Optimising the geometry of transportation networks in the presence of congestion

Matthias Dahlmanns,[1,2] Franz Kaiser,[1,2] and Dirk Witthaut[1,2,*]

[1] *Forschungszentrum Jülich, Institute for Energy and Climate Research (IEK-STE), 52428 Jülich, Germany*
[2] *Institute for Theoretical Physics, University of Cologne, Köln, 50937, Germany*

(Dated: December 14, 2023)

Urban transport systems are gaining in importance, as an increasing share of the global population lives in cities and mobility-based carbon emissions must be reduced to mitigate climate change and improve air quality and citizens' health. As a result, public transport systems are prone to congestion, raising the question of how to optimise them to cope with this challenge. In this paper, we analyse the optimal design of urban transport networks to minimise the average travel time in monocentric as well as in polycentric cities. We suggest an elementary model for congestion and introduce a numerical method to determine the optimal shape among a set of predefined geometries considering different models for the behaviour of individual travellers. We map out the optimal shape of fundamental network geometries with a focus on the impact of congestion.

I. INTRODUCTION

The structure and design of optimal transportation networks plays an important role throughout different disciplines, ranging from biological systems [1–3] to man-made networks such as hydraulic networks [4, 5], power grids [6] and urban transportation systems [7, 8]. The design of a transportation network is governed by the task it is supposed to perform, e.g. minimising the dissipation, minimising the cost to build the network or minimising the averaged travel time.

Real-world supply and transportation networks display a variety of shapes depending on their history, their task and their surroundings [9–11]. For example, in the case of urban transport systems, the design of the network may depend on the city size and the degree to which cities are historically grown or centrally planned [12]. Although most of them have grown over decades and were built in several small steps, one observes several patterns which occur frequently in various cities and have been analysed using different topological indicators [13–16]. Building an efficient urban transportation network becomes increasingly important to reduce individual traffic which leads to congestion, high emission of green house gases [17] and adverse effects on air quality and citizens' health [18].

Congestion is a central topic in traffic research and planning. Empiric and numerical studies of road traffic show how vehicle velocities decrease with density up to a complete traffic jam, which is commonly summarised in the fundamental diagram of traffic flow [19, 20]. Methods to manage or optimise traffic flows in a given network are widely studied in the literature, see, e.g., [21] for a review. Congestion effects in public transportation networks have received an increased interest in recent years [22], where it affects both travel times [23] and the user comfort [24].

A variety of models has been developed to analyse and optimise the structure of transportation networks, including traffic networks [25–27] as well as other types of technological or biological networks [1–6]. Fundamental results have been obtained for the elementary case of uncongested networks, but empirical studies reveal a growing importance of congestion [28]. Empirical results on how congestion shapes the shape of cities were presented in [11]. Abstract network models that explore the impact of congestion on the optimal structure were introduced in [29, 30].

In this paper, we study the impact of congestion on the optimal shape of transportation networks [26]. We focus on three fundamental geometries that are frequently observed in subway or tram networks. Exemplary real world networks and the corresponding regularised geometries are shown in Fig. 1: Starting with a regular star shape (e.g. Saint Petersburg metro), we extend the network by allowing for two additional geometrical features that occur frequently in subway or tram networks: a cycle around the city centre (e.g. Moscow, Paris) and the branching of tracks in the outskirts (e.g. Hanover). For each geometry, we then optimise the network structure to minimise the travel time from the city to its centre for three radially symmetric models of the population densities as sketched in Fig. 1d. Finally, we map out the optimal geometry as a function of the available resources and the importance of congestion and analyse the transitions between different optimal shapes.

Our work builds on a previous article by Aldous & Barthelemy, which addressed the corresponding problem in the absence of congestion [26]. We introduce a versatile numerical method to determine travel time in the presence of congestion and thus choose the best network. Furthermore, we extend the investigation to polycentric cities and study two central properties of supply networks: the occurrence of loops or circle lines and the branching of tracks.

II. METHODS

The central objective of this article is the structure of optimal transportation networks: Given a limited budget, what is the optimal shape of the network such that

* d.witthaut@fz-juelich.de

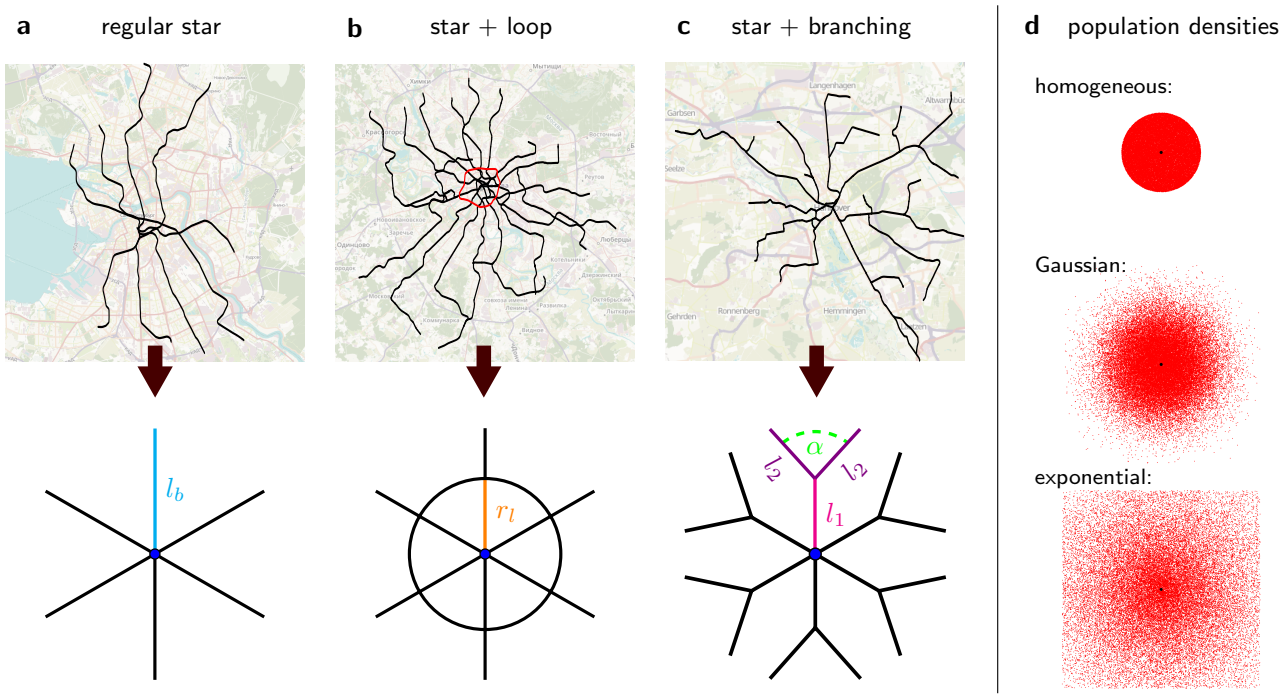

Figure 1. Structural patterns in urban transport networks. The geometries of public transport networks feature a large variety of structural elements leading to very individual shapes. However, some elementary patterns can frequently be observed. In this paper, we discuss the optimal shape of three of these fundamental patterns: (a) The shape of the Saint Petersburg metro is an example for a simple star network where several straight lines cross each other only in the city centre. This topology can be idealised to a *regular star* with n isotropical branches of length l_b. (b) The Moscow metro network features in addition a loop track around the city centre (red). In this paper, we thus secondly investigate a regular star extended by a concentric loop with radius r_l. (c) The Hanover tram network is a typical example for branching in the outskirts of the city. To study this type of geometries, we thirdly consider a star network where each branch splits at a distance l_1 from the city centre into two subbranches of length l_2 which span an angle α. (d) We study the optimal geometries for three different radial symmetric population density distributions: a homogeneous disk, a Gaussian density and an exponential density. Each red dot represents an inhabitant, which are randomly distributed according to the respective density function. The black dot marks the city centre. Map data based on OpenStreetMap [31].

the overall travel time is minimised? In the following, we will formalise this optimisation problem and introduce several key methods to solve it.

A. 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 $\mathbf{x} = (x_1, x_2)^\top \in \mathbb{R}^2$. We assume that the population density in the city is described by the function $\rho(\mathbf{x})$ and that all inhabitants want to travel from their home place to other places, which are described by a distribution of destinations $\rho_d(\mathbf{y}|\mathbf{x})$. We may then calculate the average travel time τ for all journeys in the whole city by integrating over the destination and population densities

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

where $\tau(\mathbf{x},\mathbf{y})$ is the travel time between two points $\mathbf{x}$ and $\mathbf{y}$. This expression for the average travel time τ is the central objective to be minimised. The solution of this optimisation problem depends on the properties of the city (via the functions $\rho(\mathbf{x})$ and $\rho_d(\mathbf{y}|\mathbf{x})$) as well as the methods of transportation that determine $\tau(\mathbf{x},\mathbf{y})$. In particular, we assume that the resources are limited, which is quantified by an upper bound to the total network length L. Hence, we will evaluate the optimal network geometry as a function of the total network length L throughout this article.

Cities and their spatial structures are complex systems that are subjects to ongoing research [32]. In classical urban economics, a fundamental approach to study cities is the monocentric model [33, 34], which assumes that most of the economic activities are concentrated inside a small area in the city centre, the *central business district* (CBD). We therefore focus the first part of the analysis on this fundamental city model. In the second part, we

generalise the analysis for non-monocentric cities as many cities reveal a decentralised and more complex spatial organisation [35–38].

Using the model of a monocentric city, we assume that most travellers want to go to the CBD and travellers seeking to reach other locations can be neglected. For simplicity, we assume the CBD to be point like and located in the origin $\mathbf{0}$ of the coordinate system. The distribution of densities then reads $\rho_d(\mathbf{y}|\mathbf{x}) = \delta(\mathbf{y})$. Imposing ρ_d into (1), the average travel time for a monocentric city becomes

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

Thus, it remains to minimise the average time-like distance to centre $\tau(\mathbf{x}, \mathbf{0})$ for each traveller.

B. The average travel time in multimodal traffic networks

The average travel time τ is essentially determined by the available modes of transport and their velocities. In this article, 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, then travelling along the network and then eventually walking again. Hence, the travelling time $\tau(\mathbf{x}, \mathbf{y})$ 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 $\mathbf{x}$ and $\mathbf{y}$ is simply given by $\|\mathbf{x} - \mathbf{y}\|/v_w$, where $\|\cdot\|$ denotes the euclidean distance.

Second, people may choose a transportation network, for example a subway network. This mode is typically faster than walking but also longer. We assume that travellers seek to minimise the total travel time τ, which is obtained by summing up the travel time along the path on the network.

In this article, we are especially interested in the impact of congestion within the transportation network on the average travel time τ and the optimal network structure. Congestion is taken into account by assuming that the travel time τ_l along a link l *increases* monotonically with the respective flow F_l. In particular, we assume a linear relation

$$\tau_l(F_l) = (a + bF_l)d_l\,, \tag{3}$$

where d_l is the length of link l and b a congestion parameter that is discussed in detail below. The parameter $a = 1/v_0$ is the inverse of the free-flow velocity v_0, i.e. the velocity in the limit $F_l \to 0$. For a subway network, a typical value for a is $a \approx 1/8$, assuming $v_w \approx 5\,\mathrm{km/h}$ and $v_0 \approx 40\,\mathrm{km/h}$. Throughout this article, we thus use $a = 1/8$.

The congestion parameter b describes how strongly the flow F_l increases the physical or equivalent travel time along a link l. It thus measures how susceptible the network is to congestion induced delays. The impact of congestion vastly differs between cities and times of the day (for example, rush hour vs. night). We thus keep b as a free parameter and show results as a function of $b \in [0, 4]$ in units of v_w over the flow density normalised by the city population. To get an impression of the meaning of this parameter, consider Eq. (3): The ratio a/b gives the amount of people that need to travel along a single line to double the travel time. For $a = 1/8$ and $b = 4$, we find that the travel time along a line is doubled when approximately 3 % of the city population take this line.

Before we proceed, we briefly comment on the fundamentals of the congestion model used in this study. Congestion effects are intensively studied for road traffic, where different functional relations $\tau_l(F_l)$ have been used [39, 40]. We can interpret Eq. (3) as a generic Taylor expansion of these functions up to linear order. We note that our optimisation model can easily be adapted to other congestion models by replacing the linear relation by any other function of $\tau_l(F_l)$. Congestion is also important in public transportation networks [22], where it can affect routing decisions in two ways. First, congestion can increase the physical travel times due to denied boarding or irregular vehicle arrivals [23]. Second, overcrowding reduces the comfort and thus the effective utility of travellers, which may choose alternative routes or modes of transportation [23, 24, 41]. Notably, the effect of discomfort has been quantified in terms of an equivalent increase of travel times in empiric studies [42], using a linear functional relationship as in Eq. (3).

Once we fixed the traffic flow model, it remains to determine for each traveller the stations where he enters and leaves the network. In the main part of the manuscript we assume a "lazy traveller" model, where each traveller uses the station that is next to his starting point and destination, respectively, to minimise the length of his walking path. Other routing strategies are discussed and evaluated in the appendix A. We find that that the results are very similar such that we focus on one strategy in the main text.

C. Numerical optimisation

To solve the integral in (2) for complex city and network shapes, we developed a versatile method that evaluates the average travel time τ using a discretisation of both the network, the starting points and the destinations. In particular, the solver proceeds as follows (cf. Fig. 2):

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

2. Place destinations according to the current model of the city structure by drawing at random from

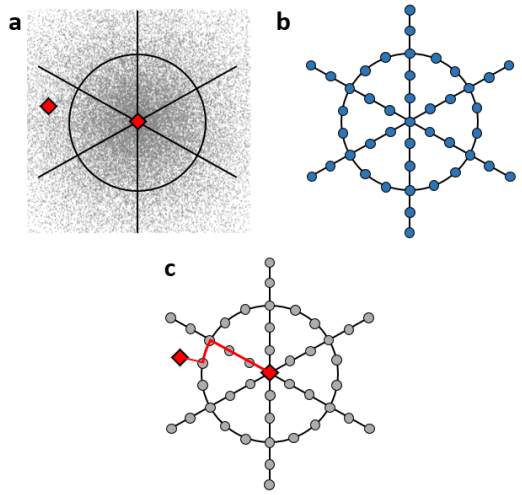

Figure 2. Main steps of the simulation scheme. (a) The origin and destination of each traveller (red diamonds) is drawn at random from the distributions $\rho(\mathbf{x})$ and $\rho_d(\mathbf{y}|\mathbf{x})$. (b) The network (black lines) is initialised using one of the three geometries shown in Fig. 1 and stations (blue disks) are placed along the network. (c) The traveller chooses a path (red line) from origin to destination according to a specified routing strategy. Finally the travel time is computed taking into account the effect of congestion, and summed up over all travellers.

the distribution $\rho_d(\mathbf{y}|\mathbf{x})$.

3. Add stations for entry and exit to a given network. Place one station in the centre (blue circle in Fig. 1b). Proceed outwards on the radial branches and place stations in intervals of length Δl. Add one additional station at the end of the radial branches. If the network contains a loop or branches, a station is added at each crossing point. Further stations are added along a loop such that their mutual distances are equal and as close to Δl as possible.

4. Routing: Compute the optimal path to the destination for each starting point according to the corresponding routing strategies. Travellers can only access or leave the network at a station. Thus, we seek the station for each traveller where they enter and leave the network.

5. Compute the flow F_l and the resulting travel time $\tau_l(F_l)$ for every segment l of the network, that is a connection between neighbouring stations. The flow F_l is directly proportional to the number of travellers using this segment of the network.

6. Sum up the travelling times for each starting point to obtain $\tau = \tau_w + \tau_s$. While the average walking time τ_w equals the Euclidean distance from each starting point to its access station of the network plus the distance from the exit station to the destination in appropriate units. The average travel time τ_s inside the network, is the sum of the travel time τ_l for each segment l of the network multiplied with the local flow F_l.

Further details of the numerical solver are described in appendix C. Throughout the paper, we use $\Delta l = 0.05\, r_0$, which corresponds to a station distance $\Delta l \approx 500\,\mathrm{m}$ for a city of size $r_0 \approx 10\,\mathrm{km}$.

We note that all users individually compute an optimal path. Traffic research generally distinguishes between user equilibrium traffic and system optimal traffic, which do not necessarily coincide [43]. The computation is substantially complicated when stochastic fluctuations or traffic information is taken into account [43, 44]. Furthermore, we note that in real public transportation networks stations are typically not placed at a constant distance but locations are adapted to the demand. The optimal placement of stations is beyond the scope of this paper.

D. Population densities

As the optimal network shape depends on the distribution of the population in the city, we performed the optimisation for three different radial symmetric population density functions $\rho(r)$ as shown in Fig. 1d:

1. a compact homogeneous disk with a constant density up to a distance r_0 from the city centre with $\rho_{\mathrm{hom}}(r) = \rho_0\,\Theta(r_0 - r)$, where Θ denotes the Heaviside step function.

2. a Gaussian density with $\rho_{\mathrm{gauss}}(r) = \rho_0 \exp(-r^2/r_0^2)$, where a small part of the population is located further away from the city centre and

3. an exponential density with $\rho_{\mathrm{exp}}(r) = \rho_0 \exp(-r/r_0)$ which yields a widely spread population.

The prefactor ρ_0 is chosen such that the overall population is normalised to $\int \rho(r)\, dr = 1$. The factor r_0 classifies the typical scale of the city. Throughout this manuscript, we set $r_0 = 1$ and express all lengths in units of this parameter. For real cities, the value of r_0 typically lays in the order of a few kilometres. We note that the exponential distribution is commonly regarded as the best model for the population density of a city, based on both empiric investigations and theoretic economic models [45].

E. Travel time without a network

To quantify the benefits of a transportation network, we compare the average travel time to the case without any transportation network.

In a monocentric city without a transportation network, we have $\tau(\mathbf{x}, 0) = \|\mathbf{x}\|/v_w$ and the integral in (2) can be evaluated in closed form for the three different models of the population density

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

The ratio $\hat{\tau} = \tau/\tau_0 \in [0, 1]$ is the reduction of the average travel time by the network and thus quantify its effectiveness. We therefore use $\hat{\tau}$ as a measure of network performance throughout this paper.

III. RESULTS

A. Optimal shape of a regular star network

We first analyse the optimal geometry of a regular star network without loops and branching. The Saint Petersburg metro is an example for such a network as visualised in Fig. 1a. The geometry is optimised by choosing the number of branches n such that, for fixed total length L of the network and congestion parameter b, the averaged travel time τ assumes its minimum. The results are plotted in Fig. 3 over the network length L.

As expected, both the optimal number of branches $n^\star$ (Fig. 3a-c) and the corresponding branch length $l_b^\star = L/n^\star$ (Fig. 3d-f) increase with the amount of available resources L. The optimal length $l_b^\star$ increases rapidly at first and then saturates, where the saturation level differs strongly for three models of the population density $\rho(\mathbf{x})$. In the homogeneous disk model, all travellers start at a distance below r_0 such that a track outside this radius will not be used. Hence, the optimal length $l_b^\star$ converges to r_0 from below. In the Gaussian model, the population density drops rapidly for $r > r_0$ such that $l_b^\star$ saturates for values slightly above 1. In the exponential model, the population density decreases much slower. Hence, $l_b^\star$ also saturates slower and at higher values. In contrast the optimal number of branches $n^\star$ does not saturate and grows almost proportionally with L. This implies that, with limited resources, one should first invest in the elongation of existing lines and only then invest into building new lines. In the following, we will mostly restrict our analysis to one population density for the sake of clarity.

When varying the congestion parameter b, we find that $n^\star$ increases with b and, consequently, $l_b^\star$ decreases. Since a growing congestion adds a penalty for strong flows on single branches, it is reasonable that in the presence of congestion more and shorter branches are preferred over fewer longer ones. Notably, congestion in real networks depends also on the frequency and capacity of trains, which is assumed to be constant in the current model.

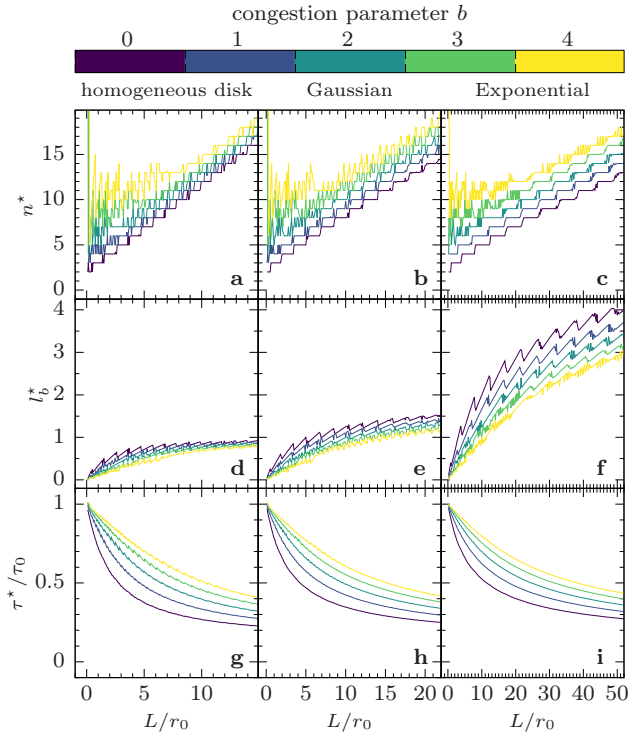

Figure 3. Impact of congestion on optimal star geometries. We plot the optimal parameters for a regular star network (cf. Fig. 1a) over the network length L for different values of the congestion parameter b (colour code): (a)-(c): The optimal number of branches $n^\star$, (d)-(f): the corresponding branch length $l_b^\star = L/n^\star$, (g)-(i): the optimal averaged travel time $\tau^\star$. We use the "lazy traveller" model for routing and compare three different models of the population density (cf. Fig. 1d). We find that the optimal branch length l_b^* saturates depending on the population density model. Congestion favours more and shorter branches over fewer and longer ones.

B. Benefit-cost ratio for a regular star

In most applications, not only the optimal shape of the network needs to be determined but the benefit-cost ratio must also be evaluated to decide whether an investment is worth it. We assume that the construction costs are proportional to the length L of the transportation network and define the benefit as the reduction of the optimum averaged travel time $\tau^\star$. The marginal

$$g(L) := -\frac{d\tau^\star}{dL} \tag{5}$$

then gives the benefit-cost ratio for enlarging the network. The minimal averaged travel time $\tau^\star$ is plotted in Fig. 3g-i in units of the averaged travel time in absence of the network τ_0. The corresponding benefit-cost ratio (5) is shown in Fig. 4 for the Gaussian population density.

For small networks, we find a high benefit-cost ratio while this value decreases approximately exponentially with growing networks, i.e. the longer the network, the

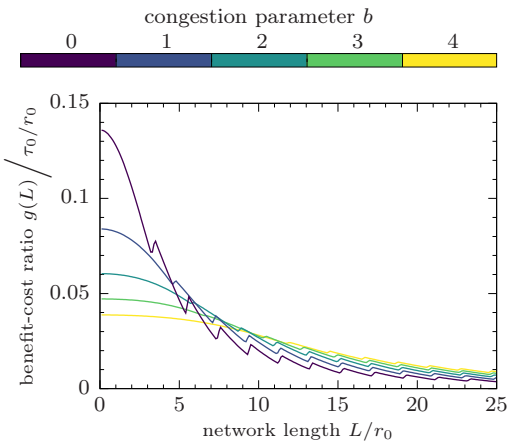

Figure 4. Cost benefit ratio for network expansion. The cost benefit ratio $g(L) = -d\tau^\star/dL$ for an optimal star generally network decreases with the network length L. In the presence of congestion, $g(L)$ is smaller for short network, but decreases slower with increasing L. Hence, extending the network becomes increasingly beneficial with higher levels of congestion, i.e. $g(L)$ increases with b, when the network is already large ($L \gg r_0$). We use the Gaussian population density model and the "lazy traveller" model.

less benefit we get from adding a unit length to the network. Considering congestion, we find that, with L fixed, $\tau^\star$ grows with the congestion parameter b as a consequence of the reduction of speed within the network due to congestion. Therefore, the benefit-cost ratio $g(L)$ is lowered in the presence of congestion for short networks. In the case of long networks, however, we find that the benefit-cost ratio is increased by b. Nevertheless, the total benefit, i.e. the value of $\tau^\star$, is always lower for stronger congestion (cf. Fig. 3g-i). Thus, when building a network from scratch, the presence of congestion always lowers the benefit for a given investment. On the other hand, if an existing network should be enlarged, congestion might even increase the benefit for a given investment.

C. Optimal loop radius

In the next step, we investigate a modification of the simple star network by adding a loop track around the city centre as sketched in Fig. 1b which is commonly observed in real public transportation networks such as the Moscow metro, the Paris metro and the Cologne tram network. We use the numerical approach introduced in the Sec. II to determine the averaged travel time τ. A regular star network with a loop is parametrised by two quantities: the number of branches n and the radius of the loop r_l. Given the total length of the network L, the branch length l_b is then determined by

$$L = nl_b + 2\pi r_l \quad \Rightarrow \quad l_b = \frac{L - 2\pi r_l}{n}\,. \tag{6}$$

To find the optimal values of the geometrical parameters $n^\star$ and $r_l^\star$, we have scanned the parameter space and selected the values which minimise τ. As before, we analyse the results in dependence of the available resources L and the congestion parameter b.

The first important finding is that a pure star network is always superior to a loopy network for given resources L and a monocentric city. When optimising over both n and r_l, we always find the optimal loop radius $r_l^\star = 0$. This result is a direct consequence of the specific optimisation problem considered in this paper: All travellers want to go 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 other positions than the city centre, cf. Sec. III E.

Nevertheless, when considering n to be fixed we find for certain parameters a finite value for the optimal loop radius. Thus, if the number of branches n is fixed, there are parameter settings where including a loop into the network becomes beneficial. Such a situation can occur in practical applications where the geographical situation or the structure of the city might bias the choice of n but also where a star shaped network is already present and space for further branches is not available. Then the question arises if the budget should be fully invested to extend the existing branches or if a loop should be established.

We now further investigate this scenario and fix $n = 4$ in the following. Furthermore, we focus on the Gaussian population density and the "lazy traveller" model, as all other models yield qualitatively similar results. The optimal loop radius $r_l^\star$ as well as the corresponding branch length $l_b^\star$ in this case are plotted in Fig. 5a-b.

Most importantly, we find a discontinuous transition for the optimal network shape. For short networks $L < L_{\text{crit}}$, the optimal structure is again a pure star with $r_l^\star = 0$. As L increases above a critical value L_{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_{crit} – i.e. the loop comes into being with a non-zero radius. Correspondingly, the optimal branch length $l_b^\star$ jumps around L_{crit} to a lower value as parts of the resources are now needed for the loop. For $L > L_{\text{crit}}$, the optimal radius $r_l^\star$ increases with L to some extent and then saturates at a value $R_l := \lim_{L\to\infty} r_l^\star$. We note that such a discontinous transition was already observed by Aldous & Barthelemy [26] and rigorously established for a different type of optimal networks by Kaiser et al. [6].

The discontinuity can be explained by plotting the averaged travel time $\tau(r_l)$ as a function of the loop radius r_l. For $n = 4$ and $b = 1$ fixed, Fig. 5e visualises these curves for different network lengths L around the critical network length, which is in this case $L_{\text{crit}} \approx 4\,r_0$. For $L < L_{\text{crit}}$, the curves $\tau(r_l)$ are strictly monotonically increasing such that the minimum is always located at $r_l = 0$. At $L = L_{\text{crit}}$, the curve becomes flat around

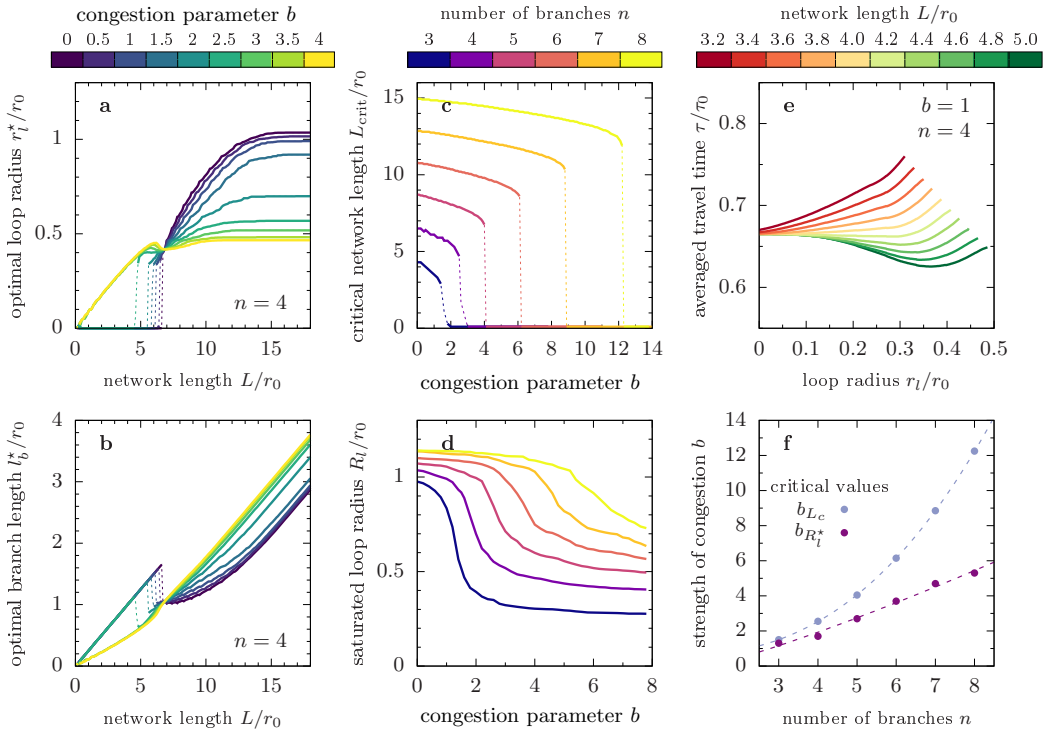

Figure 5. Transition from non-loopy to loopy transportation networks is discontinuous. We plot the optimal parameters for a star with a concentric loop (cf. Fig. 1b) over the network length L for different values of the congestion parameter b using the Gaussian population density. the number of branches n is kept fixed in the optimisation. (a) The optimal loop radius $r_l^\star$ jumps at a critical network length L_{crit} from zero (no loop) to a finite value. (b) Correspondingly, the optimal branch length $l_b^\star$ is also discontinuous at this point. (c) The critical parameter L_{crit} is also discontinuous when varying the congestion parameter b: For a critical value b_{L_c} of the congestion parameter L_{crit} jumps from a finite value to zero. (d) When considering the saturation values of $r_l^\star$ in the limit of large networks, denoted by $R_l^\star$, we also find a sharp transition from a "low congestion" to a "high congestion" phase around a critical congestion parameter $b_{R_l^\star}$. In this case, however, the transition is much smoother compared to the transitions in (a)-(c). (e) When plotting the averaged travel time τ over the loop radius r_l, we find an explanation for the discontinuity of $r_l^\star$: Increasing L flattens the $\tau(r_l)$ curve until at the critical length L_{crit}, the minimum of the curve flushes from $r_l = 0$ to a finite value. (f) While $b_{R_l^\star}$ scales linearly with the number of branches n, we find $b_{L_c} \propto n^{2.5}$.

$r_l = 0$. For $L > L_{\mathrm{crit}}$, the curves $\tau(r_l)$ have a negative slope around $r_l = 0$ and a new minimum emerges for positive values of r_l.

The saturation loop radius R_l and the critical network length L_{crit} are further investigated in Fig. 5c-d. Both quantities depend on the number of branches n and the congestion parameter b. In both cases, we can distinguish a low and a high congestion phase with relatively stable values and very sharp transitions from one phase to the other. In the low congestion phase, the saturated loop radius satisfies $R_l \approx r_0$. That is, the loop is established in the outskirts of the city if the available resources permit. We conclude that the prime function of the loop is to collect travellers from locations in the outskirts far away from the radial branches. In the high congestion regime, the loop is established much closer to the city centre, (R_l is much smaller than r_0), pointing to a very different function of the loop.

We further investigate the transition points between the low and the high congestion phase. We define the critical congestion parameters b_{L_c} and $b_{R_l^\star}$ as the points were the negative slopes of $L_{\mathrm{crit}}(b)$ and $R_l^\star(b)$, respectively, assume their maximum. Remarkably, the critical values do not agree and show a different scaling behaviour with the number of branches n. We find a linear scaling for $b_{R_l^\star} \propto n$ and a non-linear scaling $b_{L_c} \propto n^{2.5}$ with the number of branches n shown in Fig. 5f. The linear scaling of $b_{R_l^\star}$ is a consequence of the choice of units. While we define the population of the entire city as a unit of traveller, the total population within the catchment area of a single branch is just $1/n$ of the city population. Thus, the flow on a branch scales with $1/n$ and therefore the impact of congestion do. As a consequence, the critical value of the congestion parameters should scale with n. The reason for the non-linear scaling of b_{L_c} is subject to further research.

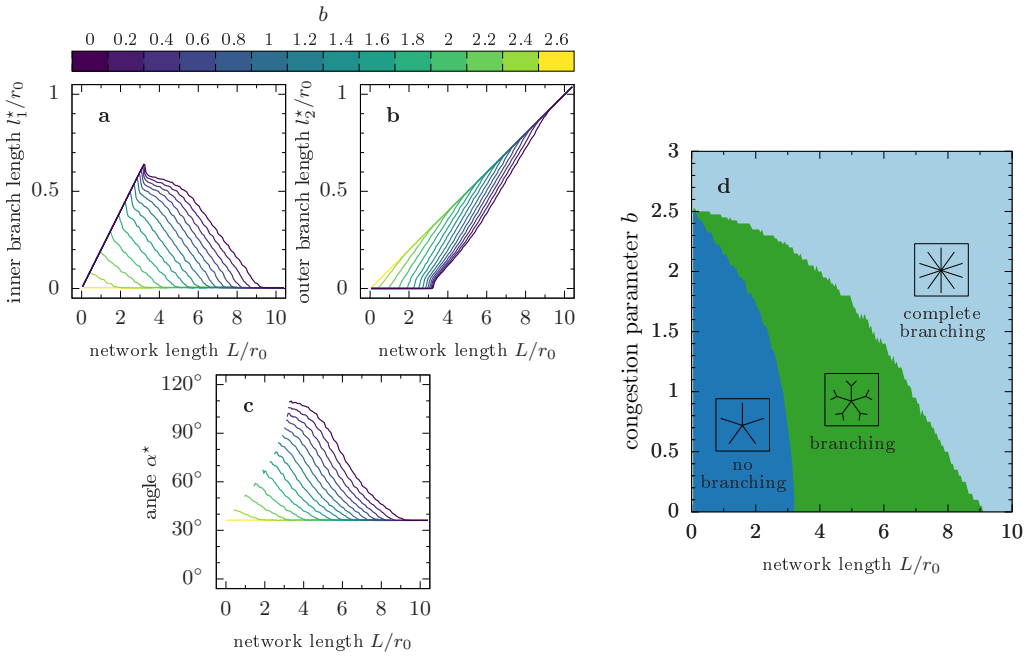

Figure 6. Transitions in optimal branching topologies are continuous. We plot the optimal parameters for a star with branching (cf. Fig. 1c) over the network length L for different values of the congestion parameter b using the homogeneous population density and fixing the number of branches to $n = 5$. (a,b) Contrary to the loopy topology (cf. Fig. 5), we observe continuous transitions from one shape to another as both the optimal inner branch length $l_1^\star$ and the optimal outer branch length $l_2^\star$ are smooth functions of L. (c) For long networks, the optimal opening angle $\alpha^\star$ spanned by the outer branches converges to $360°/2n$ as $l_1^\star$ goes to zero, i.e. the branching topology passes into a regular star with $2n$ branches. (d) The phase diagram for the optimal shape at each point in the L–b plane reveals the impact of congestion: the stronger b, the smaller the critical network lengths at with the optimal geometry switches from one shape to another. We can spot a critical value of $b \approx 2.5$ above which the regular star with $n = 10$ branches is optimal for all L. When comparing this geometry with the loopy case, we find branching to be always superior to the loop.

D. Branching

As a second extension to the simple star network, we consider a splitting of the branches at a distance l_1 from the centre (cf. Fig. 1c). Such a pattern allows the network to better reach the outskirts by saving costs through the sharing of tracks close to the city centre, where neighbouring branches are close to each other. The network shape in this case is characterised by three parameters: the number of branches n, the length of the inner branch l_1 and the angle α between the split branches. The length of the outer branches is then given by

$$l_2 = \frac{L - nl_1}{2n} \,. \tag{7}$$

The optimal parameter values are again determined using the numerical method introduced in Sec. II and a scan of the parameter space.

When optimising over n, we again find the optimal geometry to be always a regular star. As in the preceding section, we thus consider the scenario of a fixed number n. We conduct the discussion here for $n = 5$ and the homogeneous population density. Given a certain amount of resources L we now have the decision whether to enlarge or split the branches to better cover the outskirts of a city. Keeping n fixed, we can distinguish three optimal shapes:

1. n star: for small networks with $L < L_1$, the optimal length of the outer branches $l_2^\star$ is zero, i.e. the optimal shape is a regular star with n branches,

2. Branching: for intermediate networks with $L_1 < L < L_2$, we find both $l_1^\star$ and $l_2^\star$ to be non-zero, i.e. the optimal geometry contains branching,

3. $2n$ star: for large networks $L > L_2$, the length of the inner branch $l_1^\star$ diminishes and the angle spanned by the branches becomes $\alpha = 360°/2n$, i.e. the optimal network corresponds to a regular star with $2n$ branches.

The optimal geometry parameters $l_1^\star$, $l_2^\star$ and α are plotted in Fig. 6a-c over the network length L for different values of the congestion parameter b. The corresponding optimal shape for each combination of L and b is visualised in Fig. 6d. In contrast to the loopy case (cf. Fig. 5), we observe smooth transitions from one shape

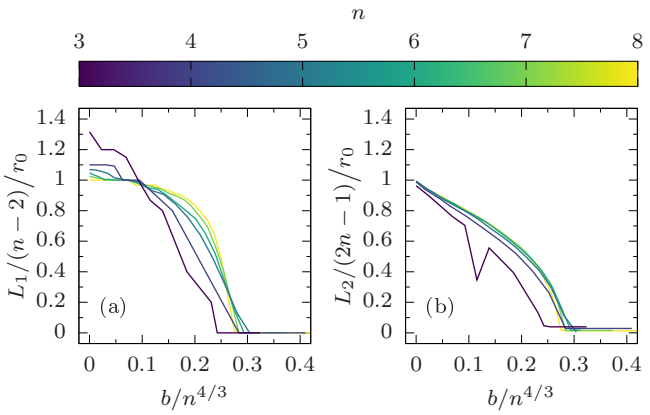

Figure 7. Impact of congestion on the critical network length $L_{1,2}$. The numerical functions $L_1(n,b)$ and $L_2(n,b)$ largely collapse when they are plotted vs. $b/n^{4/3}$ and rescaled by $(n-2)$ and $(2n-1)$, respectively. We thus conclude the scaling laws given in Eq. (8).

to another in the branching case: When passing the critical network length L_1, the outer branch length $l_2^\star$ starts growing linearly with L without any significant jumps. Similarly, $l_1^\star$ linearly diminishes when approaching L_2. Thus, both phase transitions are continuous in this geometry.

Considering the phase diagram in Fig. 6d, we find that both critical values L_1 and L_2 decrease with growing congestion parameter b until they both vanish, $L_1 = L_2 = 0$, for $b \approx 2.5$. Beyond these values of b, the regular star with $n = 10$ branches is always superior to both other shapes. This behaviour is consistent with our previous results on regular star networks. For large values of b, it is more important to distribute travellers onto many lines to mitigate congestion than to supply the outskirts of the city. Hence, congestion favours more and shorter branches over fewer longer ones.

We further quantify the dependence of the critical network lengths L_1 and L_2 on the choice of n. Empirically, we find that the impact of b scales with $n^{-4/3}$ while for b fixed, both L_1 and L_2 scale linearly with n. We thus deduce the scaling laws

$$\begin{aligned} L_1(n,b) &\approx (n-2)\tilde{L}_1(b/n^{4/3}) \\ L_2(n,b) &\approx (2n-1)\tilde{L}_2(b/n^{4/3})\,, \end{aligned} \tag{8}$$

with univariate functions $\tilde{L}_{1,2}$. Indeed, plotting $L_1/(n-2)$ and $L_2/(2n-1)$ as a function of $b/n^{4/3}$ the function largely collapse to $\tilde{L}_{1,2}$ as shown in Fig. 7. The reason for this empirically found scaling is subject to further research.

Considering the optimal angle $\alpha^\star$ between the outer branches plotted in Fig. 6c, we find that for small l_2, the angle starts at a large value and smoothly decreases to $\tilde{\alpha}^\star = 360°/2n$ for $l_1 \to 0$. Thus, the longer the outer branches, the closer they should lay together, but in any case the angle between them should be larger or equal to the angle of neighbouring branches in a $2n$ star.

E. Optimising transportation networks in polycentric cities

Many cities reveal complex spatial structures, that are shaped by a variety of parameters. In particular, growing cities typically experience a transition from a monocentric to a polycentric structure, which is strongly related to the limitation of traffic networks [37]. New subcenters or central business districts (CBDs) emerge at a certain distance, often where radial and peripheral highways cross each other (see, e.g., [36])

Here, we generalise our analysis using a model for polycentric cities. Besides the CBD in the city centre, N_c additional activity centres are isotropically distributed around the CBD at a distance R_c. In this model, we thus have $1 + N_c$ possible destinations for each traveller. We assume, that every traveller still wants to go to exactly one destination, which can be thought of, for instance, as the location of his working place. The precise mapping of each traveller's origin and destination depends on a variety of factors and is subject to current research [38]. Here, we assume the well established gravity model [46], where the probability $\rho_d(\mathbf{y}|\mathbf{x})$ of a traveller living at $\mathbf{x}$ getting mapped to a destination at $\mathbf{y}$ is proportional to the inverse of their Euclidean distance

$$\rho_d(\mathbf{y}|\mathbf{x}) \propto \frac{1}{||\mathbf{x}-\mathbf{y}||}\,. \tag{9}$$

In Fig. 8, the mapping is visualised for a city with an exponential population density and $N_c = 6$ subcentres at a distance $R_c = r_0$ from the centre.

We now investigate the impact of additional subcentres on the optimal loop radius compared to the monocentric city discussed in Sec. III C. We keep the number of subcentres $N_c = 6$ fixed but vary their distance from the city centre. The symmetry of this city suggests to implement a network with $n = N_c = 6$ branches that are placed such that each branch points into the direction of a subcentre. We thus focus the discussion on a star shaped network with 6 branches plus a single loop around the city centre. In particular, we consider the optimal loop radius r_l^* for different values of R_c (9a). We note that the monocentric case is equivalent to $R_c = 0$. To keep the analysis clear, we here discuss only the uncongested scenario $b = 0$.

We find that a pure star network is the best choice if the available resources are sparse, that is, if the total network length L is below a critical value L_{crit}. Remarkably, the critical value L_{crit} for the emergence of a loop is largely independent of the position of the subcenters R_c as long as they exist ($R_c > 0$).

If L is increased beyond L_{crit}, the optimal loop radius r_l^* rapidly increases to match the position of the subcenters R_c. This finding is intuitive as travellers can now use the loop track and exit the network directly at a respective subcenter. In the numerical results, we thus find

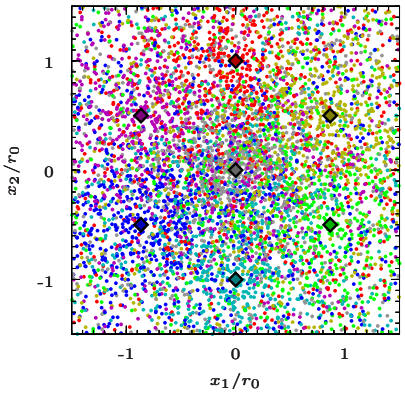

Figure 8. Destination mapping using the gravity model. In a city with a population density, that decays exponentially with the distance from the city centre, each inhabitant is mapped to exactly one destination using the gravity model given in Eq. 9. Destinations are visualised as diamonds of different colour. Each inhabitant is shown as a dot coloured according to his destination.

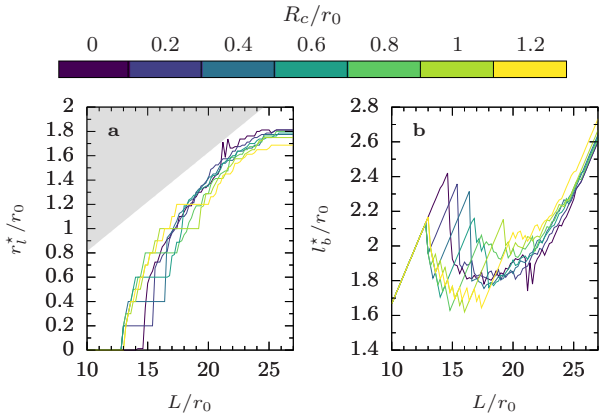

Figure 9. Optimal loop radius in a polycentric city. Assuming a city with $N_c = 6$ subcentres at a distance R_c from the city centre, we optimise the parameters of a network with $n = 6$ radial branches and one loop. (a) The optimal loop radius $r_l^\star$ is plotted over the total network length L for different values of R_c. The grey area marks the parameter range that would correspond to a disconnected network where the network branches don't reach the loop, i.e. where $r_l > l_b$. (b) The corresponding optimal branch length $l_b^\star$ is plotted over the network length L. We assume an exponential population density and the "lazy traveller" model.

a plateau where $r_l^* = R_c$. For even higher value of L, the loops leaves the subcenters and r_l^* increases beyond R_c before it finally saturates. In this regime, the loop track provides an improved accessibility to the network for travellers from the outskirts of the city. The benefits of this effect outweighs the benefits of the direct access to the subcenters.

IV. DISCUSSION

In this article, we have analysed the optimal shape of transportation networks for multimodal urban traffic in the presence of congestion. In the case where all travellers travel to the city centre, a regular star network is always superior to a network with loops or branches in terms of overall travel time. However, this strict result only holds if the number of branches n can be freely adapted as the network is extended. In practice, multiple geographic constraints exists and the question arises whether one should rather invest into simple line extensions, branches or an additional loop track. Our results show that the answer to this question strongly depends on the available resources and the importance of congestion.

Our detailed analysis of loopy and branching networks has led to four main results:

(i) The optimal shape of a loopy network is subject to discontinuous transitions. That is, the optimal loop radius $r_l^\star$ varies discontinuously, jumping from zero to a finite value, as the amount of available resources L is increased. Remarkably, strong congestion can qualitatively alter this scenario and even suppress the discontinuity.

(ii) In contrast, the optimal shape of a branching network varies smoothly with the amount of available resources L. In fact, the optimal shape evolves from no branching to finite branching to complete branching continuously as L is increased. Complete branching is always beneficial when the congestion is dominant.

(iii) Given that we focus on travelling to the city centre, branching networks are always superior to loopy networks.

(iv) Finally, congestion generally favours more and shorter branches over fewer longer ones. That is, it becomes more important to distribute the travellers on many lines to mitigate congestion than to reach the outskirts.

We believe that the value of our approach lies not only in these findings, but also in the developed methodology. We have introduced a model for congested multimodal transportation systems, extending prior studies such as [26]. To account for the mathematical complexity of the optimisation problem, we have developed a versatile numerical simulation framework, which can easily be adapted to a wide range of network geometries as well as different population densities, routing strategies and congestion models. A distinguished feature of our simulation model is that it includes the actual routing strategies of individual travellers. In the current study, we did not find a significant impact of the different strategies on the optimal network shape. However, when considering more complex scenarios, the routing behaviour might gain importance. In the case of concurrent congested transport

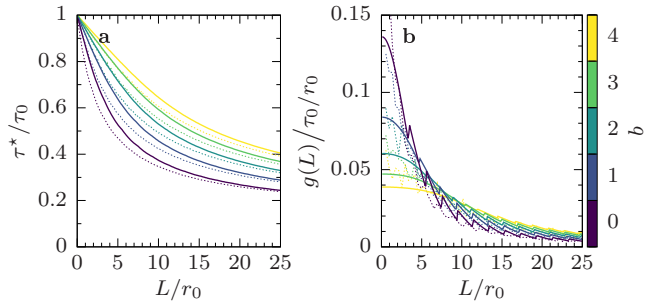

Figure 10. Comparison of different routing strategies. We plot (a) the optimal travel time and and (b) the cost-benefit ratio for different routing strategies: Results for the "lazy traveller" and the "fast traveller" are visually indistinguishable (dashed lines). Optimal travel times in the 'polar grid model" are slightly higher. Results are shown for the Gaussian density.

networks where other modes of transport such as street traffic are present, congested networks have been shown to be prone to paradoxical behaviour, where an extension of infrastructure actually increases congestion [47]. The proposed simulation model can be readily extended to investigate such phenomena and to include other nonlinear congestion models.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the Federal Ministry of Education and Research (BMBF grant no. 03EK3055B to D.W.).

Appendix A: Routing strategies

The travel time τ from a starting to a terminal point is determined by the chosen path. In general, travellers seek to minimise τ, but different constraints apply leading to slightly different routing strategies listed below. In all cases we assume that the traveller has no actual information about the load of the network. Hence, routing decisions are based on the travel times for $F_l = 0$.

1. Fast traveller model: The traveller aims to minimise his individual travel time. He is free to move in the plane and chooses the entry/exit stations such that the sum of the travel times for walking and network travel assumes a minimum.

2. Lazy traveller model: The lazy traveller aims to minimise the walking time. Hence, he always uses the entry/exit stations that are nearest to his starting or terminal point, respectively.

3. Polar grid model: We finally consider a model where travellers can only use a polar street grid. In this case, travellers will first walk inwards on a radial path until they reach the radial distance of the closest access point to the network. Then, they follow the spherical path to this point to enter the network. Using this assumption, we give an analytical expression for the average travel time in a regular star shaped network in the appendix B. This model allows for a closed analytic expression for the average travel time for certain geometries, see appendix B.

We find that the results for the "lazy traveller" and the "fast traveller" are virtually indistinguishable. The "polar grid model" includes stricter constraints such that the average optimal averaged travel time $\tau^\star$ is slightly larger and decreases slower with the network length L. (Fig. 10a). Similarly, the cost-benefit ration $g(L)$ is smaller for $L \to 0$, and decreases slower with L as for the other routing strategies. As the polar grid has been introduced to enable analytic computations, we focus on the lazy traveller model for all numerical simulations.

Appendix B: Analytical solution for travel time in congested star networks

In this appendix, we derive analytical expressions for the averaged travel time τ in a city with a subway network of length L in the shape of a regular star with n branches using the "polar grid" routing strategy.

So consider a star network with n branches of length $l_b = L/n$ as depicted in Fig. 1a. 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. Furthermore, we assume that each traveller takes the shortest possible way towards the network following a radial or spherical path.

In the following, we denote the starting point as $\mathbf{x} = (x_1, x_2)$ in cartesian coordinates or by the radius r and the azimuth θ 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 $\theta \in [-\theta_n/2, +\theta_n/2]$ with $\theta_n = 2\pi/n$. Then the travel time of a single traveller is written as

$$\tau(\mathbf{x}, \mathbf{0}) = \frac{d_w(\mathbf{x})}{v_w} + \int_0^{d_s(\mathbf{x})} \frac{dr}{v_0(r)} \,.$$

The first two quantities, d_w and v_w denote the distance and velocity, respectively, that is spent walking which are used to calculate 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 spent 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 $v_0(r) = (a + bF(r))^{-1}$,

this becomes

$$\tau(\mathbf{x},\mathbf{0}) = d_w(\mathbf{x}) + a\, d_s(\mathbf{x}) + b \int_0^{d_s(\mathbf{x})} dr\, F(r). \tag{B1}$$

The flow $F(r)$ on a branch at radial position r is proportional to the population living in the considered sector at a radial distance above v. Hence we obtain the local flow on the branch by integrating over the population density of the area in which all travellers contribute to the flow at radius d_s

$$F(r) = \int_{-\theta_n/2}^{\theta_n/2} d\theta' \int_r^\infty dr' r' \rho(r') = \theta_n \int_r^\infty dr'\, r' \rho(r'). \tag{B2}$$

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 $r \le l_b$ yields the contribution

$$\begin{aligned}\tau_1 &= \int_0^{l_b} dr\, r \int_{-\theta_n/2}^{\theta_n/2} d\theta\, \tau(\mathbf{x},\mathbf{0})\, \rho(r) \\ &= \theta_n \int_0^{l_b} dr\, r\, \rho(r) \Bigg[\left(\frac{\pi}{2n} + a\right) r \\ &\quad + \theta_n b \int_0^r dr' \int_{r'}^\infty dr''\, r''\, \rho(r'') \Bigg]\end{aligned} \tag{B3}$$

to the total travelling time. A traveller starting further outwards at a radius $r > l_b$ will first go inwards radially until the traveller is on the same radial position as 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,

$$\begin{aligned}\tau_2 &= \int_{l_b}^\infty dr\, r \int_{-\theta_n/2}^{\theta_n/2} d\theta\, \tau(\mathbf{x},\mathbf{0})\, \rho(r) \\ &= \theta_n \int_{l_b}^\infty dr\, r\, \rho(r) \Bigg[r + \left(\frac{\pi}{2n} - 1 + a\right) l_b \\ &\quad + \theta_n b \int_0^{l_b} dr' \int_{r'}^\infty dr''\, r''\, \rho(r'') \Bigg].\end{aligned} \tag{B4}$$

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 finally yields

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

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

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

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 τ and the corresponding $l_b^\star = L/n^\star$ is then computed numerically.

Appendix C: Numerical solver for congested flow networks

In this appendix, we provide additional details on the numerical solver which computes the averaged travel time τ in the city for a given parameterised network geometry. It determines the optimal parameters for the given network geometry by scanning the parameter space.

The algorithm can be split into two parts: First, the city needs to be initialised. Second, for each point in the parameter space, the network needs to be initialised and routing takes place. In the following sections we will discuss these steps. Note, that this algorithm requires fully connected networks.

Step 1: Initialisation of the city

In our model, some structures of the city do not depend on the specific implementation of the transportation network, so that we only need to initialise them once in the beginning and reuse these structures throughout the scan in the second part.

Namely, these structures comprise the starting point of each traveller as well as his destination, which are initialised within the following steps:

1. *Draw starting points.* In a first step, we need to generate a population density distributed according to an arbitrary population density $\rho(x_1, x_2)$. We mimic the exact population density by placing N travellers in the plane using equally distributed random numbers. To obtain an accurate distribution of the population, we divide the plane into a grid of sufficiently small and equally sized cells and assume the population density to be constant within a single cell. Then, each cell is associated to a bin in a lookup table, where bin width is proportional to population density ρ evaluated at the centre of

the cell. Once the lookup table is set up, we draw for each traveller three equally distributed random numbers to determine his starting point: The first random number determines via the lookup table the cell, within which the starting point is located. The second and third number define the position within this cell.

2. *Draw destinations.* In a second step, the destinations are installed in the city. The number and distribution of them strongly depends on the underlying model for the city structure. We thus keep our solver flexible at this point to allow investigations for various models. We only assume that according to a model of choice, a set of D destinations is generated. Several travellers can be mapped to the same destination, so that D can be chosen independently from the number of travellers N.

   In this manuscript, we use a toy model with one destination in the city centre surrounded by N_c subcentres located isotropically at a distance R_c around the city centre. This fundamental model allows both studying a monocentric city ($N_c = 0$) and polycentric cities ($N_c, R_c > 0$). Note, that the solver easily can be extended to locate the destinations according to an arbitrary destination density $\rho_d(y_1, y_2|x_1, x_2)$ in the same way as it draws the starting points.

3. *Map starting points and destinations.* Once all starting points and destinations are located, it remains to map each traveller having an individual starting point to a destination. We assume, that the probability, that a traveller with starting point $\vec{x}$ is mapped to a destination located at $\vec{y}$, decreases with the Euclidean distance $||\cdot||$ as

$$\rho_d(\mathbf{x}|\mathbf{y}) \propto \frac{1}{||\mathbf{x}-\mathbf{y}||^\beta}\,. \tag{C1}$$

   This approach allows to study destinations mapping ranging from completely random mapping for $\beta = 0$ to a closest destination mapping for $\beta \to \infty$. The value of β might depend again on the underlying models and parameters of the city, such as the ability of inhabitants to choose their starting points. In this manuscript, we consider $\beta = 1$ corresponding to the established gravity model [46].

After performing these three steps, we thus have a set of travellers, each having a starting point and a destination. In Fig. 11, this situation is visualised for a city with an exponential population density $\rho(\mathbf{x}) \propto e^{-\|\mathbf{x}\|/r_0}$ and a city with $N_c = 6$ subcentres at a distance $R_c = 0.5\,r_0$ from the city centre, visualised as diamonds. Travellers are mapped to a destination using different values for β and visualised as dots coloured in the colour of its destination.

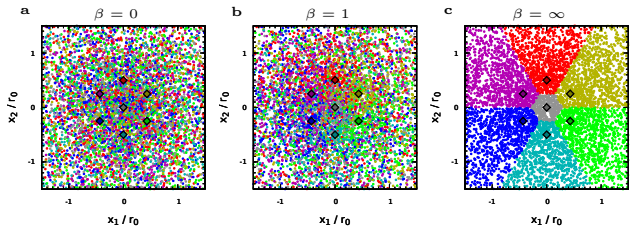

Figure 11. Comparison of destination mapping models. Each diamond corresponds to a destination. Travellers are visualised as dots in the colour of the destination they are mapped to. (a) For $\beta = 0$, travellers are randomly mapped to a destination, independent of their starting point. (b) Using a gravity model with $\beta = 1$, we observe an agglomeration of travellers close by a destination while still a significant amount of inhabitants need to travel to other parts of the city. (c) In the limit $\beta \to \infty$, we obtain extreme segregation where in each part of the city, all travellers are mapped to the closest destination.

Step 2: Initialisation of the network

Once the travellers are mapped to a starting point and a destination, we need to provide a set of possible travel modes that the traveller can chose from to commute between both points. In this manuscript, we assume the travellers only have the choice between:

1. *Walking.* Between any points in the plane, travellers can walk on any path with a constant velocity $v_w = 1$. Thus, the first possibility for a traveller to reach his destination is to walk along the straight line that connects his starting point and destination.

2. *Walking and using the transportation network.* Besides walking, there is a transport network, that allows travelling along the network edges at a higher velocity, that might differ between different edges and depend on the local flow due to congestion. Thus, the second possibility is to walk at speed v_w to an access point of the network, travelling inside the network to an other access point, and finally walking again the remaining distance to the destination.

The transportation network is modelled by a set of nodes which corresponds to stations where travellers can enter the subway. The nodes are distributed in the plane according to the given network geometry with a distance Δl to neighbouring stations along the network branches. In this paper, we set $\Delta l = 0.05 r_0$ which corresponds to a station distance of $\Delta l \approx 500\,\mathrm{m}$ when considering a city with $r_0 \approx 10\,\mathrm{km}$. Note that in some cases, the distance between neighbouring stations might differ from this value, e.g. to place the last station at the end of a branch. Each station s gets in addition to its position a list of inflow-destination pairs which count for each des-

tination the number of travellers entering the network at this station.

The edges of the network connect the nodes according to the given network geometry. For each edge of network (r, s) connecting two stations r and s we have a flow F_{rs} such that the travel time along this segment reads

$$\tau_{rs} = (a + bF_{rs})\, d_{rs} \tag{C2}$$

where d_{rs} is the distance of r and s along the network.

Step 3: Routing

Once we have initialised both the endpoints of each travellers path and the network, we need in the next step to find for each traveller the best route to travel to his destination. There are many different routing strategies that could be applied to define the best route for an individual traveller, depending on the travellers preferences, access to information and other parameters. Here, we consider three routing strategies as described in appendix A. Most numerical results are shown for the "lazy traveller" model, for which routing consists of two steps:

1. *Find closest stations.* For each starting point and for each destination, we determine the closest station in the network. Once we mapped each location to a closest station, we can compute for each station in the network the number of travellers that would enter the network here, as well as their target station, that is the closest station to their destination.

2. *Choose path inside the network..* Once the entry and exit station are known for each traveller, the shortest path within the network can be computed using a standard shortest-path algorithm.

Once the routing procedure is done, it remains to compute the number of travellers for every segment (r, s) of the network to obtain $F_{r,s}$ and $\tau_{r,s}$. Finally, we compute the travel time for each traveller, which is then averaged over all travellers to obtain τ.

[1] F. Corson, Fluctuations and Redundancy in Optimal Transport Networks, Physical Review Letters **104**, 048703 (2010).

[2] E. Katifori, G. J. Szöllősi, and M. O. Magnasco, Damage and Fluctuations Induce Loops in Optimal Transport Networks, Physical Review Letters **104**, 10.1103/PhysRevLett.104.048704 (2010).

[3] D. P. Bebber, J. Hynes, P. R. Darrah, L. Boddy, and M. D. Fricker, Biological solutions to transport network design, Proceedings of the Royal Society B: Biological Sciences **274**, 2307 (2007).

[4] A. Yazdani and P. Jeffrey, Complex network analysis of water distribution systems, Chaos: An Interdisciplinary Journal of Nonlinear Science **21**, 016111 (2011).

[5] G. Eiger, U. Shamir, and A. Ben-Tal, Optimal design of water distribution networks, Water Resources Research **30**, 2637 (1994).

[6] F. Kaiser, H. Ronellenfitsch, and D. Witthaut, Discontinuous transition to loop formation in optimal supply networks, Nature Communications **11**, 5796 (2020).

[7] M. T. Gastner and M. E. J. Newman, Optimal design of spatial distribution networks, Physical Review E **74**, 016117 (2006).

[8] M. Barthélemy and A. Flammini, Optimal traffic networks, Journal of Statistical Mechanics: Theory and Experiment **2006**, L07002 (2006).

[9] M. P. Viana, E. Strano, P. Bordin, and M. Barthelemy, The simplicity of planar networks, Scientific Reports **3**, 3495 (2013).

[10] M. Barthélemy, Spatial networks, Physics Reports **499**, 1 (2011).

[11] R. Louf and M. Barthelemy, How congestion shapes cities: from mobility patterns to scaling, Scientific Reports **4**, 5561 (2014).

[12] E. Strano, V. Nicosia, V. Latora, S. Porta, and M. Barthélemy, Elementary processes governing the evolution of road networks, Scientific Reports **2**, 296 (2012).

[13] C. Roth, S. M. Kang, M. Batty, and M. Barthelemy, A long-time limit for world subway networks, Journal of The Royal Society Interface **9**, 2540 (2012).

[14] S. Porta, P. Crucitti, and V. Latora, The network analysis of urban streets: A dual approach, Physica A: Statistical Mechanics and its Applications **369**, 853 (2006).

[15] S. Porta, P. Crucitti, and V. Latora, The Network Analysis of Urban Streets: A Primal Approach, Environment and Planning B: Planning and Design **33**, 705 (2006).

[16] P. Crucitti, V. Latora, and S. Porta, Centrality in networks of urban streets, Chaos: An Interdisciplinary Journal of Nonlinear Science **16**, 015113 (2006), publisher: American Institute of Physics.

[17] V. Verbavatz and M. Barthelemy, Critical factors for mitigating car traffic in cities, PLOS ONE **14**, e0219559 (2019), publisher: Public Library of Science.

[18] J. Fenger, Urban air quality, Atmospheric environment **33**, 4877 (1999).

[19] N. Geroliminis and C. F. Daganzo, Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings, Transportation Research Part B: Methodological **42**, 759 (2008).

[20] K. Nagel and M. Schreckenberg, A cellular automaton model for freeway traffic, Journal de physique I **2**, 2221 (1992).

[21] T. Afrin and N. Yodo, A survey of road traffic congestion measures towards a sustainable and resilient transportation system, Sustainability **12**, 4660 (2020).

[22] R. Prud'homme, M. Koning, L. Lenormand, and A. Fehr, Public transport congestion costs: The case of the paris subway, Transport Policy **21**, 101 (2012).

[23] O. Cats, J. West, and J. Eliasson, A dynamic stochastic model for evaluating congestion and crowding ef-

fects in transit systems, Transportation Research Part B: Methodological **89**, 43 (2016).
[24] A. De Palma, M. Kilani, and S. Proost, Discomfort in mass transit and its implication for scheduling and pricing, Transportation Research Part B: Methodological **71**, 1 (2015).
[25] R. Z. Farahani, E. Miandoabchi, W. Y. Szeto, and H. Rashidi, A review of urban transportation network design problems, European Journal of Operational Research **229**, 281 (2013).
[26] D. Aldous and M. Barthelemy, Optimal geometry of transportation networks, Phys. Rev. E **99**, 052303 (2019).
[27] M. Barthélemy and A. Flammini, Modeling Urban Street Patterns, Physical Review Letters **100**, 138702 (2008), publisher: American Physical Society.
[28] P. Wang, T. Hunter, A. M. Bayen, K. Schechtner, and M. C. González, Understanding Road Usage Patterns in Urban Areas, Scientific Reports **2**, 1001 (2012).
[29] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, Optimal Network Topologies for Local Search with Congestion, Physical Review Letters **89**, 248701 (2002).
[30] V. Cholvi, V. Laderas, L. López, and A. Fernández, Self-adapting network topologies in congested scenarios, Physical Review E **71**, 035103 (2005).
[31] OpenStreetMap Foundation, Openstreetmap, `https://www.openstreetmap.org` (2021).
[32] M. Barthelemy, *The Structure and Dynamics of Cities: Urban Data Analysis and Theoretical Modeling* (Cambridge University Press, 2016).
[33] M. Fujita and H. Ogawa, Multiple equilibria and structural transition of non-monocentric urban configurations, Regional Science and Urban Economics **12**, 161 (1982).
[34] P. Kemper and R. Schmenner, The density gradient for manufacturing industry, Journal of Urban Economics **1**, 410 (1974).
[35] D. P. McMillen and S. C. Smith, The number of subcenters in large urban areas, Journal of Urban Economics **53**, 321 (2003).
[36] V. Dökmeci and L. Berköz, Transformation of istanbul from a monocentric to a polycentric city, European Planning Studies **2**, 193 (1994), https://doi.org/10.1080/09654319408720259.
[37] R. Louf and M. Barthelemy, Modeling the polycentric transition of cities, Phys. Rev. Lett. **111**, 198702 (2013).
[38] C. Roth, S. M. Kang, M. Batty, and M. Barthélemy, Structure of Urban Movements: Polycentric Activity and Entangled Hierarchical Flows, PLOS ONE **6**, e15923 (2011).
[39] United States Bureau of Public Roads, *Traffic assignment manual for application with a large, high speed computer*, Vol. 2 (US Department of Commerce, 1964).
[40] A. Nagurney and Q. Qiang, Robustness of transportation networks subject to degradable links, EPL (Europhysics Letters) **80**, 68001 (2007).
[41] L. Haywood, M. Koning, and R. Prud'Homme, The economic cost of subway congestion: Estimates from paris, Economics of transportation **14**, 1 (2018).
[42] M. de Lapparent and M. Koning, Analyzing time sensitivity to discomfort in the paris subway: an interval data model approach, Transportation **43**, 913 (2016).
[43] S. Peeta and H. S. Mahmassani, System optimal and user equilibrium time-dependent traffic assignment in congested networks, Annals of Operations Research **60**, 81 (1995).
[44] J. N. Prashker and S. Bekhor, Route choice models used in the stochastic user equilibrium problem: a review, Transport reviews **24**, 437 (2004).
[45] A. Bertaud and S. Malpezzi, The spatial distribution of population in 48 world cities: Implications for economies in transition, Center for urban land economics research, University of Wisconsin **32**, 54 (2003).
[46] C.-i. Hua and F. Porell, A critical review of the development of the gravity model, International Regional Science Review **4**, 97 (1979).
[47] F. Zhang, H. Yang, and W. Liu, The downs–thomson paradox with responsive transit service, Transportation Research Part A: Policy and Practice **70**, 244 (2014).