta X$ is a Gaussian random variable with mean 0 and variance $\sqrt{2}\sigma$.

Analogue, one finds that $X_1 + X_2$ is a Gaussian random variable with mean $2\mu$ and variance $\sqrt{2}\sigma$.

## A.2. Means of absolute values of Gaussian random variables

Consider a Gaussian random variable $X$ with mean $\langle X \rangle = \mu$ and $\mathrm{Var}(X) = \sigma$, i.e. it's PDF is

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{A-1}$$

The mean of it's absolute value then ready

$$\begin{aligned}
\langle |X| \rangle &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathcal{R}} dx\, |x| \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad | \quad y := \frac{x-\mu}{\sqrt{2\sigma^2}} \\
&= \frac{1}{\sqrt{\pi}} \int_{\mathcal{R}} dy\, \left|\sqrt{2\sigma^2} y + \mu\right| \exp\left(-y^2\right) \\
&= \frac{1}{\sqrt{\pi}} \left( - \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma^2}}} + \int_{-\frac{\mu}{\sqrt{2\sigma^2}}}^{\infty} \right) dy\, \left(\sqrt{2\sigma^2} y + \mu\right) \exp\left(-y^2\right) \\
&= \frac{\sigma}{\sqrt{2\pi}} \left( \int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma^2}}} - \int_{-\frac{\mu}{\sqrt{2\sigma^2}}}^{\infty} \right) dy\, \underbrace{(-2) y \exp\left(-y^2\right)}_{= \partial_y \exp(-y^2)} \\
&\quad + \frac{\mu}{\sqrt{\pi}} \Bigg( - \underbrace{\int_{-\infty}^{-\frac{\mu}{\sqrt{2\sigma^2}}} dy \exp\left(-y^2\right)}_{= \sqrt{\pi}/2\left(1 + \mathrm{erf}\left(-\mu/\sqrt{2\sigma^2}\right)\right)} + \underbrace{\int_{-\frac{\mu}{\sqrt{2\sigma^2}}}^{\infty} dy \exp\left(-y^2\right)}_{= \sqrt{\pi}/2\left(1 - \mathrm{erf}\left(-\mu/\sqrt{2\sigma^2}\right)\right)} \Bigg) \\
&= \frac{\sigma}{\sqrt{2\pi}} \left( \exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \right) - \mu\, \mathrm{erf}\left(-\frac{\mu}{\sqrt{2\sigma^2}}\right)
\end{aligned}$$

$$= \frac{2\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \mu \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right)$$

Hence:

$$\langle |x| \rangle = \frac{2\sigma}{\sqrt{2\pi}} \exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \mu \operatorname{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right) \tag{A-2}$$

## A.3. Second momentum of Gaussian random variables

An independent and identically distributed Gaussian random variable $X \sim \mathcal{N}(\mu, \sigma)$ has the PDF

$$p(X = x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{A-3}$$

Hence their expectation value is $\langle X \rangle = \mu$ and the variance is $Var(X) = \sigma$. In the following, some interesting values are calculated.

$$\begin{aligned}
\langle X^2 \rangle &= \int_{\mathbb{R}} x^2 p(x)\, dx \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int x^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \qquad \mid\ y := x - \mu \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int (y^2 + 2\mu y + \mu^2) e^{-\frac{y^2}{2\sigma^2}}\, dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int y^2 e^{-\frac{y^2}{2\sigma^2}}\, dy + \frac{2\mu}{\sqrt{2\pi\sigma^2}} \underbrace{\int y e^{-\frac{y^2}{2\sigma^2}}\, dy}_{=0} + \underbrace{\frac{\mu^2}{\sqrt{2\pi\sigma^2}} \int e^{-\frac{y^2}{2\sigma^2}}\, dy}_{=\mu^2} \\
&= \mu^2 + \int \underbrace{y}_{=:u} \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} y\, e^{-\frac{y^2}{2\sigma^2}}}_{=:v'}\, dy \qquad \mid\ v = -\sigma^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} \\
&= \mu^2 + [u \underbrace{v}_{=0}]_{-\infty}^{\infty} - \int \underbrace{u'}_{=1} v \\
&= \mu^2 + \sigma^2 \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \int e^{-\frac{y^2}{2\sigma^2}}\, dy}_{=1} \\
&= \mu^2 + \sigma^2
\end{aligned}$$

# B. Calculations

In the following, some calculations are done which are to lengthy to include them in the main text.

## B.1. Loopyness - Topological phase transition

### B.1.1. Optimal capacities in the tree case

In this section, the analytical expressions for the optimal capacities $\kappa_l$ from sec. IV.3 are derived. The $\kappa_l$ are given by

$$\kappa_l = \frac{\left( \langle F_l^2 \rangle + \lambda \langle |F_l| \rangle \right)^{\frac{1}{1+\gamma}}}{\left( \sum_l \left( \langle F_l^2 \rangle + \lambda \langle |F_l| \rangle \right)^{\frac{\gamma}{1+\gamma}} \right)^{\frac{1}{\gamma}}} K$$

which can be simplified with the definition of

$$A_l := \left( \langle F_l^2 \rangle + \lambda \langle |F_l| \rangle \right)^{\frac{1}{1+\gamma}} \tag{B-1}$$

using the symmetry between the two branches, one finds $A_1 = A_3$ and $A_2 = A_4$. Thus

$$\kappa_l = \frac{A_l}{(2A_1^\gamma + 2A_2^\gamma)^{\frac{1}{\gamma}}} K = \left( \frac{2\left(A_1^\gamma + A_2^\gamma\right)}{A_l^\gamma} \right)^{-\frac{1}{\gamma}} K$$

so that we get:

$$\kappa_1 = \kappa_3 = \left[ 2\left( 1 + \left( \frac{A_2}{A_1} \right)^\gamma \right) \right]^{-\frac{1}{\gamma}} K \tag{B-2}$$

$$\kappa_2 = \kappa_4 = \left[ 2\left( 1 + \left( \frac{A_1}{A_2} \right)^\gamma \right) \right]^{-\frac{1}{\gamma}} K \tag{B-3}$$

with

$$\langle |F_1| \rangle = \langle |F_3| \rangle = \langle |P_2 + P_3| \rangle = \frac{2\sigma}{\sqrt{\pi}} \exp\left( -\frac{\mu^2}{\sigma^2} \right) + 2\mu \operatorname{erf}\left( \frac{\mu}{\sigma} \right)$$

| case | flows | cycle flow |
|---|---|---|
| a | $F_1, F_3 \geq 0$ | $f_a$ |
| b | $F_1 > 0 > F_3$ | $f_b$ |
| c | $F_3 > 0 > F_1$ | $-f_b$ |
| d | $F_1, F_3 \leq 0$ | $-f_a$ |

Table B.1.: Different cases for the cycle flow $f$

$$\begin{aligned}\langle|F_2|\rangle = \langle|F_4|\rangle = \langle|P_3|\rangle &= \frac{2\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \mu\mathrm{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right)\\ \langle F_1^2\rangle = \langle F_3^2\rangle &= 4\mu^2 + 2\sigma^2\\ \langle F_2^2\rangle = \langle F_4^2\rangle &= \mu^2 + \sigma^2\end{aligned}$$

one find:

$$\left(\frac{A_2}{A_1}\right)^\gamma = \left(\frac{\mu^2 + \sigma^2 + \lambda\frac{2\sigma}{\sqrt{2\pi}}\exp\left(-\frac{\mu^2}{2\sigma^2}\right) + \mu\mathrm{erf}\left(\frac{\mu}{\sqrt{2\sigma^2}}\right)}{4\mu^2 + 2\sigma^2 + \lambda\frac{2\sigma}{\sqrt{\pi}}\exp\left(-\frac{\mu^2}{\sigma^2}\right) + 2\mu\mathrm{erf}\left(\frac{\mu}{\sigma}\right)}\right)^{\frac{\gamma}{\gamma+1}} \tag{B-4}$$

Inserting (B-4) into (B-2) and the inverse of (B-4) into (B-3) respectively allows to compute the optimal capacities for given parameters $\gamma$, $\sigma$ and $\lambda$.

## B.1.2. Derivation of the cycle flow

The scenario demands that the flow is arranged such that the total travel time $\tau$ is minimised, i.e. the value of $f$ is defined by $\partial_f \tau_{loop} = 0$. As the links 2 and 4 are not affected by $f$, they can be omitted in the following and only the reduced travel time $\tau_{loop}(f, F_1, F_3)$ is considered, which reads using $\kappa_1 = \kappa_3$:

$$\begin{aligned}\tau_{loop}(f, F_1, F_3) &= \frac{1}{\kappa_1}\left(\tilde{F}_1^2 + \lambda|\tilde{F}_1| + \tilde{F}_3^2 + \lambda|\tilde{F}_3|\right) + \frac{1}{\kappa_5}\left(\tilde{F}_5^2 + \lambda|\tilde{F}_5|\right)\\ &= \frac{1}{\kappa_1}\left((F_1 - f)^2 + (F_3 + f)^2 + \lambda(|F_1 - f| + |F_3 + f|)\right) + \frac{1}{\kappa_5}\left(f^2 + \lambda|f|\right)\end{aligned} \tag{B-5}$$

To find the optimal $f$, we split the problem up into 4 cases, each of them covering one quadrant of the $F_1$-$F_3$ plane. Using the symmetry $\tau_{loop}(f, -F_1, -F_3) = \tau_{loop}(-f, F_1, F_3)$, only the cases $F_1, F_3 \geq 0$ and $F_1 \geq 0 > F_3$ need to be studied, the values of $f$ in the other quadrants are related to these cases as shown in tab. B.1.

### B.1.2.1. Case a: $F_1, F_3 \geq 0$

In this region, $\tau_{loop}(f_a)$ is continuous and has 4 smooth components in the intervals

$$\begin{aligned}I_1^a &= (-\infty, -F_3) & I_2^a &= (-F_3, 0)\\ I_3^a &= (0, F_1) & I_4^a &= (F_1, \infty)\end{aligned} \tag{B-6}$$

Using these intervals, $\tau_{loop}$ can be decomposed into its smooth components:

$$\tau_{loop}(f_a) = \frac{F_1^2 + F_3^2 + 2f_a^2 + 2f_a(F_3 - F_1)}{\kappa_1} + \frac{f_a^2}{\kappa_5} + \lambda \begin{cases} \kappa_1^{-1}(\ \ F_1 - F_3 - 2f_a) - \kappa_5^{-1} f_a & , f_a \in I_1^a \\ \kappa_1^{-1}(\ \ F_1 + F_3) - \kappa_5^{-1} f_a & , f_a \in I_2^a \\ \kappa_1^{-1}(\ \ F_1 + F_3) + \kappa_5^{-1} f_a & , f_a \in I_3^a \\ \kappa_1^{-1}(-F_1 + F_3 + 2f_a) + \kappa_5^{-1} f_a & , f_a \in I_4^a \end{cases}$$

To find the global minimum, we are looking for the minima of each component:

$$\partial_{f_a} \tau_{loop}(f_a) = \frac{4f_a + 2(F_3 - F_1)}{\kappa_1} + \frac{2f_a}{\kappa_5} + \lambda \begin{cases} -2\kappa_1^{-1} - \kappa_5^{-1} & , f_a \in I_1^a \\ -\kappa_5^{-1} & , f_a \in I_2^a \\ \kappa_5^{-1} & , f_a \in I_3^a \\ 2\kappa_1^{-1} + \kappa_5^{-1} & , f_a \in I_4^a \end{cases}$$

The derivatives vanish in the following points:

$$f_a = \frac{F_1 - F_3}{2 + \frac{\kappa_1}{\kappa_5}} - \frac{\lambda \kappa_1}{2\left(2 + \frac{\kappa_1}{\kappa_5}\right)} \begin{cases} -2\kappa_1^{-1} - \kappa_5^{-1} & , f_a \in I_1^a \\ -\kappa_5^{-1} & , f_a \in I_2^a \\ \kappa_5^{-1} & , f_a \in I_3^a \\ 2\kappa_1^{-1} + \kappa_5^{-1} & , f_a \in I_4^a \end{cases}$$

Introducing

$$\rho := 2 + \frac{\kappa_1}{\kappa_5} \tag{B-7a}$$

$$\Delta F := F_1 - F_3 \tag{B-7b}$$

$$\Delta F_c := \frac{\lambda \kappa_1}{2\kappa_5} \tag{B-7c}$$

we can write the minimum of each of the 4 smooth components as

$$\tilde{f}_{a1} = \frac{\Delta F + \Delta F_c + \lambda}{\rho} \qquad \tilde{f}_{a2} = \frac{\Delta F + \Delta F_c}{\rho}$$

$$\tilde{f}_{a3} = \frac{\Delta F - \Delta F_c}{\rho} \qquad \tilde{f}_{a4} = \frac{\Delta F - \Delta F_c - \lambda}{\rho}$$

Each component is a quadratic polynomial in $f_a$, it has therefore only one minimum. If the corresponding $\tilde{f}_{ai}$ value lays outside the interval where the component is defined, the minimum of $\tau_{loop}$ in the corresponding interval $I_i^a$ is located at the intervals border which is closest to $\tilde{f}_{ai}$. Using that the $\tilde{f}_{ai}$ decrease with increasing $f_a$ and that $\tau_{loop}$ is continuous, we find that the global minimum of $\tau_{loop}$ must either lay inside one of the smooth intervals at $\tilde{f}_{ai}$ or, if both $\tilde{f}_{ai}$ of neighbouring intervals lay inside the other interval, the minimum is located at the corresponding limit between those intervals.

Comparing this result with the interval borders in (B-6), we find that there are three intervals where the global minimum is stationary:

$$\begin{aligned}
\Delta F \in [-\rho F_3 - \Delta F_c - b, -\rho F_3 - \Delta F_c] \quad &\rightarrow \quad \tilde{f}_1 > -F_3 \ \wedge \ \tilde{f}_2 < -F_3 \quad \rightarrow \quad f = -F_3 \\
\Delta F \in [-\Delta F_c, \Delta F_c] \quad &\rightarrow \quad \tilde{f}_2 > 0 \ \wedge \ \tilde{f}_3 < 0 \quad \rightarrow \quad f = 0 \\
\Delta F \in [\rho F_1 + \Delta F_c, \rho F_1 + \Delta F_c + b] \quad &\rightarrow \quad \tilde{f}_3 > F_1 \ \wedge \ \tilde{f}_4 < F_1 \quad \rightarrow \quad f = F_1
\end{aligned}$$

In between of these intervals, the global minimum is identical to the corresponding local minimum $\tilde{f}_{ai}$. We can therefore summarise:

$$f_a(F_3, F_1) = \frac{1}{\rho} \begin{cases} \Delta F + \Delta F_c + b & , \ \Delta F \in (-\infty \ , \ -\rho F_3 - \Delta F_c - b] \\ -\rho F_3 & , \ \Delta F \in [-\rho F_3 - \Delta F_c - b \ , \ -\rho F_3 - \Delta F_c] \\ \Delta F + \Delta F_c & , \ \Delta F \in [-\rho F_3 - \Delta F_c \ , \ -\Delta F_c] \\ 0 & , \ \Delta F \in [-\Delta F_c \ , \ \Delta F_c] \\ \Delta F - \Delta F_c & , \ \Delta F \in [\Delta F_c \ , \ \rho F_1 + \Delta F_c] \\ \rho F_1 & , \ \Delta F \in [\rho F_1 + \Delta F_c \ , \ \rho F_1 + \Delta F_c + b] \\ \Delta F - \Delta F_c - b & , \ \Delta F \in [\rho F_1 + \Delta F_c + b \ , \ \infty) \end{cases} \tag{B-8}$$

Using the constrains $F_1, F_3 \geq 0$, we find that

$$\begin{aligned}
\Delta F = F_1 - F_3 \leq \quad & F_1 < \rho F_1 + \Delta F_c \\
\Delta F = F_1 - F_3 \geq -F_3 > \quad & -\rho F_3 - \Delta F_c
\end{aligned}$$

Therefore only the 3th to the 5th cases from (B-8) occur in the region where $f_a$ is defined. It can therefore be compactified to

$$f_a(\Delta F) = \frac{\Delta F - \operatorname{sgn}(\Delta F)\, \Delta F_c}{\rho} \Theta\left(|\Delta F| - \Delta F_c\right) \tag{B-9}$$

#### B.1.2.2. Case b: $F_3 \geq 0 > F_1$

$\tau_{loop}$ is also in this case continuous with 4 smooth components in the intervals

$$\begin{aligned}
I_1^b &= (-\infty \ , \ -F_{max}) & I_2^b &= (-F_{max} \ , \ -F_{min}) \\
I_3^b &= (-F_{min} \ , \ 0) & I_4^b &= (0 \ , \ \infty)
\end{aligned} \tag{B-10}$$

with $F_{min} = \min(-F_1, F_3)$ and $F_{max} = \max(-F_1, F_3)$. With this, we can again decompose (B-5) into its smooth components:

$$\tau_{loop}(f_b) = \frac{F_1^2 + F_3^2 + 2f_b^2 + 2f_b(F_3 - F_1)}{\kappa_1} + \frac{f_b^2}{\kappa_5} + \lambda \begin{cases} \kappa_1^{-1}\big(F_1 - F_3 - 2f_b\big) - \kappa_5^{-1} f_b & , f_b \in I_1 \\ \kappa_1^{-1}\big(\pm(F_1 + F_3)\big) - \kappa_5^{-1} f_b & , f_b \in I_2 \\ \kappa_1^{-1}\big(-F_1 + F_3 + 2f_b\big) - \kappa_5^{-1} f_b & , f_b \in I_3 \\ \kappa_1^{-1}\big(-F_1 + F_3 + 2f_b\big) + \kappa_5^{-1} f_b & , f_b \in I_4 \end{cases}$$

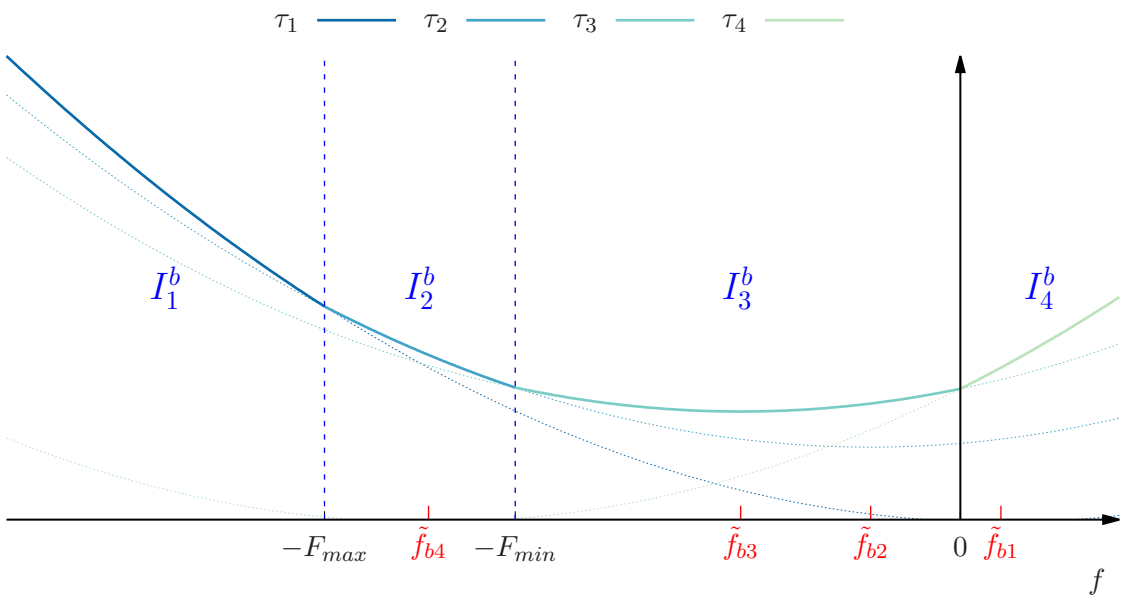



Figure B.1.: The smooth components $\tau_1$ to $\tau_4$ of $\tau_{loop}$ for $\kappa_1 = 0.12$, $\kappa_5 = 0.05$ and $\lambda = 0.9$. The solid line marks $\tau_{loop}$ while the dotted lines are the continuations of the smooth components outside the interval where they are defined. The minima of each smooth component is marked on the x-axis. In this example, the global minimum lays at $\tilde{f}_{b3}$ as it lay inside the correct interval, $I_3^b$.

where the $\pm$ is $+$ if $|F_3| > |F_1|$ and $-$ else. The derivatives read:

$$\partial_{f_b}\tau_{loop}(f_b) = \frac{4f_b + 2(F_3 - F_1)}{\kappa_1} + \frac{2f_b}{\kappa_5} + \lambda \begin{cases} -\kappa_5^{-1} - 2\kappa_1^{-1} & , f_b \in I_1 \\ -\kappa_5^{-1} & , f_b \in I_2 \\ -\kappa_5^{-1} + 2\kappa_1^{-1} & , f_b \in I_3 \\ \kappa_5^{-1} + 2\kappa_1^{-1} & , f_b \in I_4 \end{cases}$$

which vanish in the points

$$\tilde{f}_b = \frac{F_1 - F_3}{2 + \frac{\kappa_1}{\kappa_5}} - \frac{\lambda\kappa_1}{2\left(2 + \frac{\kappa_1}{\kappa_5}\right)} \begin{cases} -2\kappa_1^{-1} - \kappa_5^{-1} & , f_b \in I_1^b \\ -\kappa_5^{-1} & , f_b \in I_2^b \\ 2\kappa_1^{-1} - \kappa_5^{-1} & , f_b \in I_3^b \\ 2\kappa_1^{-1} + \kappa_5^{-1} & , f_b \in I_4^b \end{cases}$$

Using the definitions from (B-7), these minima can be rewritten:

$$\tilde{f}_{b1} = \frac{\Delta F + \Delta F_c + \lambda}{\rho} \qquad \tilde{f}_{b2} = \frac{\Delta F + \Delta F_c}{\rho}$$
$$\tilde{f}_{b3} = \frac{\Delta F + \Delta F_c - \lambda}{\rho} \qquad \tilde{f}_{b4} = \frac{\Delta F - \Delta F_c - \lambda}{\rho}$$

Since $\Delta F_c \geq 0$ and $\lambda \geq 0$, the value of the minima decrease with increasing $f_b$ values as plotted in fig. B.1. Using this visualisation we can find the global minimum with the following algorithm: If the minimum $\tilde{f}_{bi}$ of a smooth component lays left to the corresponding interval $I_i^b$, the global minimum of $\tau_{loop}$ must also lay left or on the left end of this interval. The same holds vice versa when $\tilde{f}_{bi}$ lays right of the $I_i^b$.

Using this method, we can restrict the area were the minimum can be. We have $\Delta F = F_1 - F_3 \leq 0$. Hence $\tilde{f}_{b4} \leq 0$, so the minimum can't lay in $I_4$. We find also

$$\tilde{f}_{b1} = \frac{\Delta F + \Delta F_c + \lambda}{\rho} \geq \frac{\Delta F}{\rho} > \frac{\Delta F}{2} \geq -F_{max}$$

using $\rho > 2$, $\Delta F < 0$ and $\Delta F \geq -2F_{max}$. $\tilde{f}_{b1}$ is therefore always located right of $I_1^b$ so that the global minimum must lay in $I_2^b$ or $I_3^b$. For the remaining analysis we need to separate two cases: When $\Delta F < -2F_3$, we have $F_{min} = F_3$. When $\Delta F \in [-2F_3, -F_3]$, we get $F_{min} = -(\Delta F + F_3)$.

#### B.1.2.2.1. case $\Delta F < -2F_3$

The condition for $\tilde{f}_{b2}$ to lay in $I_2^b$ is

$$\frac{\Delta F + \Delta F_c}{\rho} < -F_{min} = -F_3 \quad \Leftrightarrow \quad \Delta F < -\rho F_3 - \Delta F_c$$

The condition for $\tilde{f}_{b3}$ to lay in $I_3^b$ is

$$\frac{\Delta F + \Delta F_c - \lambda}{\rho} > -F_{min} = -F_3 \quad \Leftrightarrow \quad \Delta F > -\rho F_3 - \Delta F_c + \lambda$$

In the interval $[-\rho F_3 - \Delta F_c, -\rho F_3 - \Delta F_c + \lambda]$ where both $\tilde{f}_{b2}$ and $\tilde{f}_{b3}$ lay in the opposite interval, the global minimum is located on the border between $I_2^b$ and $I_3^b$, i.e. $f_b = -F_3$. The result is therefore:

$$f_b(F_3, \Delta F) = \frac{1}{\rho} \begin{cases} \Delta F + \Delta F_c & , \quad \Delta F \in (-\infty \ , \ -\rho F_3 - \Delta F_c] \\ -\rho F_3 & , \quad \Delta F \in [-\rho F_3 - \Delta F_c \ , \ -\rho F_3 - \Delta F_c + \lambda] \\ \Delta F + \Delta F_c - \lambda & , \quad \Delta F \in [-\rho F_3 - \Delta F_c + \lambda \ , \ -2F_3] \end{cases}$$

#### B.1.2.2.2. case $\Delta F \in [-2F_3, -F_3]$

Now, the condition for $\tilde{f}_{b2}$ to lay in $I_2^b$ is

$$\frac{\Delta F + \Delta F_c}{\rho} < -F_{min} = \Delta F + F_3 \quad \Leftrightarrow \quad \Delta F > \frac{\Delta F_c - \rho F_3}{\rho - 1}$$

and for $\tilde{f}_{b3}$ to be in $I_3^b$ we need

$$\frac{\Delta F + \Delta F_c - \lambda}{\rho} > -F_{min} = \Delta F + F_3 \quad \Leftrightarrow \quad \Delta F < \frac{\Delta F_c - \rho F_3 + \lambda}{\rho - 1}$$

So in the interval $[\frac{\Delta F_c - \rho F_3}{\rho - 1}, \frac{\Delta F_c - \rho F_3 + \lambda}{\rho - 1}]$, the minimum is the knee on the interval border, i.e. $f_b = -F_{min} = \Delta F + F_3$. We can sum up:

$$f_b(F_3, \Delta F) = \frac{1}{\rho} \begin{cases} \Delta F + \Delta F_c & , \ \Delta F \in [-2F_3 \ , \ \frac{\Delta F_c - \rho F_3}{\rho - 1}] \\ \rho(\Delta F + F_3) & , \ \Delta F \in [\frac{\Delta F_c - \rho F_3}{\rho - 1} \ , \ \frac{\Delta F_c - \rho F_3 + \lambda}{\rho - 1}] \\ \Delta F + \Delta F_c - \lambda & , \ \Delta F \in [\frac{\Delta F_c - \rho F_3 + \lambda}{\rho - 1} \ , \ -F_3] \end{cases}$$

Hence we can write

$$f(F_1, F_3) = \begin{cases} f_a(F_1 - F_3) & , \ F_1, F_3 \geq 0 \\ f_b(F_3, F_1 - F_3) & , \ F_3 \geq 0 > F_1 \\ -f_b(F_3, F_1 - F_3) & , \ F_1 \geq 0 > F_3 \\ -f_a(F_1 - F_3) & , \ 0 > F_1, F_3 \end{cases} \tag{B-11}$$

with

$$f_a(\Delta F) = \frac{\Delta F - \operatorname{sgn}(\Delta F)\, \Delta F_c}{\rho} \Theta(|\Delta F| - \Delta F_c)$$

and

$$f_b(F_3, \Delta F) = \begin{cases} f_b^\alpha(F_3, \Delta F) & , \ \Delta F < -2F_3 \\ f_b^\beta(F_3, \Delta F) & , \ \Delta F \in [-2F_3, -F_3] \end{cases}$$

$$f_b^\alpha(F_3, \Delta F) = \frac{1}{\rho} \begin{cases} \Delta F + \Delta F_c & , \ \Delta F \in (-\infty \ , \ -\rho F_3 - \Delta F_c] \\ -\rho F_3 & , \ \Delta F \in [-\rho F_3 - \Delta F_c \ , \ -\rho F_3 - \Delta F_c + \lambda] \\ \Delta F + \Delta F_c - \lambda & , \ \Delta F \in [-\rho F_3 - \Delta F_c + \lambda \ , \ -2F_3] \end{cases}$$

$$f_b^\beta(F_3, \Delta F) = \frac{1}{\rho} \begin{cases} \Delta F + \Delta F_c & , \ \Delta F \in [-2F_3 \ , \ \frac{\Delta F_c - \rho F_3}{\rho - 1}] \\ \rho(\Delta F + F_3) & , \ \Delta F \in [\frac{\Delta F_c - \rho F_3}{\rho - 1} \ , \ \frac{\Delta F_c - \rho F_3 + \lambda}{\rho - 1}] \\ \Delta F + \Delta F_c - \lambda & , \ \Delta F \in [\frac{\Delta F_c - \rho F_3 + \lambda}{\rho - 1} \ , \ -F_3] \end{cases}$$

# C. Detailed description of the numerical solver for the distance-to-centre problem

In this appendix, we will provide additional details on the numerical method introduced in section V.1.5.

Step 2: Discretisation of the network. For the numerical solution, the network is modelled by a set of nodes which corresponds to stations where people can enter the subway. For each segment of network $(i, j)$ connecting two stations $i$ and $j$ we have a flow $F_{ij}$, such that the travel time along this segment reads

$$\tau_{ij} = l_{ij}(a + bF_{ij}),$$

where $l_{ij}$ is the distance of $i$ and $j$ along the network.

Step 3: We have to solve the routing problem for every starting point $l = 1, \ldots, N$. We assume that the traveller walks on a straight line from the starting point $\mathbf{x_l}$ to a station $s$ at $\mathbf{x_s}$ such that the walking time reads $\tau_w(l, s) = \|\mathbf{x}_l - \mathbf{x}_s\|/v_w$. Then he takes the shortest path along the network which requires the time $\tau_n(s)$, for which we use the uncongested travel time along the network as the flows are not known at routing time. Physically this corresponds to travellers who plan their travel path using the scheduled travel times of the subway. Hence the total travel time reads

$$\tau(\vec{x}_l, \vec{0}) = \|\mathbf{x}_l - \mathbf{x}_s\|/v_w + \tau_n(s). \tag{C-1}$$

Using a discretisation of the transportation network, the routing problem reduces to the choice of the optimal entry station $s$ to the network. Which station is chosen depends on the preferences of the traveller: They chooses $s$ to either minimise the walking time $\tau_w$ (lazy traveller model) or the total travel time $\tau(\vec{x}_l, \vec{0})$ (fast traveller model). Since the set of stations is finite, the optimal station $s^*$ can simply be determined by checking all alternatives. Exploiting the symmetry of setting, we can reduce the computation time for a network with $n$ branches to $1/n$ if we only check the stations from the correct branch as each branch covers a well defined section of the plane.

Step 4: Once the path to the centre has been decided for each of the $N$ starting points in step 3, the network flows are computed in the following way. Starting

from the centre, the flow on each link is just the sum of the inflows of all stations which lay on the same branch further outside. In general, one would need to solve a numerically expensive Quadratic Programme (QP) to get the optimal flows (cf. Sec. II.1). However, since we are considering in this distance-to-centre analysis only with trees and quasi-trees, the optimal flows are fully determined by Kirchhoff's current conservation law. In the loopy network we profit from the $Z_n$ symmetry, i.e. the symmetry of the cyclic group, of the city which ensures that the load on each branch is equal so that all the travellers entering the network on the loop will choose the path via the closest branch so that we can numerically handle the loopy network like a tree.

After calculating the flows, one can now update the $\tau_s(s_i)$ considering also the congestion term which are coupled to the $F_{0i}$. After calculating the flows, the $\tau_n(s)$ can be updated recursively: Given a station $i$, which is an outer neighbour of station $j$, one finds

$$\tau_n(i) = \tau_n(j) + l_{ij}(a + bF_{ij})$$

once we have $\tau_n(j)$. Starting at the central station 0, for which $\tau_n(0) = 0$, we can call this function for each outer neighbour of 0 and their outer neighbours consecutively.

This numerical algorithm includes two parameters: The number $N$ of randomly chosen starting points and the discretization of the public transportation network $\Delta l$, which can be interpreted as the distance between two stations. The limits $N \to \infty$ and $\Delta l \to 0$ restore the setup from Sec. V.2. We note, however, that keeping $\Delta l$ finite may be practically more meaningful as real transportation networks have only a finite number of access points. In all simulations we have set $\Delta l = 0.05\, r_0$; for a city with spatial extend $r_0 \approx 10\, km$, this corresponds to a distance of $\Delta l \approx 500\, m$ between two stations. Furthermore, variations of $\Delta l$ in this regime has only small influence on $\tau$. The number of starting points is chosen as $N = 10^4$ in all simulations, which is a good compromise between accuracy and numerical efficiency. To test the reliability of results, each computation is repeated several times and the standard deviation of these realisations is used as a measure of the uncertainty of the results. All error bars in the plots refer to this standard deviation. We usually averaged over 20 or 32 realisations.

# D. Additional figures

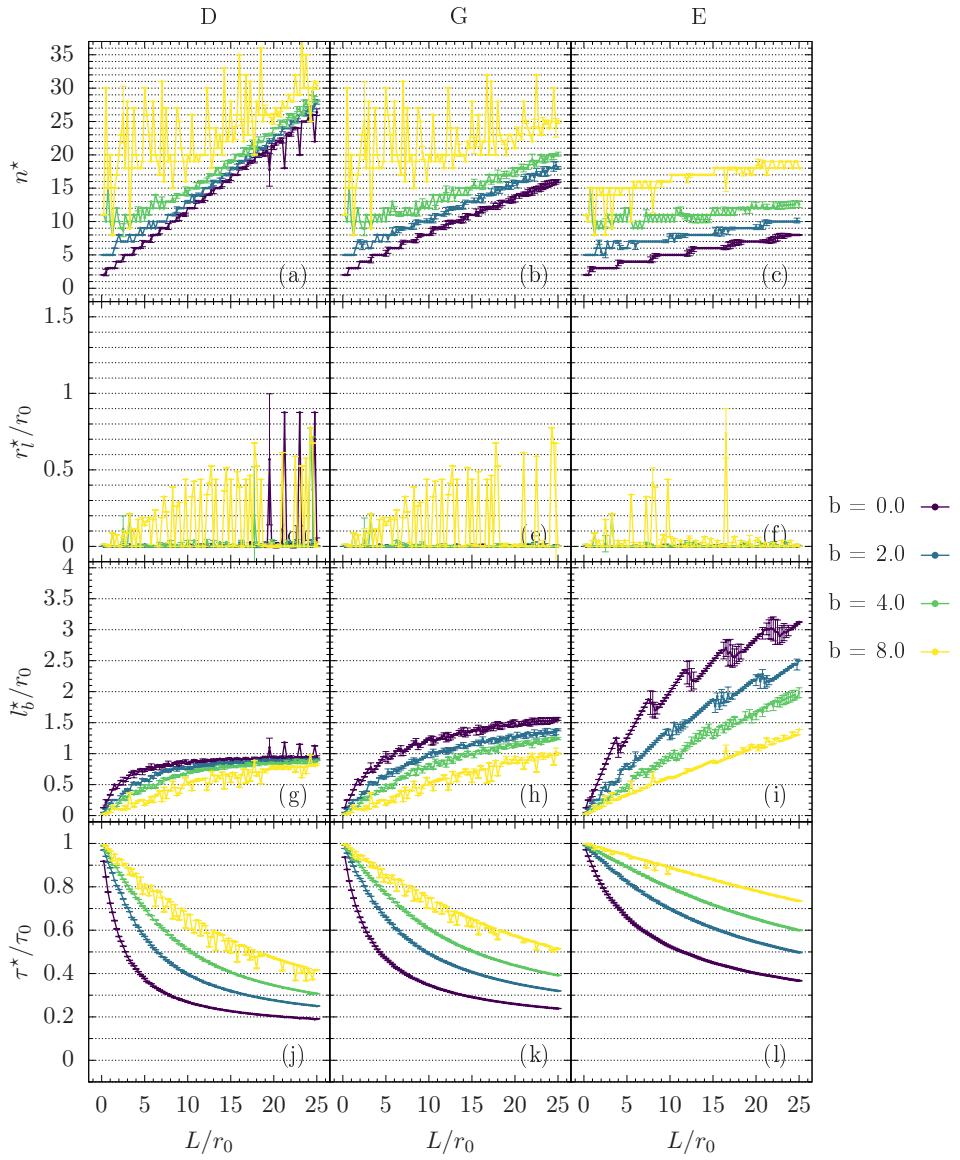



Figure D.1.: Optimal parameters for a star network with loop in the fast traveller model.