UNIVERSITÁ DEGLI STUDI DI TRENTO
Physics Department

Doctor of Philosophy in Physics

Doctoral Thesis submitted to the Doctoral School of Physics - XXXVI Cycle

Optimal structures and collective dynamics of human flows in transportation networks

Supervisors:
Manlio De Domenico
Bruno Lepri

Candidate:
Sebastiano Bontorin

June 2024

I am grateful to my supervisors Manlio De Domenico and Bruno Lepri, for providing support and guidance through the PhD journey.
To my lab mates during these years, from CoMuNe, MobS and CHub, you have been above all great friends.
To the boys of Casa Veneto: Jack, Matte and Comi, in order of cooking skills (sorry Comi).

To my family. Papà, Mamma, grazie.
To zia Tiziana and Davidino, for your unyielding support.
To my sister Giovanna, for teaching me kindness.

0.1 Chapters Outline

1. **Introduction.** We introduce the context of the doctoral research. We first remark the relevance of the physics of complex systems applied to the study of cities and human mobility. We discuss the aim of this thesis, where we focus on the interplay between transportation network properties and collective human flows, from a theoretical perspective to a more data driven approach.

2. **Theoretical background.** We provide a minimal basis of shared literature and theoretical background on network science and statistical physics of human mobility for the following chapters. This chapter covers core concepts and metrics in network theory, spatial and transportation networks, statistical physics of human mobility, and urban systems from the perspective of complexity science.

3. **Multi-pathways network effective distances and prediction of human mobility on diffusive scales.** We study here the role of multiple pathways between origin-destination in the definition of effective network distance metrics [1]. We test these metrics in the task of prediction of human mobility on diffusive scales, with an application in a pipeline to assess the pandemic potential of a COVID-19 variant of concern [2].

4. **Optimal topologies in flow-weighted transportation networks.** Focusing on transportation networks, in this chapter we show how via optimization processes we can reconstruct different network topologies from simple constraints on a lattice planar substrate [3]. We study the role of traffic congestion, and the integration of spatial attractiveness in simulating human mobility patterns, illustrated through a case study on the Greater London Area urban structure.

5. **Mixing individual and collective behaviors to predict out-of-routine mobility.** In this chapter we focus on collective mobility dynamics, exploiting a large scale dataset of individual trajectories in three US cities [4]. We combine individual and collective information in human mobility models to improve next-location prediction in the context of novel mobility, including the impact of COVID-19 restrictions. Moreover, we investigate the statistical properties of collective mobility in proximity of attractive urban areas in predicting out-of-routine individual behaviors.

6. **Discussion.** We remark the core messages of this thesis, providing also a perspective on future extensions of the discussed works.

Chapter 1: Introduction

Many systems, both natural and man-made, can be seen as a complex system of interacting units [5]. The often non-linear nature of these interactions, paired to complex topologies of connections, leads to emergent functional properties of the system that can not be understood by only knowing the rules that govern single units. Biological systems, from proteins as networks of amino acids [6], gene regulatory networks [7, 8], up to entire food webs and ecosystems [9] represent in nature systems of networked units where their collective dynamics leads to system-level functional properties. In the context of man-made artificial or social systems, from infrastructure and transportation networks, the internet [10], up to networks of social interactions both online and physical, also represent paradigmatic examples of complex systems. Between the units, a plethora of different quantities are exchanged and can be modeled: fake news or consensus and endorsements in real and online social networks [11, 12], packages of data between servers, regulatory pathways in genes [6], electrical signals in neuronal networks [13], up to human beings travelling from one origin to destination in a transportation network [14]. This diverse and vast set of systems [15], spanning orders of magnitudes in spatial and temporal scales, can be treated under the same framework of network theory, non-linear dynamics and statistical physics. As a primary example of this is the mapping of these very diverse systems into universality classes in the way information propagates [16, 17].

In the present day, given also the access to large scale data of different phenomena, this system modeling has the validity and the usefulness for problems that concern our society [18, 19]. Among alia, the assessment of the robustness of human infrastructure systems [20] against failures which may originate from climate driven events is one paradigmatic example. Moreover, modeling the resulting cascade of effects that can spill into other layers such as economic and trade networks is a pressing necessity. A network science perspective is crucial to anticipate the impact of climate change on human social and infrastructure networks [20]. Similarly, the stability and recovering capacity of ecosystems can also be assessed by studying their response to external perturbations as networked systems [21]. Understanding their inherent chaotic and complex nature helps predicting and possibly avoiding systemic collapses. In other scales, network medicine [22] is another relevant field. Here for example the network perspective of activation or inhibition of specific genes in signaling pathways [6] represents a core modeling step

to improve drug design. In 2020 we have also seen that the world-scale spread of pathogens [2] represents a fundamental case where the system perspective is essential. Proactive measures can be designed with efficacy only when a broad view is taken into account, and simple epidemic models that fail to address this complexity are limited. Comprehensive views that embed together the different aspects of epidemic propagation represent fundamental tools for policy makers. Finally, the modeling of cities as adaptive and dynamic systems of interacting humans and infrastructure represents another critical topic for complexity-based studies.

The relevance of the study of human mobility and cities as complex systems.

The study of human mobility [23] and its interplay with spatial transportation networks [14], especially in the dimension of urban structures and cities [24], has seen in the last two decades relevant advances thanks to the powerful modeling tools coming from the physics of complex systems and network theory. Fundamental results such as the statistical properties of human trajectories [25, 23], the modeling of human migration [26] or the scaling rules and the hidden hierarchical structures behind complex cities [27], allowed novel understanding of these systems. Human mobility occurs across various scales [28], and is mediated by intrinsically multi-layer network structures spanning from air-transportation [29, 14] to complex structures at the urban level [30]. This intricate human dynamics significantly influences a set of diverse phenomena, among alia the design of sustainable and efficient transportation, the dynamics of social interactions, pathogen spread [2] and the prediction of individual movements [31, 32].

The design of sustainable and livable cities can follow principles and be guided by the system perspective [33]. As it is estimated that two thirds of human beings will live in urban areas by 2050 [34], building cities that are sustainable and livable is a pressing challenge. Among alia, considering the interplay of urban land use, multi-modal transportation networks and mobility patterns is essential in the design of public transit and active mobility infrastructure [35, 36] that can push for large cities to be less reliant on cars [37, 38]. The non-linear dependencies between traffic patterns and congestion [39], the morphological structure of cities [33] and the consequence on efficient transportation, segregation patterns and accessibility [40], are all intertwined dimensions where the insights provided by complexity science modeling are essential in informing policy making.

The interconnection between human mobility and the underlying transportation network [41] is pivotal, and when focusing on the urban scale it is intertwined with the spatial arrangement of land use and mobility hot spots [32, 42]. The notion of "structure-dynamics interplay" is exemplified in urban structures by the mobility flows, the traffic congestion dynamics and the multi-layer velocity structure of public transit [43]. These processes are all interdependent and ultimately interact in determining an effective temporal distance [41, 44] to travel from one part to another in a metropolitan city. The predictive modeling and understanding of this complex interplay, as well as the

understanding of the of structural elements and how these can emerge from simple rules [45, 46, 47], are foundational aspects of this doctoral research.

Aim: Explore the interplay of collective human mobility and transportation networks on multiple scales.

In this thesis we discuss these aspects starting from **(1)** a broad and more theoretical viewpoint, with the definition of effective distances on networked dynamical processes with a focus on human mobility seen as diffusive flows. Here we aim to show that an effective distance computed on the network of flows can predict the arrival times of a signal in reaction-diffusion dynamics, showing an application on a global epidemics scenario. Subsequently, we narrow the focus into the smaller scales of urban mobility **(2)**. We show how the optimization of an effective temporal distance on spatial networks, biased by traffic-like flows that mimic collective mobility, leads to the emergence of complex topologies characteristic of urban systems. Our aim here is to uncover the minimal conditions that lead to the emergence of optimal features on a planar substrate. Finally **(3)** we aim to exploit collective mobility behaviors to enhance predictability of individual trajectories in the context of next-location prediction. Here we adopt a data driven approach using a large scale mobility dataset of three United States cities. We show that collective mobility properties are intimately related to the urban space and aid the prediction of individual mobility. We describe these 3 specific objectives more in detail in the following paragraphs and in Figure 1.1.

(1) Knowledge of multiple pathways leads to prediction of arrival times in diffusive dynamics mediated by collective mobility flows. In Chapter 3 we start with the discussion of effective distances for networked dynamical systems, and explore how multiple pathways determine the estimation of effective distances. Effective distances on networks are fundamental as they capture and synthesize the interplay between nodes' interactions and the underlying structure. Their knowledge is essential to predict signal propagation or embed the network in a latent geometry. We introduce and discuss how a multi-pathways temporal distance [1] naturally encodes the concerted behavior of the ensemble of paths connecting two nodes in conveying perturbations, with applications ranging from protein-protein dynamics in the proteome to epidemic spreading in social networks. In the specific case for diffusive dynamics we introduce the Information Distance metric exploiting the Laplacian matrix. As collective human mobility flows on large scales such as the World Airtransportation Network can also be represented as proxies for diffusion processes, they represent a key modeling step in reaction-diffusion systems such as epidemics models with meta-population nodes connected by air transportation. We discuss how the Information Distance is employed to compute the effective probability for infectious seeds to reach a target airport using data from cross-national human mobility flows. The Information Distance metric with its analytical closed form allows for fast estimation of effective distance while embedding information of multiple pathways in the application of the Laplacian operator. We also use it as a benchmark against other metrics as a part in a general pipeline [2]

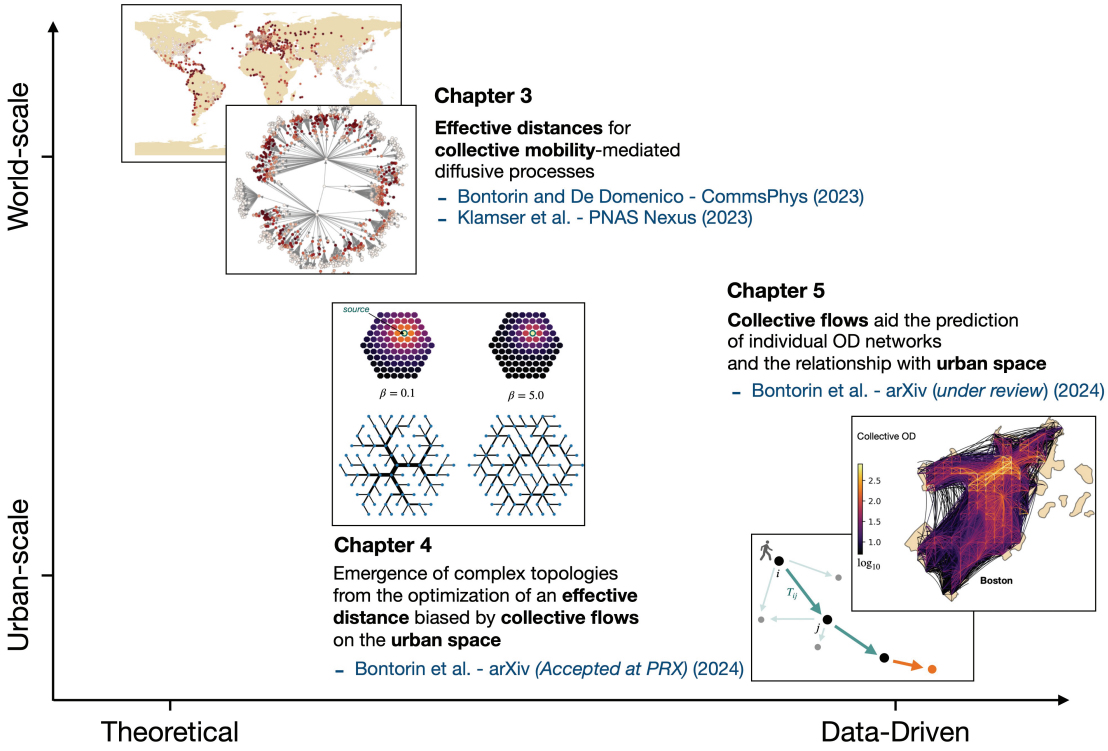

Figure 1.1: Thesis structure. The three main chapters of this thesis are plotted on a bi-dimensional space based on the extent of analytical and theoretical insight and the reliance on data used to generate networks and mobility patterns. Keywords that also highlight connections between chapters are mapped in bold. The reference papers on which the chapters are based are mentioned. Chapter 3 is based on the works we published in Ref. [1] and in Ref. [2]. Chapter 4 is based on our work in Ref. [3] and finally Chapter 5 is based on Ref. [4].

where national genomic surveillance was integrated with global mobility and large scale epidemics modeling to quantify the pandemic potential of a variant of concern in the context of the SARS-CoV-2 pandemic.

(2) Optimization of a temporal distance weighted by collective human mobility leads to emergence of complex spatial topologies. In Chapter 4 we then focus into the modeling of spatial transportation networks, which represent the arteries for human movements, especially on urban scale. We aim to understand the network topologies that emerge when a substrate network is optimized in different mobility scenarios. Transportation and distribution networks are often characterized by the presence of complex structures such as central loops paired with peripheral branches, which can appear both in natural and man-made systems, such as subway and railway networks. In this study, we investigated the conditions for the emergence of these non-trivial topological structures [3] in the context of human transportation in cities. We propose a minimal model for planar networks generation, where a network lattice

acts as a spatial planar substrate and edge velocities and distances define an effective temporal distance which quantifies the efficiency in exploring the urban space. To this aim we implemented a simulated annealing algorithm to obtain the optimal network configurations of velocities to guarantee the minimum effective distance given the mobility flows. We study how the interplay between a flow probability between two nodes in space and the associated travel cost influences the resulting optimal network. In the perspective of urban transportation we simulate these flows by means of human mobility models to obtain Origin-Destination matrices. We find that when using simple lattices, the obtained optimal topologies transition from tree-like structures to more regular networks, depending on the spatial range of flows. Remarkably, we find that branches paired to large loops structures appear as optimal structures when the network is optimized for an interplay between heterogeneous mobility patterns of small range travels and longer range ones typical of commuting. Furthermore we show that our framework is able to recover the statistical spatial properties of the Greater London Area subway network. We extended the analysis by also considering the effect of congestion and user-equilibrium principles in traffic routing. We derived analytical conditions for a tree structure to transition to one where multiple alternative paths emerge to mitigate the temporal delays introduced by congestion.

(3) Exploit collective dynamics and its dependence on the urban space to improve predictive capabilities of individual mobility models. Finally in Chapter 5 we focus on predictive models of individual and collective human mobility in urban space, adopting a more data-driven approach. Performing next-location prediction at the individual level is a challenging task, which is critical for numerous research and practical applications, such as urban planning, public health and sociology. Intrinsic properties of mobility datasets limit the predictability of individual-level human mobility, such as the visitation frequency to places following a long-tail distribution. This leads to a core problem in predicting locations not seen in training, a challenge where even state-of-the-art deep recurrent models fail. In this chapter we propose to support next-location predictors via collective origin-destination matrices [4]. We study minimal Markov models and exploit a normalized Shannon's entropy as an effective definition of individual movement's unpredictability seen in training. We demonstrate the effectiveness of this approach by using a large-scale dataset that contains trajectories of more than 2 million users collected over eight months in 2020 in New York City, Seattle and Boston provided by Cuebiq. Our results show that this approach improves the accuracy of next-location prediction compared to existing state-of-the-art methods. Moreover, we shed light on the features of collective mobility linked to the urban environment: out-of-routine movements in urban areas in proximity of POIs are better predicted by collective behaviors. Furthermore, we investigate the changes in human mobility behavior and routines during the COVID-19 pandemic.

Chapter 2:
Theoretical background

2.0.1 Chapter Overview

In this chapter we provide a minimal theoretical background and concepts from the literature on complexity science applied to spatial networks, human mobility and urban systems. The aim is to delineate the common literature and theoretical framework on which the following chapters will be discussed and developed.

We begin introducing basic elements of network science in Section 2.2 and providing some of the network observables and operators that will be used throughout the thesis. We then discuss in Section 2.3 the study of spatial networks and of optimal spatial topologies, of which transportation networks represent a peculiar case, especially for human mobility and in cities. We then discuss some of the fundamental findings in the study of human mobility and its statistical properties in Section 2.4. In this section we also discuss some fundamental models that will be adopted in Chapters 4 and 5. Finally in Section 2.5 we discuss the modeling of cities as complex systems, introducing a minimal literature background to Chapters 4 and 5 and providing a general perspective on why the study of urban systems represents an important challenge.

2.1 Physics of complex systems and Network Science

Complex systems can be described as collections of interacting elements. Here, the dynamics of individual elements may be well-understood in isolation, but this knowledge is insufficient to grasp the system's dynamics as a whole [5]. In more formal terms, the laws governing complex systems cannot be derived as a linear combination of laws governing each individual element as an isolated entity. The interactions between these elements are fundamental. The final system's behavior arises from the interplay between the non-linear dynamics [16] that govern how a single unit's state depends on its neighbors, and the overall structure of these connections.

A paradigmatic example in the physics of complex systems is represented by spin glasses [48]. These are magnetic materials that exhibit complex emergent features from quenched disorder properties to frustrated dynamical states. These general concepts are fundamental to understand the emergence of complex behavior in many physical systems [49, 50]. In the last decades these modeling tools have been extended for example in larger scales such as in biological systems, from the complex dynamics of flocks [50, 51], the study protein-protein physical interactions [22], the non-linear dynamics of neuronal networks [52], or to map the functional hierarchical structure of gene regulatory networks in cells [53]. In system biology, the network-science perspective can be applied from the study of transport properties of fungal networks [54] up to the multi-layered structure of ecosystems [15, 55]. Within these ecosystems, mathematical tools from dynamical systems can model complex behaviors such as predator-prey dynamics, and even quantify how it can affect other ecosystems' layers [56] such as the presence of certain plant species within the layer of vegetation species. Human-made systems like infrastructure networks and the internet [57], with its routing protocols, also can be better understood via network science. Even social interactions, online and offline, form networks that drive information flow and reveal communication patterns [12]. Furthermore, the rich variety (and heterogeneity) of interactions [58] within these networks allows for the propagation of diverse signals and information exchange between the systems' units. Despite the remarkable disparity in scale and temporal dynamics across these systems, the framework based on graph theory and complexity science represents a powerful tool to model these intricate connections [16, 17].

While not all complex systems can be described using the networks formalism, in many the network science perspective allows to uncover fundamental complex properties [57]. Real world systems are characterized by connectivity patterns that are heterogeneous and fat-tailed [58]. The presence of mesoscale communities [59, 57], hierarchical connectivity patterns and complex information propagation patterns [17] represent some of the systems' functional properties that can be uncovered via network science. Moreover, from percolation theory [20] up to approaches based on the definition of a network entropy [60] also constitute sophisticated tools to assess the robustness and functionality of these systems.

2.2 Mathematics of Networks: Core concepts and metrics

A network [57] can be defined using a graph $G(N, E)$, where N is the number of nodes, entities that can be seen as points and represent the units of the system. Edges $e = (i, j) \in E$ define the pair-wise connections (lines) between nodes. This information is mapped to the adjacency matrix A, which in the case of a weighted graph:

$$A_{ij} = \begin{cases} w_{ij} & \text{if } (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

Where in the case of a binary network $w_{ij} = 1$, or in a weighted network is a scalar value such as the number of travellers between stations i and j. If $w_{ij} \neq w_{ji}$ we refer to the resulting graph as a directed network. From the knowledge of A_{ij} the connectivity properties can be computed.

Connectivity properties. The simplest connectivity feature of a graph is the number of connections of a node, the degree k. For an undirected network:

$$k_i = \sum_{j=1}^{N} A_{ij}. \tag{2.2}$$

In a directed network the degree can be separated between *in-degree* and *out-degree*. In mobility matrices, these can for example represent out-flows or in-flows of travellers from or to a station [23]. Nodes having orders of magnitude more connections than the average $\langle k \rangle$ are defined as "hubs", where $k_{hub} >> \langle k \rangle$.
When all connections are present: $A_{ij} \neq 0 \ \ \forall (i,j)$ a network is said to be complete and $\dim E = \binom{N}{2}$. G is instead a connected network only if a path exists between any pair of nodes. Networks can also be sparse [60], resulting in a low density of connections, e.g. when $\dim E \sim N^{\alpha}$ with $\alpha < 2$. In the statistical approach to complex networks, the definition of the probability that a node has degree k, $P(k)$, allows to characterize asymptotic regularities such as networks having similar distributions $P(k)$. Different generative random networks models can for example be classified by their $P(k)$. Finally, real networks in nature often exhibit sparse connections and a power law scaling $p(k) \sim k^{-\gamma}$ which is related to the presence of hubs, although this represents a debated observation [61].

The Graph Laplacian. From the knowledge of A_{ij} and the set of nodes' degrees $\{k\}$, matrices such the Laplacian can be built as $L = D - A$. Where D is the diagonal matrix of degrees $D_{ij} = k_i \delta_{ij}$, and therefore has entries:

$$L_{ij} = \begin{cases} k_i & \text{if i = j} \\ -A_{ij} & \text{if i} \neq \text{j} \\ 0 & \text{otherwise} \end{cases} . \tag{2.3}$$

L is a positive semi-definite matrix [11] (with eigenvalues $\lambda_i \geq 0$, $\forall i$). The terminology stems from the analogy with the heat diffusion in a continuous medium, where the Laplacian operator (∇^2) acts as a diffusion operator on the system state: $\partial_t u(x,t) = \partial_x^2 u(x,t)$. It can be shown that by mapping a diffusion-like equation on a network, where the rate of change dx_i/dt for each node is driven by the difference of x_i with the neighbors x_j, leads to the formulation with the Laplacian graph operator:

$$\frac{dx_i}{dt} = -\sum_j A_{ij}\,(x_i - x_j) = -\sum_j L_{ij} x_j. \tag{2.4}$$

Moreover solutions to 2.4 can be decomposed as linear combinations $x_i = \sum_{i=1}^{N} c_i(t)\mathbf{v_i}$, where $c_i(t) = c_i(0)e^{-\lambda t}$ and $\{\lambda_i, \mathbf{v_i}\}$ represents the eigenspectrum of L. Therefore the diffusion dynamics will be governed by the smallest eigenvalues of L, and the first non-vanishing eigenvalue λ_1 determines the rate of speed for these networked systems to relax to equilibrium. Other features, such as mapping the network to low dimensional embedding [57], the definition of density matrices [60], and other dynamical features can be derived from the Laplacian. Overall, this shows how from the knowledge of A_{ij} and $\{k\}$, which represents information often directly encoded in the data, powerful operators such as the Laplacian can be computed. And from L, even mesoscale properties of a network can be inferred [59]. Although we mention that the inference of the network connections A_{ij} from data in many context represents instead a non-trivial task [62].

A normalization of the Laplacian matrix as $L^{RW} = D^{-1}L$ allows to study random walks on networks and other processes such as synchronization dynamics [63]. A random walk can be introduced as a discrete-time dynamics where the probability of walker to be in node i at time t is:

$$p_i(t+1) = \sum_j P_{ij} p_j(t), \tag{2.5}$$

where P_{ij} is the transition matrix for the discrete-time Markovian process. The probability of jumping to one of the neighbors is $P_{ij} = A_{ij}/k_i$. The continuous-time mapping of Eq. 2.5 is therefore:

$$\dot{p}_i = \sum_j (P_{ij} - \delta_{ij}) p_j = -\sum_j L^{RW}_{ij} p_j. \tag{2.6}$$

Ultimately this leads to the master equation: $\dot{\mathbf{p}}(t) = -\mathbf{p}(t)L^{RW}$ with solution $\mathbf{p}(t) = \mathbf{p}(0)e^{-tL^{RW}}$. This can also be generalized to consensus dynamics and synchronization near a stable state [63, 59]. L^{RW} will be employed in Chapter 3 to introduce the Information Distance and apply it to predict the arrival time of infectious seeds diffusing on the routes of the World Airtransportation Network [29], further highlighting the relevance of Laplacian in capturing the fundamental dynamics of and on networked systems.

Paths. Sequences of connected edges that connect an origin node i to another destination node j define a path Π_{ij}. A path of length L, $\Pi(L)$ can be also defined via the nodes it traverses:

$$\Pi(L) = \{n_1, \ldots, n_L\}, \tag{2.7}$$

or the set of edges $(n_i, n_{i-1}) \in E$ that connect these nodes. The shortest path length, L_{ij} between nodes i and j is the minimum number of edges that need to be traversed to reach node i starting from node j. From knowledge of L_{ij}, $\forall i, j$, we can also define global metrics such as the average path length:

$$\langle L \rangle = \frac{1}{N(N-1)} \sum_{i,j} L_{ij}. \tag{2.8}$$

Albeit simple, metrics like $\langle L \rangle$ quantify the ability of a network to propagate in few steps information between pairs of nodes.

Network effective distances. From the knowledge of the set of paths $\mathbf{P}(i \to j)$ that connect two nodes $i-j$, sophisticated metrics such as effective distances can be computed, related to the general topic of network geometry [64, 65]. Here we briefly mention the one introduced in Ref. [29] as it also represents a precursor of the metrics presented in Chapter 3. The effective distance devised by Brockmann and Helbing [29] for diffusive systems can be computed directly from the mobility network of air transportation where $A_{ij} = F_{ij}$ and F_{ij} is the number of passenger on routes between airports i and j. They devised a metric built on the definition of an edge probabilistic distance as:

$$d_{eff}(n, m) = d_0 - \log P_{nm}, \tag{2.9}$$

where $P_{nm} = F_{nm} / \sum_m F_{nm}$ and d_0 is a constant. From this, the effective distance of a path can be computed as $D_{eff}(\Pi) = \sum_{e \in \Pi_{ij}} d_{eff}(e)$ and finally the effective distance between two nodes will be computed on the path $\Pi \in \mathbf{P}(i \to j)$ having the smallest D_{eff}. In Chapter 3 we discuss more in details these effective distances. While in Chapter 4 we work with a definition of effective distance based on the travel velocity of an edge w_{ij}, and the final effective temporal distance will result from the path Π having the fastest travel time [41, 44].

Generative Models: ER and B-A. Generative models of random networks can be devised to reproduce different $P(k)$ with the aim of providing benchmarks with which real networks can be compared. One of the simplest generative model is the Erdos-Rényi (ER) network, in which given a set of N nodes, an ensemble of networks can be generated by sampling an edge between each nodes' pair (i, j) with probability ρ_{ER} [66]. Hence a degree k has an associated probability of appearing which is equal to the probability of a node having k connections and $N - 1 - k$ absent connections. $P(k)$ is therefore a Binomial distribution:

$$P(k) = \binom{N-1}{k} \rho_{ER}^{k} (1-\rho_{ER})^{N-1-k}, \tag{2.10}$$

which, in the thermodynamic limit of the number of particles/nodes $N \to \infty$ and by taking $\rho_{ER} \cdot N = \langle k \rangle$ it can be approximated as a Poisson process: $P(k) = \mathrm{e}^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$. Thus the presence of edges in a Erdos-Rényi graph is associated to a random white noise process. If instead edges are not drawn independently, but rather following rules based on other nodes' features, different $P(k)$ can appear. A paradigmatic example of these rules is the preferential attachment process in the Barabasi-Albert [58] model for generating networks with a heavy-tailed $P(k)$ distribution. These models allow to capture the fundamental mechanisms that generate networks with heavy-tailed degree distributions often seen in real systems.

The different properties of $P(k)$ distributions then reflect in actual differences in functional properties of the networks. For example, by recalling the definition of $\langle L \rangle$, these different generative models allow to estimate how $\langle L \rangle$ scales with network size based on $P(k)$. For example, in a lattice of dimension D, $\langle L \rangle \propto N^{1/D}$, while it scales $\propto \log N$ in ER networks. Finally in scale-free networks the presence of hubs guarantees that $\langle L \rangle$ remains low even when the size (number of nodes) of the network N grows. This is quantified by $\langle L \rangle \propto \log \log N$.
In Chapter 3 we rely on these models to perform simulations and testing different effective distances. In Chapter 4 we also employ the ER model as a substrate for the study of optimal topologies as a simple case of non-spatial network.

2.3 Spatial and Transportation Networks

Spatial networks [14] are a class of graphs where nodes are embedded in a D-dimensional metric space and can be located with a set of D coordinates. Many real spatial networks, such as road, railways, or fungal networks [54], have edges that are supported by the 2 dimensional plane and do not intersect. When edges can not intersect or overlap due to physical constraints we have a *planar* network. In spatial networks, several topological features are influenced by the inherent cost of building long-distance links [67]. Long connections are favored only if the payoff is significant, such as connecting to a hub that favors communication [67]. These spatial constraints limit the number of connections, therefore leading to a degree distribution $P(k)$ that peaks at a certain number of connections and is thus characterised by a specific scale [14]. In planar networks these constraints are even more relevant and have an effect on how topological properties such as $\langle L \rangle$ scale. In 2D planar networks, it scales similarly to regular lattices ($\langle L \rangle \propto \sqrt{N}$). Although in non-planar ones, the presence of long distance links introduces topological short-cuts and the emergence of small-world properties, and $\langle L \rangle$ scales with a logarithmic dependency.

Most of the transportation networks present in nature exhibit planar characteristics, existing on a two-dimensional plane (despite eventual elevation changes). River networks and leaves' venation [68], for instance, are characterized by hierarchical branching patterns with minimal intersections, creating a primarily unidirectional flow. Another layer of complexity can be added when also accounting for capacity variations due to river width and depth. Fungal networks [54], on the other hand, are characterized by a web-like structure with bi-directional flow through mycelium threads. Similar features can be found also in human made transportation networks [30, 43]. In Figure 2.1 we can see the three layers for the urban transit system of London [43].

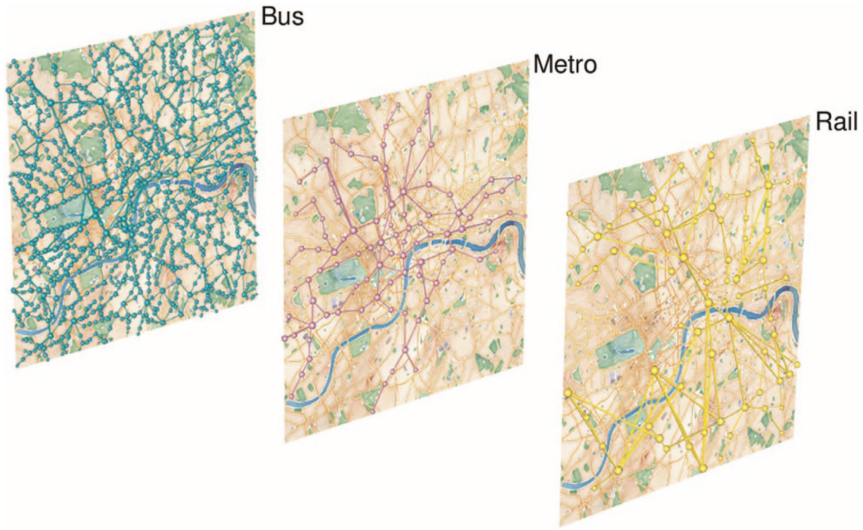

Figure 2.1: Multi-layer transit network of London. The sub graphs of transit networks in London for the bus, metro and rail layers. We can visually appreciate the lattice-like topology of the bus network, while the metro exhibits a structure characterized by a connected core paired to peripheral mono-dimensional branches. Figure taken from [43].

The development of models to recover these features allow to understand and optimize the dynamics on and of these spatial networks. As an example, in Chapter 4 we study how to arrange the edges' velocities on a planar substrate such that the travellers flow is optimized [3].

2.3.1 Optimal networks

Different network features can be generated by imposing optimality constraints on spatial embedded networks [69, 67]. For example, the hub-and-spokes structure represent a key topological structure in optimal distribution networks, where long-distance connections morph into high-capacity arteries, primarily interconnecting regional hubs, which in turn distribute traffic within their local neighbor nodes [67]. Other structural properties can be understood via optimization processes, such as fluctuations and resilience to disruptions within optimal networks, which naturally lead to the formation of loops [68]. While a

plethora of models to generate optimal networks exist, we mention here a relevant one that also serves to introduce the work presented in Chapter 4.

Distance versus topology optimization. A simple model was studied in Ref. [69] in the case of non-planar spatial networks, where different topologies emerge when the links are drawn to either have the cumulative shortest distance or building a topology that minimizes the betweenness centrality [57]. Given a set of N nodes in space, the objective is to find the tree connecting them that minimizes the quantity:

$$E = \sum_{e \in G} b_e^{\mu} d_e^{\nu}. \tag{2.11}$$

Where μ and ν are exponents that regulate the role of betweenness (topological measure) or the role of distance (euclidean metric measure). Samples of the resulting topologies are shown in Fig. 3.1 reproduced from [69].

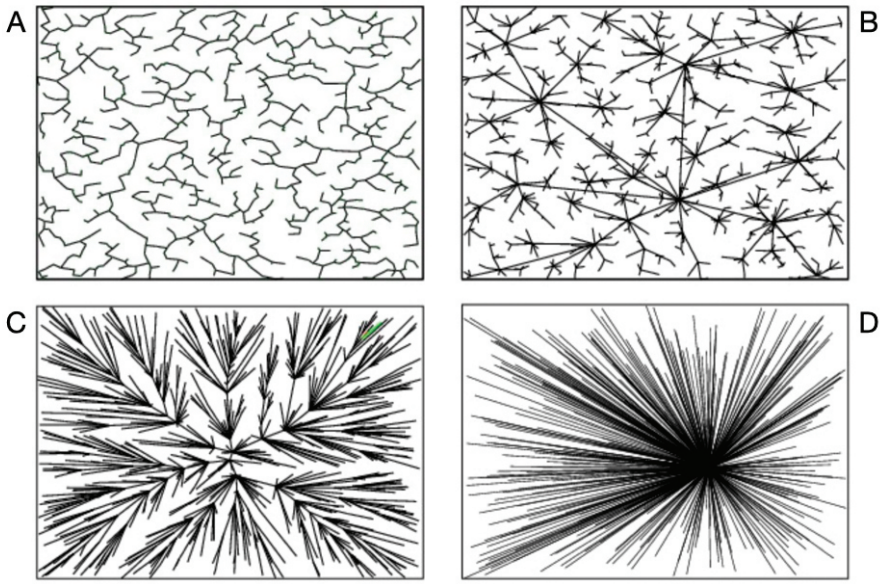

Figure 2.2: Resulting topologies. Different optimal network from the optimization of Eq. 2.11. **(a-d)** show topologies when (μ, ν) values are set to (0,1),(1/2,1/2),(1,1),(1,0) respectively. The different topologies therefore transition from the Minimum Spanning Tree MST (0,1), where the total distance is minimized, to a star-like topology, the Shortest Path Tree SPT (1,0). The figure is taken from [69].

Interestingly, the in-between scenario with ($\mu = 1/2$,$\nu = 1/2$) is a non-planar spatial optimal tree that is characterized by the presence of hubs connecting different areas, with smaller degrees nodes connected to them. Remarkably, it exhibits a hierarchical structure, recovering a mesoscale property typical of natural systems. This hierarchical

topology is therefore the trade off between the efficient Shortest Path Tree (SPT, with high cumulative edges distances) and the Minimum Spanning Tree (MST). In Chapter 4 we adopt an approach where we also observe a transition from tree-like structures (similar to the one obtained in these model for $(\nu = 1, \mu = 1)$) to lattice-like with short range links $(\mu = 0, \nu = 1)$. Although we work on planar triangular lattices, with the aim of optimizing edges' velocities biased by traffic flows.

We also mention edges' capacity-based optimal models [70], used to study for example in botanics the evolution of leaves [68]. Highlighting that the resilience to damage also naturally induces a high density of loops.

2.4 Statistical physics of human mobility

In recent years, the large availability of data of human mobility, mostly from GPS trajectories [23], allowed to uncover regular patterns and develop sophisticated mobility models for both individuals [71] and groups of people [25, 72, 26]. These in turn have become crucial for various applications, including traffic prediction [41, 39] and congestion mitigation [39], urban planning, and even to understand segregation dynamics in cities [40].
To this aim, the analytical framework provided by network science and statistical physics allows to model and capture the complex spatio-temporal inter dependencies that shape these patterns [73]. For instance, while agent-based models (ABM) treat individuals as autonomous agents programmed with rules for destination choice and navigation, statistical models identify universal patterns [32] in movement data and derive mobility laws [25, 72] that can be generalized and exploited in many use cases. Moreover, network-based models represent the spatial physical environment (at multiple scales) as a network of flows between locations [74]. For example, in urban planning these insights can be exploited to design cities that facilitate the efficient flow of people or commodities.

2.4.1 Observations

From empirical individual trajectories several observations can be derived [23], such as travel time and distance distributions, mean square displacements, the radius of gyration, energy-related metrics, up to origin-destination matrices. Most of these are metrics derived from the study of the physics of multi-particle systems. We briefly discuss the ones that will be also employed in the following chapters. We start by introducing the concept of distribution of travel distances r, fundamental develop individual mobility models based on diffusion dynamics.

Distribution of travel lengths. The distribution $P(r)$ [23] of distances r travelled by individuals represents the likelihood of an individual to travel a specific distance r. To estimate $P(r)$ a set of very diverse sources have been used, from the tracking of bank notes, mobile phone records, to now commonly used GPS spatio-temporal points. While

there is a general consensus for the power-law distribution property of $P(r)$, studies have found that actually meaningful scales in human mobility exist [75] and the power law behavior might result from an aggregation of these jump length log-normal distributions. More specifically, fits from the tracking of trajectories of dollar bills $P(r)$ was shown to be a power law $P(r) \approx r^{-(1+\beta)}$ with $\beta = 0.59 \pm 0.02$ [76], while more in general for mobile phone data, $P(r)$ can be fitted with a truncated power law [23]:

$$P(\Delta r) = (\Delta r_0 - \Delta r)^{-(1+\beta)} \cdot e^{-\Delta r/\kappa}. \tag{2.12}$$

We will adopt a power-law fit in Chapter 5 to fit the different $P(r)$ of human mobility in different regions in the cities under study.

Other mobility regularities and the network-based perspective. Other fundamental statistical regularities in human movements have been uncovered in recent years. For example, we mention the paradigm of "returners" versus "explorers" [77, 32], two classes of individuals identified from GPS data. Returners' mobility is characterized by repeated visits to a limited number of preferred locations. Their $P(r)$ and radius of gyration are stable in time, while explorers tend to visit diverse locations, with an expanding radius of gyration. Furthermore, the study of mobility networks from individual trajectories identified only 17 trip sub-networks [78] structures that can ideally describe all daily human mobility patterns. Therefore even if each individual visits a different set of locations, the same set of mobility motifs are common in our everyday movements, further highlighting the relevance of the network-based modeling. Finally, these individual networks can be aggregated in population-level Origin-Destination (OD) matrices, where entries encode the flow of individuals from one origin location to another [23]. We discuss in the followings different physics-inspired models to reproduce these ODs.

2.4.2 Modeling Individual and Collective mobility

Now we briefly discuss some exemplary models devised to either reproduce individual mobility or collective mobility patterns previously introduced.

Individual models, the EPR and other paradigms. Individual movement exhibits inherent stochasticity due to free will and randomness in trip patterns. Initially this lead to minimal models based on diffusive dynamics, such as the Continuous Time Random Walk (CTRW) [23] where also the number of movements between location in a time δt is also a stochastic variable. The ability to model travelers via diffusion-based models will be something we will discuss also in Chapter 3. Significant regularities in individual trajectories highlighted by recent research allows for the construction of more sophisticated generative models for individual movement. For example, visitation patterns in human individual trajectories follow a rank-frequency Zipf's law: ($f_k \sim k^{-\xi}$) [23], in contrast with the set of locations a random walker would visit, gradually drifting

from the origin. More sophisticated mechanisms have therefore been added to modify the CTRW, such as the Exploration and Preferential Return (EPR) [79] where the tendency of individuals to visit a set of preferred locations daily (preferential return) is mixed with the complementary probability of exploring novel locations $P_{new} = \rho S^{-\gamma}$. Where ρ, γ are model parameters and S increases at each new visit of a location. Recent research also highlighted modification of the EPR model including a preferential direction [32]. We will discuss in Chapter 5 the role of collective behaviors in predicting instead novel individual mobility patterns in the task of next-location prediction.

Collective Models: from the the gravity to the radiation model. Individual mobility data can be aggregated to generate population-level flows across spatial and temporal scales. This allows understanding from migration patterns between nations to commuting at the urban scale. In Chapter 3, we will develop an effective distance [64] for a diffusive process mediated by the OD network of collective mobility flows between airports. Here we discuss more in detail the gravity-like models [23], with a focus on the model for spatial interaction [74] at the urban scale, which we will then use in Chapter 4.

Zipf in 1946 introduced a model to estimate migration flows between communities. It was based on the assumption that locations with large population are both sources and targets (high attractiveness) of large flows of people, while greater distance represents instead a deterrent factor in migration. This lead to the first gravity-like model $T_{ij} \propto (P_i P_j)/r_{ij}$. It can also be generalized to consider other variables (m_j) such as the GDP of destination j. Other distance-based deterrence functions can be considered, leading to a more general form:

$$T_{ij} = k m_i m_j f(r_{ij}). \tag{2.13}$$

While widely adopted, it shows some limitations such as the estimation of fitted parameters, sensitive to data quality. Later, other non-parametric models such as the radiation model [26] were introduced to better fit migration patterns and models based on deep neural networks are also studied to instead learn from data these dependencies [80]. We discuss now the development of a gravity-like model from entropy-based principles, for human movements that will be then thoroughly used in Chapter 4.

Gravity model at the urban scale: the spatial interaction model. Alan Wilson [42, 74] provided a derivation within the framework of classical transportation theory and entropy maximization of the functional form of the gravity model in Eq. 2.13. The argument was based on the concept that in absence of more information, the most probable origin-destination (OD) matrix (denoted by $\{T_{ij}\}$), is the one that maximizes the number potential configurations of trips associated with it, all while respecting a set of constraints. First, $\Omega(\{T_{ij}\})$ represents the total number of distinct individual arrangements (configurations) that can generate the flow set $\{T_{ij}\}$ as:

$$\Omega(\{T_{ij}\}) = \frac{(\sum_{ij} T_{ij})!}{\prod_{ij} T_{ij}!}. \tag{2.14}$$

They then tried to find an equation for T_{ij} to maximize $\Omega(\{T_{ij}\})$ by also constraining the outflow $O_i = \sum_j T_{ij}$ to a specific value (an empirical observed value such as the population of a tile i), and similarly the destination inflow and the total travelling cost $C = \sum_{ij} T_{ij} C_{ij}$ can be constrained. By using Lagrange multipliers, in the limit of large number of trips $T = \sum_{ij} T_{ij}$ this leads to a doubly-constrained gravity model: $K_i O_i L_j D_j e^{-\beta C_{ij}}$ [23, 74]. Interestingly, the power law decays $f(r) \propto r^{-\beta}$ can instead be obtained by setting costs with a logarithmic dependence: $C_{ij} = ln(r_{ij})$. Thus recovering the gravity-like form from an a-priori statistical approach. Following this process, the spatial interaction model was devised by Wilson [74] by also determining that destination locations have an attractiveness W_j that can be quantified using for example the commercial land use area in node j, resulting in:

$$T_{ij} = K P_i^{\alpha_1} W_j^{\alpha_2} \cdot e^{-\beta c_{ij}}. \tag{2.15}$$

Where α_2, α_1, β are fit parameters. Albeit simple, the general approach of gravity-like modeling, using the number of Points of Interest (POIs) as attractiveness metric, was also found to be the closed form modeling that better represents OD data in urban contexts when evaluated via a Bayesian setting [81]. Moreover, it has shown good results even when compared to deep learning baselines, further corroborating the validity of this closed form in generating realistic collective flow patterns in a city. This highlights the intricate relationship between a city's spatial characteristics and the dynamical processes happening on it, a hallmark of cities as complex systems. In the next section we briefly discuss the study of cities from the perspective of complexity science.

2.5 Cities as complex systems.

Cities are intricate socio-economic, cultural, and spatial entities wherein the interactions between myriad agents—ranging from individuals to institutions—give rise to emergent phenomena [24]. Cities can be viewed as complex systems with an intricate interdependence of several layers [18]. Land-use patterns exhibit spatial organizations, which then represent functional areas used by the human population [24]. Analogous to biological information within a circulatory system, humans navigate the transportation network connecting these functional zones. Furthermore, information exchange and social interactions within these zones generate emergent social phenomena and result in socio-economic outputs of the system. Interestingly, these systems exhibit scaling laws [45] reminiscent of metabolic laws observed in biological organisms. Therefore, the city-as-organism analogy, while metaphorical, may hold some validity at specific scales. Among alia, examples of emergent collective effects include the clustering of functional areas [82], social segregation phenomena [40], congestion dynamics in traffic [39], and the diffusion of ideas [45]. These collective phenomena also culminate in allometric scaling properties with the city size [45], further establishing the interdependence of all the different layers that constitute a city.

2.5.1 Scaling properties of cities

Cities exhibit scaling properties between different indicators and their size. The first example is Zipf's law $R_i \sim N_i^{-\gamma}$ (with $\gamma = 1.019$), which suggests a rank-size relationship. In simpler terms, the population of the second-largest city is roughly half the population of the largest. Other properties of social, infrastructure and economic indicators scale with the city size. The analysis of scaling properties [83] of urban systems focuses on the non-linear (power-law) relationship between the population in a city (denoted as N, as the city size) and a set of indicators (Y) which quantify the socio-economic activity of a city. Ultimately this relationship is expressed as $Y = Y_0 \cdot N^{\beta}$, where Y_0 is a constant and β is the scaling exponent [45]. There are different scaling regimes in a city: super-linear scaling ($\beta > 1$, with $\beta \sim 1.15$ [45, 84]) characterizes the socio-economic activities of a city. Meaning that the amount of socio-economical dynamics to which a single individual is exposed is larger in more populated ones. While sub-linear scaling instead characterizes infrastructure related metrics (e.g. $\beta \sim 0.85$) [45, 84], such as the number of gas stations. As a consequence, in large cities less infrastructure is needed to provide the same amount of services to the single individual [27, 85]. This "economy of scale" lies behind many complex systems and further establishes complexity as a fundamental tool to study cities. We also mention that there is ongoing research highlighting limitations in these studies of scaling laws [86].

Mathematical models for scaling laws. Models have been devised to go beyond empirical observational descriptions, and rather propose the emergence of these scaling laws as the result of the multiplicative interplay and non-linear interactions between socio-economic activities and properties of infrastructure networks. [83].
As an example, in Ref. [87] they study the allometric scaling of the centers in a city. They start from a model in spatial economics for individual choices based on the maximization of a joint function of benefit/attractiveness W_j (quantified as benefit of working in a tile j) and the cost of commuting there: $Z_{ij} = W_j - C_{ij}$ [88]. Then they consider the Bureau of Road Traffic [41] function to model the travel time cost affected by congestion dynamics $C_{ij} \propto r_{ij}\,[1 + (T_{ij}/c)^{\mu}]$. An increase in population size is also associated to an increase in congestion in a monocentric structure (single centre with high W_i). This approach allowed to quantify the critical values of the population size such that due to an interplay of congestion and infrastructure the city transitions from a monocentric to a polycentric structure. This was condensed in the derivation of the scaling property $N^* \sim c\left(\frac{l}{\sqrt{A}N_c}\right)^{\mu}$, where N_c is the number of sub-centers, A is the city area, l a commuting distance, and μ, c instead regulate the non-linearity and transportation capacity. While other models also propose relevant mechanisms to explain scaling laws in cities [83], we discussed specifically this one as it exploits a model from spatial economic and one from traffic science. These models were derived in two separate sub-fields to gather insights on a complex systems were both layers interact. Moreover, both the definitions of cost C_{ij} (in Chapter 4) and attractiveness W_j are used in Chapter 4 and Chapter 5.

2.5.2 Morphology of the city

Another fundamental area of research is represented by the complex modeling relating the morphology of a city to network metrics and in turn socio-economic indicators. Relating these properties to real urban morphologies can inform policy-makers to plan more efficient and livable cities. These properties range from polycentricity versus monocentricity [87, 24], the studies on 15 minutes cities and their limitations [89], up to the process of urban sprawl [33].

As an example, the study of urban sprawl paired to morphology can quantify the increase in commute times and associated energy demands for transporting population within these urbanized areas. Leading to insights such as the prioritization of the development of compact and well-connected urban forms towards more livable and sustainable cities [33]. Other studies have highlighted also how morphology impacts the city livability: cities with well-connected, diverse street networks are more walkable and foster community development [90]. Dense, mixed-use neighborhoods with green spaces boost social interaction and economic activity. Even street design matters as fractal patterns improve accessibility [91, 92]. Moreover, studies on active mobility and the efficiency of public transit networks [35] to provide accessibility [93] also provide important insights to build sustainable cities.

Chapter 3:
Multi-pathways network effective distances and prediction of human mobility on diffusive scales

3.0.1 Chapter Overview

In this chapter we provide a theoretical view of information propagation on complex networks, with a focus on the definition of effective distances to predict human mobility-mediated processes. We study the interplay between the network connectivity structure and the interaction dynamics between nodes in determining patterns of perturbation propagation. We highlight the relevance of integrating the information of the ensemble of pathways or routes between origin and destination. Moreover, by leveraging the formalism of dynamics on networks this chapter explores how human collective movements can be conceptualized as flows between nodes, embedded in OD networks such as World Airtransportation Network (WAN), providing the arteries for the diffusion of information such as infectious seeds in a global pandemic. This chapter stems from our works in Ref. [1] and in Ref. [2].

In Section 3.2 we begin by providing analytical insights via a general study of perturbation propagation on networked dynamical models near steady state. By decomposing the perturbations across the ensemble of pathways, we highlight the significance of considering multiple pathways (MP) in the propagation of information between origin and destination nodes. Through the definition of a temporal distance, the hidden geometric structures underlying information propagation are unveiled. Subsequently, we focus into the specific application of MP distance in diffusion dynamics.
In Section 3.3 we then exploit the Laplacian operator to devise metrics better tailored for general diffusion dynamics, and introduce the concept of Information Distance D_{ID}. Through the integration of the Susceptible-Infectious-Recovered (SIR) model and diffusion dynamics within WAN nodes, human collective flows between airports are modeled as proxies for diffusion processes. We discuss how parsimonious metrics such as D_{ID} embed multiple pathways in their analytical structure, enabling predictions comparable to

existing methodologies, but with the advantage of reduced computational complexity. Finally, in Section 3.4 we present the application of effective distances derived from human mobility networks in assessing the pandemic potential of Variants of Concerns [2]. By integrating effective distances into a comprehensive pipeline for pandemic intelligence, effective distances such as D_{ID} is compared as benchmark against alternative methodologies, showcasing its efficacy in discerning pandemic spread patterns.

3.1 Information propagation on networks: interplay between structure and dynamics.

Non-linear differential equations that model the state activity of an element, and the formalism of networks that embed the dependence (either physical or functional) of an element's state to other elements' are two powerful mathematical descriptor that are often paired to model and simulate the behavior of intricate interdependent systems [94]. On this modeling framework, a theoretical analysis of reaction-diffusion mechanisms, synchronization processes, and more general complex dynamical models has started to uncover of some fundamental behaviors of these systems. In particular, analytical approaches that capture the interplay between the system dynamics and its structural connectivity have been of interest in recent years [16, 17]. Among others, the synchronization properties being contingent upon both the Laplacian spectrum and the non-linear equations for systems of oscillators [63, 95, 59] represent a relevant example. While in this chapter we will focus mostly on diffusion dynamics [59] and its application on human mobility scenarios, a general approach that explores the universal mechanisms that lie behind not only the movement of people but also other information spreading processes in nature is fundamental. Within these networked systems, connections can represent different pairwise interactions, ranging from mass action kinetics [8] to switch-like processes [6]. The specific choice of functions to emulate real-world mechanics yields distinct node responses and different patterns even on top of identical networks. Consequently, discerning the functional significance of each node, given the specific process, requires more than just knowledge of connection topology. Among other things, the interplay between topology and dynamics regulates the system's transient temporal response to a perturbation and the propagation of information. Therefore predicting arrival times of a signal [17, 96, 97, 98] is a critical challenge and a general analytical treatment that exploits universal properties is relevant. Furthermore, by understanding arrival times, we can uncover the hidden geometries induced by the dynamics [29, 65].

Recent research has explored how perturbations introduced in an equilibrium state of a networked dynamical system can shed light on this interplay of structure and dynamics governing information propagation [99, 16]. Analytical treatment under linear response approximation of these perturbations reveals the scaling laws governing the system's response to perturbations. This approach effectively disentangles the role of network topology (a node's degree k) from dynamics (the universal exponents obtained from the model's differential equations) and synthesizes the complex interplay determining several dynamical properties such as the system's stability [21]. This paradigm of scaling laws identifies different systems into universality classes of dynamical regimes, where units respond to and propagate information similarly [16]. While several metrics have been defined to quantify different aspects of a nodes' impact or efficiency in this framework [100], they in particular estimate a node's characteristic time response τ_i from a perturbation in a neighboring node as a scaling law of the node's degree with a universal exponent: $\tau_i \sim k_i^{\theta}$ [17]. We will exploit this framework to study in depth the

role of the ensemble of paths that connect source and target in conveying perturbation for a general networked dynamics. In the context of human mobility, we will also discuss the application on diffusion dynamics where we can see this perturbation as signal being transported by human flows.

3.2 Multiple pathways decomposition of perturbation propagation

We first introduce more in detail the general analytical framework devised for the study of these universality classes [99, 16]. We employ this framework to investigate the role of network paths in the propagation of a signal [1], and we employed the same notation for consistency.

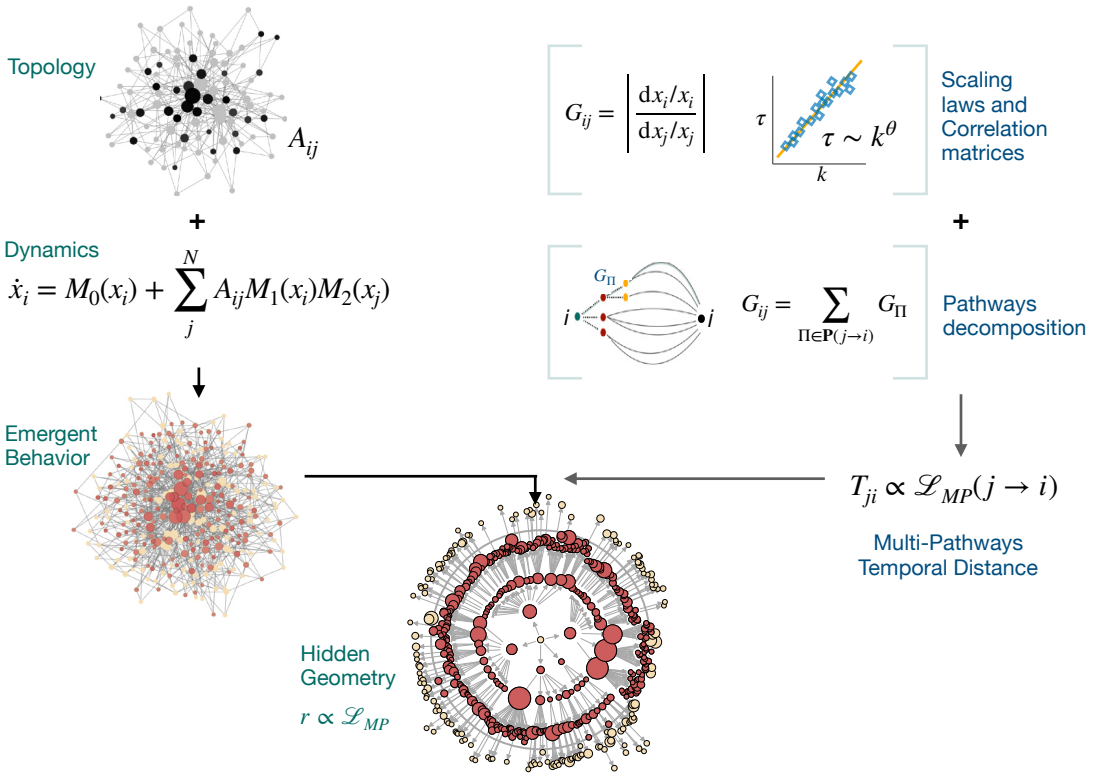

Figure 3.1: Hidden geometry of propagation patterns. The interplay between network structure A_{ij} and non-linear interactions (ODE) results in an complex propagation patterns. We combine analytical knowledge of scaling laws [17] and correlation matrices [16] with the decomposition of propagation on the ensemble of pathways to define $\mathcal{L}_{MP}$ and unravel the signal propagation. Figure is adapted from [1].

The analytical framework

Here each node's state can be described variable $x_i(t)$, $i = 1, ..., N$. This variable can encode a large set of quantities, such as fraction of genes in a regulatory network [6] or the fraction of infected individuals in an airport network [29]. The system's dynamics is encoded in a set of non-linear functions $\mathbf{M} = \{M_0(x), M_1(x), M_2(x)\}$, and the time evolution of $x_i(t)$ is driven by a set of non-linear Ordinary Differential Equations (ODEs):

$$\dot{x}_i(t) = M_0\left(x_i(t)\right) + \sum_{j=1}^{N} A_{ij} M_1\left(x_i(t)\right) M_2\left(x_j(t)\right). \tag{3.1}$$

Where $\mathbf{M}$ indicates a set of general functions: M_0 is element i's self dynamics, while M_1 and M_2 functions require that the interaction between i and j to be factorized. Expansions of $\mathbf{M}$ functions in steady state condition of Eq. 3.1 as power series of node degrees is the fundamental analytical step in recovering the scaling laws and universal exponents [16, 17, 101], such as the exponent θ in the aforementioned scaling law $\tau_i \sim k_i^{\theta}$. In this work we studied specifically three dynamical processes that represent different systems in nature, as presented in Fig. 3.2 with their relative θ exponent [17]. While in Section 3.3 we will discuss more in detail the case for diffusion dynamics. A small constant perturbation in a source node j's steady state ($x_j \to x_j + \mathrm{d}x_j$) is used as an analytical probe to derive analytical relationships linking the nodes' degrees to the exponents derived from the ODEs 3.1 and $\mathbf{M}$ functions. The system's response to this constant perturbation is to be shifted to a new steady state (e.g. $x_i \to x_i + \mathrm{d}x_i$). This is encoded definition of the Global Correlation Function matrix given the source node j [99]: $G_{ij} = \left|\frac{\mathrm{d}x_i/x_i}{\mathrm{d}x_j/x_j}\right|$. We remark that singularities in Eq. 3.2 do not pose an issue because steady state activities $x_i \neq 0, \forall i$ in the systems under study.

	Regulatory $\mathbb{R}$	Epidemic $\mathbb{E}$	Population $\mathbb{P}$
ODE	$\dot{x}_i = -x_i + \sum_{j=1}^{N} A_{ij} \frac{x_j^h}{1 + x_j^h}$	$\dot{x}_i = -x_i + \sum_{j=1}^{N} A_{ij}(1 - x_i)x_j$	$\dot{x}_i = -x_i^a + \sum_{j=1}^{N} A_{ij} x_j^h$
Exponent θ $\tau_i \sim k_i^{\theta}$	$\theta = 0$ $(h = 1/2)$	$\theta = -1$	$\theta = +1$ $(a = 1/2,\ h = 1/5)$

Figure 3.2: Models ODE. Ordinary Differential Equations of the three dynamical models representing different processes in nature. $\mathbb{R}$ Regulatory dynamics in gene transcription regulatory networks via Michaelis-Menten model [6], where x_i represents the expression of a gene. $\mathbb{E}$ Epidemic dynamics via Susceptible-Infected-Susceptible (SIS) compartmental model [102] where x_i represents the fraction of infected individuals in node i and $(1 - x_i)$ of the susceptible. Finally we also study birth-death dynamics, as a case of continuous time Markovian process where species' concentration x_i is increased/decreased at given rates. As it can be applied in demography we refer to it as $\mathbb{P}$ Population dynamics. The choice of exponents (a,h) is taken in order to obtain three different classes of time response regimes encoded in the exponent θ [17] of the scaling $\tau_i \sim k_i^{\theta}$, where τ_i is the typical response time of a node from a nearby perturbation. For example, this helps define whether an hub responds slowly or not in the specific dynamics. These models are also among the set of dynamics used in Ref. [99, 16, 17].

The concept of the emergence of a complex pattern from the interplay of structure and dynamics, and the process of embedding the propagation pattern in a predictive metric that encodes and predicts the interplay such that inherent geometries are obtained are presented in Fig. 3.1. We now introduce a path-driven approach to describe G_{ij} matrix elements. This involves decomposing this matrix as a summation across the ensemble of paths linking nodes i and j.
To this aim, we define $\mathbf{P}(j \to i)$ as the collection of all paths Π linking the source node j to the destination node i. This definition considers paths that may traverse vertices multiple times (walks) [103]. We quantify the correlation carried along each path, denoted as G_Π, as the product of local correlation elements R_{ij}. Specifically, G_Π is computed on the set of edges in the path and is determined by the expression:

$$R_{ij} = \left| \frac{\partial x_i / x_i}{\partial x_j / x_j} \right|, \tag{3.2}$$

as outlined in [16]. Here, R_{ij} captures node i's state dependency on a change in its direct neighbor j's state, as evaluated through the partial derivative in the steady state of Equation 3.1. Here we can also define a path Π of length L as a set of L nodes, sorted by visiting order: $\Pi(1 \to L) = \{n_1, \ldots, n_L\}$. Therefore we have that: $G_\Pi(L) = R_{L,L-1} \times \ldots \times R_{2,1}$:

$$G_\Pi(L) = \prod_{i=1}^{L-1} \left| \frac{\partial x_{i+1} / x_{i+1}}{\partial x_i / x_i} \right| = \prod_{i=1}^{L-1} R_{i+1,i}, \tag{3.3}$$

and G_{ij} can be rewritten as an sum over the ensemble of paths:

$$G_{ij} = \sum_{\Pi \in \mathbf{P}(j \to i)} G_\Pi. \tag{3.4}$$

This summation can be also extended by considering neighbors NN_i of target node i and taking out the sum over R_{ik}:

$$G_{ij} = \sum_{k \in NN_i} R_{ik} \cdot \left(\sum_{\Pi \in \mathbf{P}(j \to k)} G_\Pi(j \to k) \right) \tag{3.5}$$

$$= \sum_{k \in NN_i} R_{ik} \cdot G_{kj}, \tag{3.6}$$

of which the equality with G_{ij} can be understood since terms $\sum_{\Pi \in \mathbf{P}(j \to k)} G_\Pi(j \to k) = G_{kj}$, therefore by iteration we can recover the definition of G_{ij} proposed in Ref. [16]. A visual representation of this decomposition is present in Fig. 3.3.

Decomposition of G_{ij} on few paths of relevance

One of the scaling properties uncovered in Ref. [16] for systems described by Eq. 3.1 is that the average correlation G of target nodes at a topological distance l from the

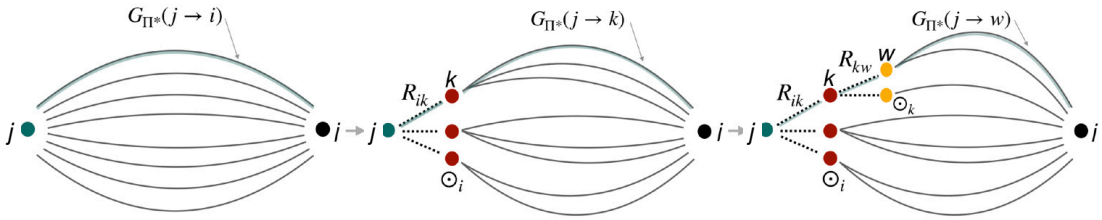

Figure 3.3: Decomposition of G_{ij}. The ensemble of paths between target i and source node j and their G_Π. Regrouping the paths via R_{ij} elements of set of node i's neighbors, allows to recover the definition of G_{ij} as in Eq. 3.4. Figure is adapted from [1].

source follows an exponential decay, such that $G(l) \approx exp(-l/\lambda)$. As a consequence, if we consider that $L_{ij} = L$ represents the topological distance (minimum number of edges) between nodes i and j, we expect that only paths of length L (or $L+1$) in the set $\mathbf{P}(j \to i)$ will significantly contribute to G_{ij}. We can employ Eq. 3.4 to reconstruct correlation matrix elements G_{ij} using a limited subset of path between each pair of nodes. Consequently, since $\mathbf{P}(j \to i)$ can be divided into subsets of paths, each with the same length l: $\mathbf{P}(j \to i) = \bigcup_{l=1}^{\infty} \mathbf{P}_l(j \to i)$, we aim to reconstruct the propagation from j to i by considering only the set of shortest paths in $\mathbf{P}_L(j \to i)$. Formally, this means restricting the summation in Equation 3.4 to the subset of shortest paths of length L_{ij}:

$$G_{ij} \approx \bar{G}_{ij} = \sum_{\Pi \in \mathbf{P}_L(j \to i)} G_\Pi. \tag{3.7}$$

We validate this assumption numerically by comparing exact values for G_{ij} with approximate values $\bar{G}_{ij}$ for the set of dynamical models described in Fig. 3.2. We quantify the accuracy of the reconstruction by calculating the absolute relative errors $\delta G = |G_{ij} - \bar{G}_{ij}|/G_{ij}$. Initially, we consider only single terms G_{Π^*} associated with the shortest path carrying the largest amount of correlation ($\Pi^* = \max_\Pi(G_\Pi)$). Then we include in the summation all paths having shortest path length $L_{ij} = L$ between the pair $i - j$: $\mathbf{P}_L(j \to i)$. Finally, we integrate in the summation also paths of successive orders $L_{ij} + N$.
We compare analytical predictions from Eq. 3.7 with simulations in Fig. 3.4, showing how we can approximate G_{ij} with just a small subset of paths, reducing the error δG in function of the order of paths considered. These findings suggest that when analyzing perturbation propagation between two nodes, focusing solely on the set of shortest paths (and potentially extending to a successive order $L+1$) is sufficient to capture the relevant pathways in the spatiotemporal propagation of perturbations and in general information that decays exponentially with a networks' path's length, as a trade-off between a single shortest path and considering all paths.

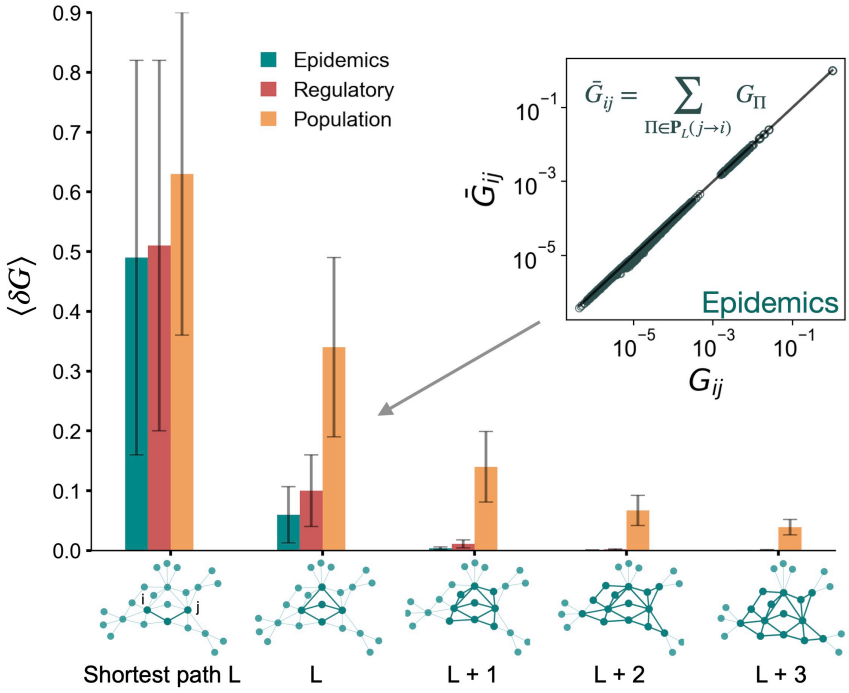

Figure 3.4: Error in reconstructing G_{ij} with few paths. For each dynamics we test the comparison of G_{ij} versus the approximation, on an Erdős-Renyi (ER) [66] synthetic network (N = 100, $\rho_{ER} = 0.15$). Average of absolute relative errors $\langle \delta G \rangle$ are presented versus the path orders in Eq. 3.4. The shortest path on average justifies only 50 % of the total perturbation propagated. By considering the entire set $\mathbf{P}_L(j \to i)$ of available shortest paths brings the approximation to $\sim 90\%$ retaining the relevant information. Adding next orders allows to reproduce G_{ij} more accurately. **Inset**) Matrix elements G_{ij} values against approximations $\bar{G}_{ij}$ for the case of SIS $\mathbb{E}$ dynamics, computed considering the subset $\mathbf{P}_L$. Figure is adapted from [1].

Temporal distance capturing the hidden geometry

As previously stated, as a consequence of the propagation dynamics, a target node i will shift its steady state to a new perturbed state, forced by the constant perturbation in node j, denoted as $x^{new}i(t \to \infty) = xi + \Delta x_i(t \to \infty)$. In Ref. [17], they propose an operative definition of temporal distance between the nodes $T(j \to i)$ as the time taken for node i to achieve a fraction η of its total response: $\Delta x_i(t = T(j \to i)) = \eta \Delta x_i(t \to \infty)$. By exploiting the scaling that predicts the typical response of a node from a nearby perturbation as a power-law of its degree $\tau_i \sim k_i^\theta$, a path Π cumulative lag time:

$$\tau_\Pi = \sum_{\substack{p \in \Pi(j \to i) \\ p \neq j}} k_p^\theta \tag{3.8}$$

can be introduced. In Ref. [17] they considered the shortest τ_Π as the candidate temporal distance from j to i:

$$\mathcal{L}(j \to i) = \min_{\Pi(j \to i)} \{\tau_\Pi\}, \tag{3.9}$$

which can be used to predict $T(j \to i)$.
Given that the overall response in G_{ij} stems from the collective behavior of all paths, where additional paths may contribute significantly (non negligible G_Π), it becomes fundamental to consider their temporal delays τ_Π in perturbation propagation. Specifically, we can first compute the time lag across each path τ_Π, then we assign weights to each path based on the proportion of perturbations it carries G_Π/G_{ij}, effectively normalizing such that $\sum_\Pi G_\Pi/G_{ij} = 1$. Consequently, we integrate effectively the two relevant descriptors of information propagation in this framework: local perturbations R_{ij} and relaxation times τ_i, and we obtain a novel temporal metric $\mathcal{L}_{MP}(j \to i)$ as a weighted linear combination of the cumulative time delay across each pathway:

$$\mathcal{L}_{MP}(j \to i) = \sum_{\Pi \in \mathbf{P}(j \to i)} \tau_\Pi \cdot W_\Pi \tag{3.10}$$

where:

$$W_\Pi = \frac{G_\Pi}{G_{ij}}. \tag{3.11}$$

As a blueprint for understanding propagation patterns, we remark that R_{ij} and τ_i can be derived analytically [101, 16]. Moreover, as we have shown in Eq. 3.7, the summation in Eq. 3.10 can also be restricted to only the set of shortest paths of length L with largest contribution G_Π, thus also providing analytical tractability. These analytical tools can be derived for any dynamics that can be casted into the model in Eq. 3.1.

Experiments

We assess the predictive performance of $\mathcal{L}_{MP}(j \to i)$ estimations in comparison to the prior temporal metric $\mathcal{L}(j \to i)$ [17] and the conventional shortest path length denoted as Lij [57]. To conduct our evaluation, we utilize a network model characterized by heterogeneous connectivity, specifically employing a Scale-Free network A_{ij} generated through the Barabási-Albert preferential attachment procedure [58]. We first numerically compute the equilibrium states of node activities for Epidemic, Regulatory, and Population dynamics (refer to Fig 3.2). We then select a source node and introduce a small constant perturbation in its steady state, and numerically solve Eq. 3.1 given the perturbation. This induces shifts in the steady states of other network units, and we determine response times $T(j \to i)$ [17] as defined previously. The resulting set of empirical perturbation times $T(j \to i)$ is then employed to evaluate the efficacy of candidate metrics. The accuracy is assessed using Spearman's correlation coefficient denoted as ρ_s. The average values of ρ_s over a set of experiments on different network generative models are presented in Table 3.1. We observe that ρ_s for our proposed metric closely approaches 1, despite the approximations of using scaling laws $\tau_i \sim k_i^\theta$ in estimating time responses, indicating that $\mathcal{L}_{MP}$ effectively predicts temporal distances [1].

	$\mathbb{E}$	$\mathbb{R}$	$\mathbb{P}$
L_{ij}	0.83 ± 0.01	0.88 ± 0.01	0.40 ± 0.02
$\mathcal{L}$	0.56 ± 0.01	0.88 ± 0.01	0.85 ± 0.01
$\mathcal{L}_{MP}$	0.99 ± 4e−4	0.99 ± 2e−5	0.99 ± 4e−4

Table 3.1: Accuracy of temporal metrics. We compute the average Spearman's ρ_s over multiple perturbation instances for Epidemic ($\mathbb{E}$), Regulatory ($\mathbb{R}$), and Population ($\mathbb{P}$) dynamics. We generate 10 synthetic networks (scale free [58] N=300 nodes). Within each network, we randomly select 10 nodes as perturbation sources and propagate a constant perturbation obtaining $T(s \to i)$. Subsequently, ρ_s is calculated for the predicted times versus empirical arrival times. Effective distances are computed using the topological L_{ij}, the single path $\mathcal{L}$ in Eq. 3.9 and the multi-pathway distance $\mathcal{L}_{MP}$. The average ρ_s and a 95% confidence interval are presented in the table for each dynamic.

In Figure 3.5, testing on a scale-free synthetic graph [58] A_{ij} with $N = 500$ nodes, $\mathcal{L}_{MP}(j \to i)$ exposes the underlying geometry shaped by the interaction set $\mathbf{M}$. We embed target nodes in a new geometry [17, 29], with a radius from the source $r \propto \mathcal{L}_{MP}(i \to j)$. By testing under different dynamics (Fig. 3.5) we can appreciate the adaptive nature of $\mathcal{L}_{MP}$. In Population dynamics $\mathbb{P}$, hubs respond slowly ($\theta > 1$), and are their embedded distance is far from the source, whereas in Regulatory dynamics $\mathbb{R}$, the degree-independence of $\tau(k)$ yields a discretized propagation pattern influenced by the topological distance Lij [17]. In this embedding the intricate propagation of the perturbation front mirrors the behavior of a travelling wave solution similar to a reaction-diffusion system described by a Fisher-like equation [104].

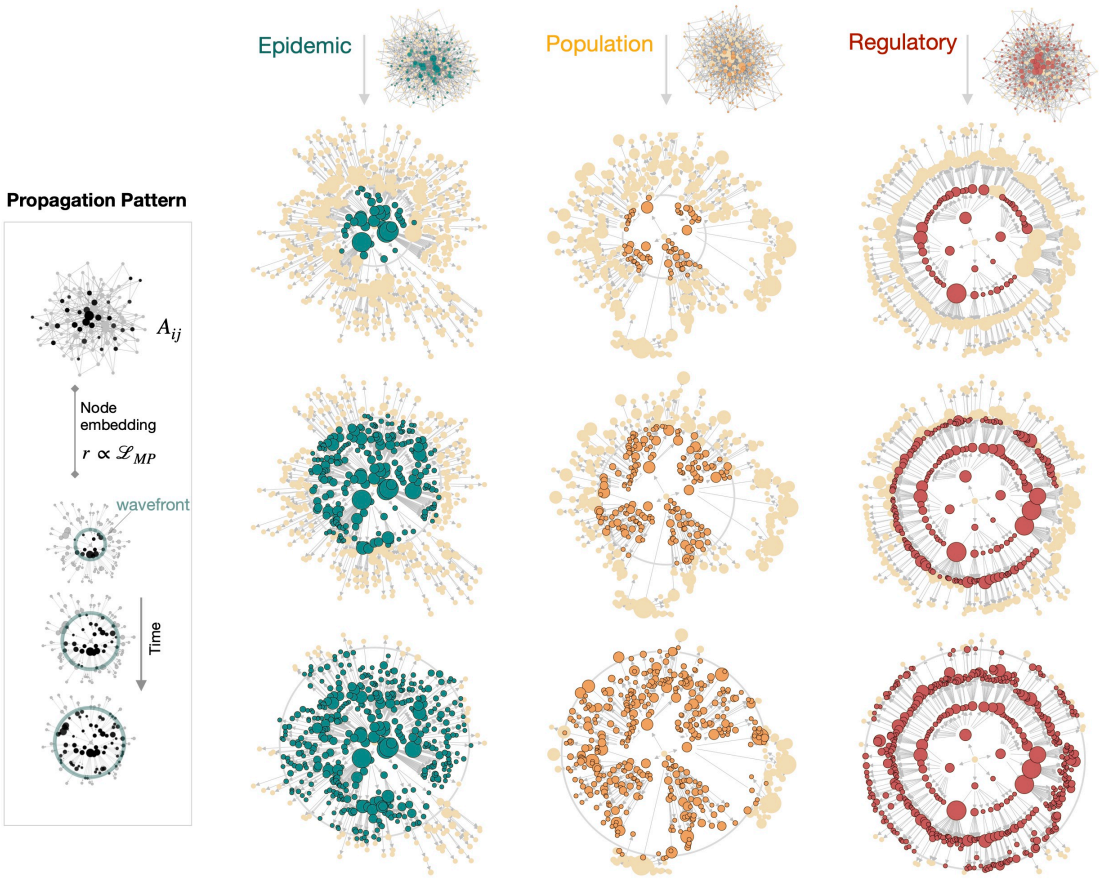

Figure 3.5: Embedding of spatiotemporal perturbation propagation. We embed the target nodes in a space where the radial distance r from the source node (placed in the center) is $r \propto \mathcal{L}_{MP}$. This embedding unveils a the hidden geometry of complex propagation, shaping it into a concentric pattern [17] of perturbed nodes, akin to the traveling wave solutions observed in reaction-diffusion systems [104, 29]. Notably, diverse perturbation patterns emerge from the same underlying topology Aij. Node sizes correspond to their degrees. Drawn edges are selected from the path Π^* with the highest correlation G_{Π^*}. This concentric geometry can also be used to infer the source given the propagation pattern [29]. Figure adapted from [1].

3.3 Diffusion dynamics and Information Distance

The study of epidemic, regulatory and population dynamics served as representative cases of different classes of propagation. In this section we focus on the case of diffusive-like processes. We study in this case the application of $\mathcal{L}_{MP}$ and introducing a novel metric D_{ID}, highlighting advantages and limitations in predicting complex spreading processes mediated by collective mobility.

Diffusion dynamics. The simplest case of a diffusion process between nodes in this framework is represented by the equation:

$$\dot{x}_i = +D \sum_j A_{ij}(x_j - x_i), \tag{3.12}$$

where D represents the diffusion constant. We derived the universal exponent for the scaling law of the temporal response $\tau_i \sim k_i^{\theta}$ following the methodology introduced in [17]. Namely, by computing small perturbations in the steady state in Eq. 3.1 in a mean field approximation paired to configuration model assumption, an analytical derivation of functions (built using the set of **M** functions) that govern the temporal response of δx can be obtained [16, 17]. Hans' series expansion of these functions yields the power law dependency of τ with the degree of a node which is perturbed. We study the **M** functions in Eq. 3.12 in this analytical framework and obtained the exponent $\theta \sim 0$. Therefore, in the case of diffusion dynamics the temporal response of a node is not dependent on its degree. We can apply the steady state condition of x_i setting Eq. 3.12 equal to 0, and deriving it with respect to node j's state, yielding:

$$R_{ij}^{diff} = \left| \frac{1}{k_i} \cdot \frac{x_j}{x_i} \right|. \tag{3.13}$$

The R_{ij}^{diff} element therefore takes the form of uniform transition probability from node i to any of its neighbors, normalized by the steady state activities x_i. The generalization to a weighted adjacency matrix (where $A_{ij} = w_{ij}$) is straightforward, with R_{ij} being proportional to a transition probability. We study empirical versus predicted distance on a SF and two ER networks. Results are illustrated in Fig. 3.6 where we compare L_{ij}, $\mathcal{L}$ and $\mathcal{L}_{MP}$. The decomposition of information propagation on multiple pathways improves the prediction in the case of diffusion dynamics up to a Spearman's of 0.97.

In the next section we study a more complicated model for reaction-diffusion, where collective human flows on the World Air-transportation Network (WAN) mediate the infectious seeds in a meta population SIR model. This model corresponds to the one described in [29], and we leverage the same dataset to reconstruct the WAN. It is important to note that while $\mathcal{L}_{\mathcal{MP}}$ has demonstrated efficacy within the theoretical framework for which it was designed, it was primarily intended to examine the role of

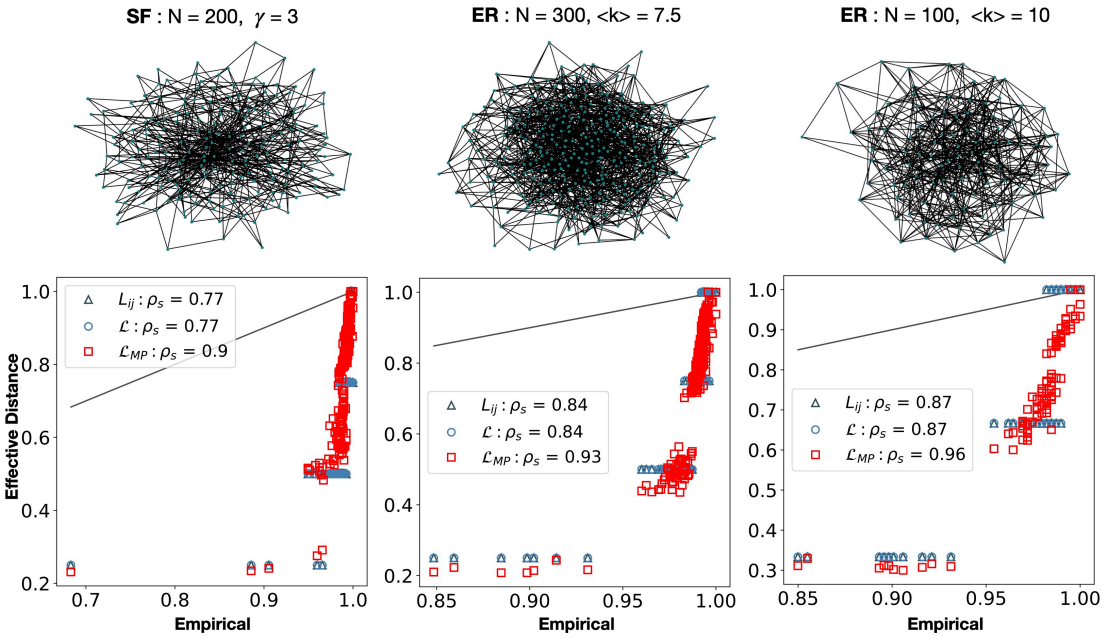

Figure 3.6: Diffusive dynamics and effective distances: We test $\mathcal{L}_{MP}$ on a simple diffusion model, against single path $\mathcal{L}$ and topological distance L_{ij} predictions. An example of a single instance of signal propagation is used to compare in a scatter plot the re-scaled effective distances of target nodes versus re-scaled measured empirical times $T(j \to i)$. This is done for three network generative models, a Scale Free (SF) and two Erdos-Renyi (ER) of different sizes and average degree $\langle k \rangle$ (see introductory Chapter 2). We highlight the improvement in ρ_s nodes' ranking when multiple paths are considered.

pathways in conveying perturbations from a steady state. We now introduce a general metric for dynamical processes mediated by diffusion dynamics.

3.3.1 Information Distance

We devise an alternative definition of distance on top of a network which embeds information from multiple-pathways diffusion. We formulate here an effective distance devised for diffusive processes exploiting the Laplacian operator [57, 59] of which powers embed information from multiple pathways. In recent years, there has been interest in distances derived from the mathematics of diffusion processes [65, 98]. Among others, we remark the concept of Diffusion Distance [59], which estimates a distances between two nodes by assessing the similarity in the exploration patterns of random walkers utilizing those nodes as starting points. In the embedding space induced by this distance the mesoscale structure [57] is recovered for the time scales in which the the functional communities is explored.

Building from Diffusion Distance and recent advances on operators built on the Laplacian such as the density matrix [60], we propose a metric tailored to rank and predict arrival

times of a random walker within the network. As $\mathbf{p}(t \mid i)$ defines the probability of a walker originating from node i to be in some other node at timestamp t, this can be recovered as successive applications (powers) of the random-walk Laplacian $L^{RW} = D^{-1}L = \mathbb{I} - D^{-1}A$. We introduce a new measure that also merges the insights from the effective distance in Ref. [29], namely, the idea that low probabilities $p_k(t \mid i)$ are associated with large distances. This can be embedded by taking the negative of the logarithm of the probability. We now introduce this candidate measure for diffusive dynamics which we define Information Distance D_{ID}:

$$D_{ID}^{(s\to k)}(t) = -\log_{10} p_k\left(t \mid s\right). \tag{3.14}$$

Where $p_k\left(t \mid s\right)$ denotes the k-th component, corresponding to node k, within the probability state $\mathbf{p}(t \mid s) = \mathbf{v_s} \cdot e^{-tL^{RW}}$, that we can therefore compute analytically from the knowledge of L^{RW}.
We can also see $p(t \mid i)_j$ as probabilities of finding an infected individual (walker) in node j at time t. Here, $\mathbf{v_s}$ represents the initial probability distribution for the walker originating from node s, defined as the canonical vector with the s-th component set to 1 ($\mathbf{p}(0) = \mathbf{v_i}$). The term of the random walk normalized Laplacian L^{RW} [57] encapsulates the probabilities of transitioning from node i to node j in its matrix elements. The off-diagonal elements of L^{RW} can be computed as the negative values of P_{ij}, which are directly derived from the WAN weighted number of passengers. We evaluate the metric across different timescales t to determine the scale where $D^{ID}(t)$ demonstrates its optimal performance (e.g. in the following section is $t = 1.0$). The exponential operator $e^{-tL^{RW}}$ naturally embeds the multiple-pathways decomposition of information propagation. This can be better formalized by expanding the Laplacian via Taylor expansion:

$$e^{-tL^{RW}} = \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \left(L^{RW}\right)^n. \tag{3.15}$$

Where the powers $(L^{RW})^n$ return the probability sum of the pathways on length n. Specifically, entry $i - j$:

$$\left[\left(L^{RW}\right)^n\right]_{ij} = \sum_{\Pi_{ij} \in \mathbf{P}_{(L=n)}} \mathbb{P}(\Pi_{ij}). \tag{3.16}$$

Where $\mathbf{P}_{(L=n)}$ if the set of paths Π_{ij} of topological length $L = n$ between nodes i and j, and $\mathbb{P}$ is the probability of that path computed as product of each random-walk transition probability along the edges of Π_{ij}.

3.3.2 Predicting arrival times of a reaction-diffusion model on the World Airtransportation Network

In this section we test $\mathcal{L}_{MP}$ and D^{ID} on a reaction-diffusion system that models a more realistic scenario where a network reconstructed from collective human flows acts as a proxy for the spreading process. Here we exploit a dataset of real collective human flows of the World Airtransportation Network (WAN) [29].

WAN Network

Original dataset [29] sourced from https://www.oag.com consists of 3924 airport nodes and 51541 links (each link having the associated number of seats per year), corresponding to a total flux of 4202 million seats per year. To ease the computation, we work with the dataset pruned of airports representing only 5% of the total traffic, yet constituting approximately 70% of the total nodes. The pruned WAN network consists of 965 airports and 33369 routes, with a total flux of 3808 million seats per year. The WAN represents a weighted undirected network as each edge (route) is weighted by daily number of seats F^*_{nm}, where n and m indicate the airports. As weights in routes direction are similar $F^*_{nm} \approx F^*_{mn}$, detailed balance is satisfied and we consider: $F_{nm} = (F^*_{nm} + F^*_{mn})/2 = F_{mn}$. the airport network is presented in Fig. 3.7

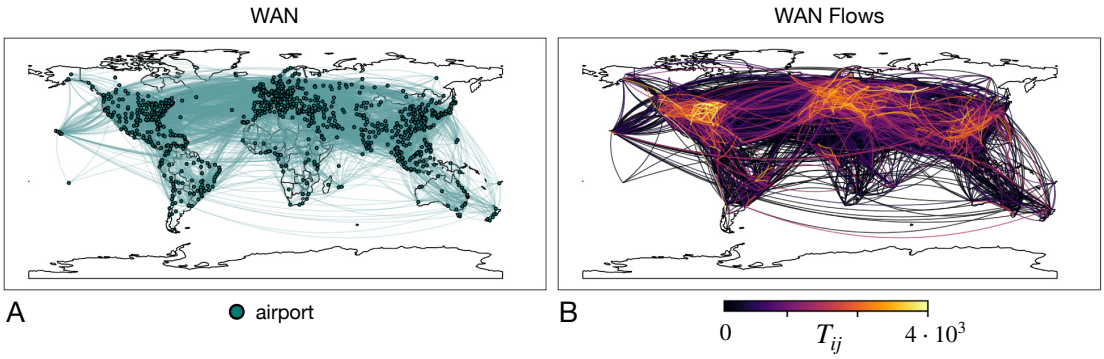

Figure 3.7: World Airtransportation Network. A) Network of airports (nodes) connected by edges (routes) [29]. **B**) Weighted edges represent OD flows T_{ij}, as number of seats per day (F_{ij} in the model).

Now we introduce the meta-population model where in each node a local population of susceptible individuals are infected by the seeds (infectious individuals) propagated by collective mobility. This modeling was introduced and thoroughly described in Ref. [29] as test benchmark for effective distances. We remark that in this section we indicate network nodes/airports using n and m indices.

SIR coupled to Diffusion model

First, in each node n we adopt a SIR [102] compartmental model. Standard modeling assumptions are adopted, such as population N_n is considered constant, and each individual's state belongs to one of three compartments: susceptible (S), infected (I), or recovered (R), subject to the normalization condition: $S_n + R_n + I_n = N_n$. Finally, the homogeneous mixing assumption is adopted. Transition probabilities among these states are encoded as:

$$\begin{aligned} I_n + S_n &\xrightarrow{\alpha} 2I_n \\ I_n &\xrightarrow{\beta} R_n \end{aligned} \tag{3.17}$$

Here, α and β are the transmission rate and average recovery rate respectively, both measured in units of [day^{-1}]. These determine the Reproduction Rate $R_0 = \frac{\alpha}{\beta}$, which indicates the expected number of secondary cases produced by a single infected individual. Moreover, individuals in a node can transition from node (population) m to n at a rate denoted by w_{nm}. By considering the normalized states in each population, such as $i_n = I_n/N_n$ being the fraction of infected individuals in node n, we obtain the final model [29]:

$$\begin{aligned} \partial_t s_n &= -\alpha s_n i_n \sigma\left(i_n/\epsilon\right) + \gamma \sum_{m \neq n} P_{mn}\left(s_m - s_n\right) \\ \partial_t i_n &= \alpha s_n i_n \sigma\left(i_n/\epsilon\right) - \beta i_n(t) + \gamma \sum_{m \neq n} P_{mn}\left(i_m - i_n\right). \end{aligned} \tag{3.18}$$

Here the r.h.s. divides into two terms: a self-dynamics term (e.g. infection dynamics within population i) and a diffusive term governing the inter-population movement. Fraction of recovered r_n can be derived from the constant population condition ($r_n = 1 - s_n - i_n$). Rates w_{nm} can be determined from WAN data F_{nm}, and the population size of the originating node N_m: $w_{nm} = F_{nm}/N_m$. A detailed balance condition can also be recovered under the assumption that the number of passengers traveling each day is, on average, the same in both directions: $F_{nm} = w_{nm}N_m = w_{mn}N_n = F_{mn}$. Finally N_m can be computed from the assumption that only a fraction γ of the total population N_n travels from m daily, hence:

$$F_n = \sum_m F_{mn} = \gamma N_n. \tag{3.19}$$

Inserting this information in the detailed balance condition $(N_m/N_n)w_{nm} = w_{mn}$, allows to reformulate the diffusion rates as follows: $w_{mn} = F_{mn}/N_n = \gamma F_{mn}/F_n$. Resulting in:

$$w_{mn} = \gamma P_{mn}, \tag{3.20}$$

where P_{mn} are actual transition probabilities. Finally, we note that in Eq. 3.18 an additional switch-like function is added to initiate infections in a node solely once the proportion of infected individuals surpasses a predetermined threshold ϵ, employing a sigmoid function:$\sigma(x) = \frac{x^\eta}{1+x^\eta}$.

Simulation of Epidemic Spread and Prediction of Infection Times

We now predict the arrival time of the infection versus effective distances estimations. The arrival time $T(OL \to n)$ is determined numerically when the epidemic curve in node n reaches its peak: $i_n\left(t = T(\mathrm{OL} \to n)\right) = \mathrm{peak}\left(i_n(t)\right)$, where OL is the Outbreak Location.

We first set the epidemic and mobility parameters of the model in Eq. 3.18. Specifically we set $\alpha = 0.42 \;\; [d^{-1}]$, $\beta = 0.285 \;\; [d^{-1}]$, $\gamma = 3 \cdot 10^{-3}$, and $\epsilon = 1 \cdot 10^{-6}$. The epidemic parameters α and β result in a reproduction rate $R_0 = 1.5$. We simulate two instances of contagion processes, the first chosen Outbreak Location (OL) for the contagion is the Federico Garcia Lorca Airport in Granada (IATA airport code GRX) and the second

is the Venice Marco Polo Airport in Venice (VCE). These selections are motivated by their status as nodes with different outflows $F_i = \sum_j F_{ij}$, VCE has an average annual outflow equal to 5.3 Million passengers while GRX 415 Thousands passengers. Fraction of infected individuals is set at $i_{OL}(t=0) = 10^{-3}$ and we monitor the progression of $i_n(t)$ through numerical solution of Equations 3.18.

Results. In Fig. 3.8, we show the scatter plot of effective distances against peak infection times. Recent metrics [65] have also exploited arrival time statistics devised for these diffusion-like processes [98], and allow to decompose information propagation on multiple pathways in the case of flows governed by probability transitions. We compare the metric $\mathcal{D}_{MP}^{RW}$ alongside the D_{ID} and $\mathcal{L}_{MP}$ metrics. In Appendix 7.1 we discuss the limits of extending $\mathcal{L}_{MP}$ to the SIR model in Eq. 3.18. We also calculated the geographical distance $\mathcal{D}_{Geo}$ obtained using the haversine formula for spherical coordinates. Remarkably, D_{ID} provides results comparable to the state of the art $\mathcal{D}_{MP}^{RW}$ which was specifically devised for these processes, while representing a simple analytical parsimonious modeling requiring only the Laplacian of the mobility matrix.
Finally, we remark that while this metric is built on top of the flows between airports on a world scale, their application can be extended also on more national-level scale. In the seminal work in Ref. [29], the mobility networks of Germany were used to predict the OL (outbreak location) of a national scale pandemic.

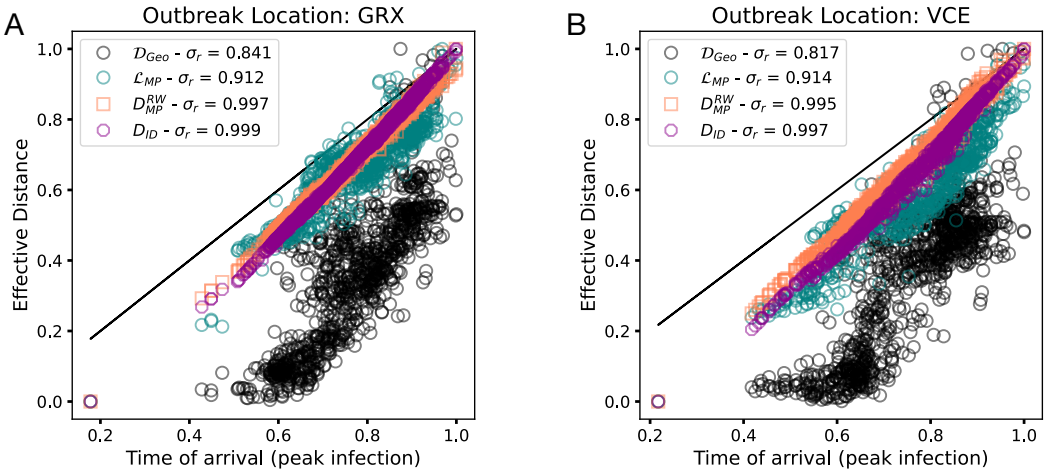

Figure 3.8: The case for diffusion dynamics and SIR modeling. A) Empirical infection times $T(\text{OL} \to n)$ as peaks of $i(t)$ curves of the target nodes are compared with re-scaled distances. Two Outbreak Locations are considered GRX (Granada) and VCE (Venice). Spearman's rank is used to evaluate the accuracy of metrics. **A)** Single path distance $\mathcal{D}_{SP}$ [29] and Multi-path distance $\mathcal{L}_{MP}$ are the worst performing. The metrics $\mathcal{D}_{MP}^{RW}$ [65] is compared with Information Distance D_{ID} with similar scores. The scores ~ 1.0 show that these metrics can accurately predict arrival times when the diffusion process is mediated by mobility flows.

3.4 Application on a Pipeline for Pandemic Intelligence

We now discuss the application of effective distances in a pipeline to assess the pandemic potential of a SARS-CoV-2 variant [2]. The SARS-CoV-2 pandemic rapidly spread globally [105], prompting a pandemic declaration by the World Health Organization on March 11, 2020, and resulted in millions of fatalities despite public health measures [106] such as widespread testing, masking, and vaccination [107]. The task of controlling the spreading was further complicated as novel variants with heightened immune evasion and/or transmissibility emerged, with the additional consequence of rendering modeling forecasts outdated. Ongoing research, surveillance, and global cooperation are crucial to mitigate the impact of the pandemic and preventing future outbreaks.

In this work we integrated phylogenetic data from national genomic surveillance coming from a limited set of initial sequenced samples, with knowledge of human mobility networks to infer country-specific projections of epidemic curve scenarios and ultimately provide an early assessment of the pandemic potential of a Variant of Concern (VoC) [2]. This pipeline represents a blueprint for pandemic intelligence, a complexity-based modeling framework that merges computational and analytical tools to provide a comprehensive understanding of epidemic spreading. We have validated the methodology using Alpha, Delta and other variants [108]. This allowed to also obtain insights into the pandemic potential of BA.5, BA.2.75, and other related sub-lineages. In comparison to conventional country-level epidemic surveillance, we revealed that only via an integrated multi-scale approach can the pandemic potentials of various lineages be accurately compared.

3.4.1 Pipeline

In this section we briefly discuss the pipeline, which we also briefly describe in Fig. 3.9. Then we focus on how the effective distance can be used as a fundamental step in this pipeline to provide the number of infectious seeds that traveled to a target country. Here D_{ID} is employed as a test metric and compared to more sophisticated metrics that model behavioral elements such as the exit probability from a node (airport), as the Import Risk distance [109].

(1) Epidemic parameters and estimation of VoC origin from sequencing data. SARS-CoV-2 sequences associated with the Alpha B.1.1.7, Delta B.1.617.2, and Omicron B.1.1.529 (BA.1), BA.2, BA.5, and BA.2.75 lineages were collected from GISAID [110]. On these sequences a Bayesian evolutionary reconstruction of timed phylogenetic history [111] was performed to derive posterior distributions for the growth rate, molecular clock parameters, and the time of the most recent common ancestor (tMRCA). This approach enabled us to estimate the time of the initial unreported case t_0, along with other epidemic parameters. Using these estimates, we calculated the effective reproduction number and generation interval. These parameters were then fed to the next modules of the pipeline (see Fig. 3.9).

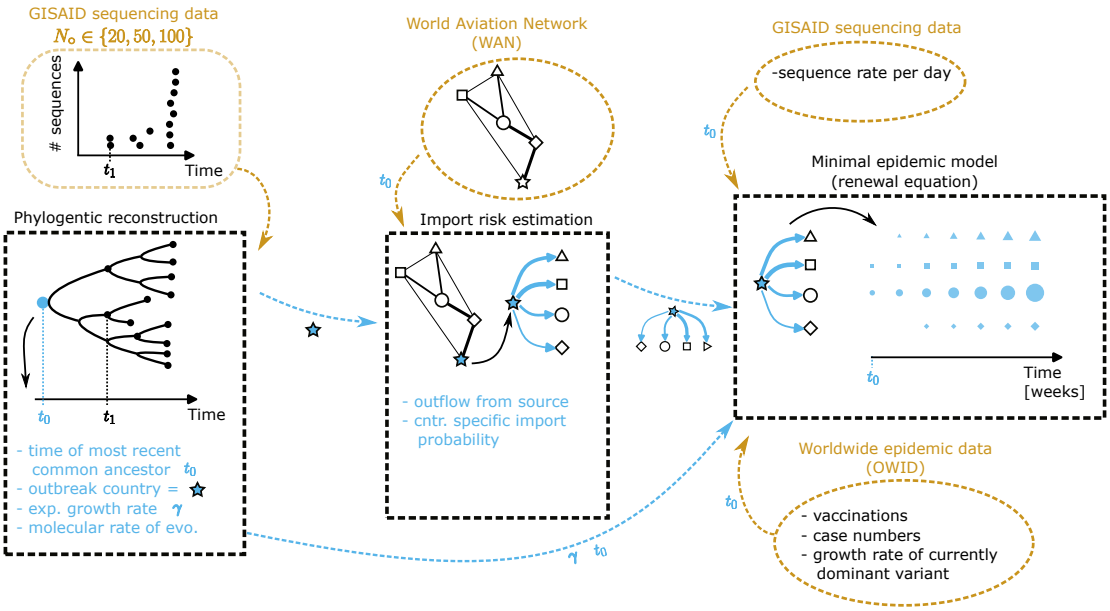

Figure 3.9: Pipeline. Three core modules of the pipeline (dashed boxes in black). Each module takes both inputs from external data sources (orange modules) or predictions from previous modules (dashed blue arrows). In each module the epidemic and mobility estimates inferred from data are reported in blue. Figure is taken from Ref. [2].

(2) Infectious seeds estimation. Monthly seat capacities in the WAN dataset for each month of 2020, 2021 and 2022 were used to estimate effective distances (see Section 3.3). Effective distances proportional to the mean first arrival time such as Import Risk [109] are computed from WAN as $D^{IR} = -\log p_\infty(n|n_0)$ thus probabilities that a traveler from source n_0 ends to a target country n can be directly derived. Here $p_\infty(n|n_0)$ is a multi-pathways metric that also models the probability that a random walker exists from a node along a specific path before jumping to the next node. As these distances are derived for the network of airports, nation-level distances can be computed by aggregating distances from each airport in a country m and weighting its distance using its outflow F_m. From the information of airport's outflows in WAN data, we can weight this probability and compute the mean number of travelers to target nodes. The probability of these travelers to be infected is computed by taking the daily incidence of new cases (7 day moving average) times the proportion of sequences of that VoC. We also considered possible under counting scenarios. Therefore the final number of infected seeds travelling from n_0 to target is computed as the number of travelers times the proportion of infected individuals (VoC) in source country. Correlations between effective distances and reported arrival times is shown in Subsection 3.4.2.

(3) Epidemic modeling. Finally, to forecast new daily infections we adopted a renewal model [112] in the main pipeline (a multi-strain SIR-like model was used for validation). The renewal equation requires minimal data: the daily reproduction number

$\mathcal{R}_s$, its generation interval distribution Γ_s, and $I_{out}(t)$ the number of infected individuals (seeds) traveling from the source country to the target (as estimated from previous step). Finally, in the model the number of infected at time t is therefore estimated as $I(t) = I_{out}(t) + I_{in}(t)$ and $I_{in}(t)$ is modeled as:

$$I_{in}(t) = \sum_{s=t_0}^{t} \Gamma_s \mathcal{R}_s I(s), \tag{3.21}$$

where t_0 is the day of arrival of first infected cases. A more complex modeling was also devised to include inhomogeneous immunization landscape due to vaccination and previous infections processes from other variant. Here we limit the discussion to the Renewal as a core example on how the pipeline parameters and estimates can be merged.

3.4.2 Seeds estimation with effective distances

Geographical distance D_{geo} and effective distances built on WAN, specifically the effective random walk distances ($D^{(N)}_{eff,MP}$ [29] and $D^{(N)}_{RW,MP}$) derived for reaction-diffusion models [65], the Information Distance D_{ID} and import risk D^{IR} [109] are used to estimate the number of infectious seeds traveling from source country to a target country. In step (2) of the pipeline the steps required to estimate the number of travelers from these probabilities have been presented. We note that in the operational pipeline, estimates from import risk distance D^{IR} [109] were used, as the integration of modeling aspects such as exit probabilities allowed to better infer the arrival of infectious seeds when we performed validation (see Fig. 3.10). We remark that the Information Distance $D_{ID,MP}$ provided estimates similar to more sophisticated metrics, simply by computing the logarithm of an operator built on the random-walk Laplacian L^{RW} of the WAN.

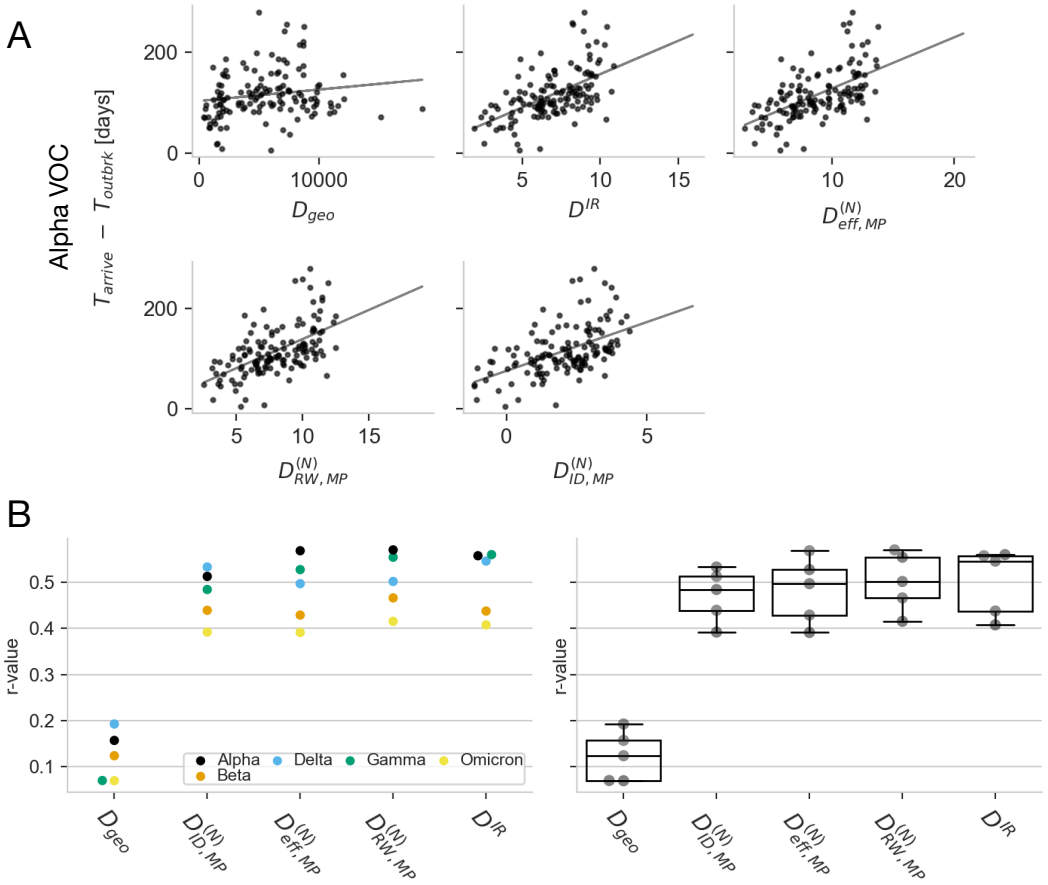

Figure 3.10: Comparison of effective distances estimation versus arrival times of the variants. A) Distance measures vs. arrivals for Alpha variant. The distance measures are the geographic distance D_{geo}, the import risk distance D^{IR}, the effective distance $D^{(N)}_{eff,MP}$, the random walk distance $D^{(N)}_{RW,MP}$, and the information distance $D^{(N)}_{ID,MP}$. **B)** Correlation comparison between different distance measures and different arrival times of VoCs under study. The uncertainties in reporting official arrival times by national authorities (due for example to low genome sequencing rates) and the underlying stochasticity in the mobility of passengers are possible reasons behind correlations remaining moderate (up to r-value ~ 0.56) even in the best performing metrics. Figures is taken from [2].

3.5 Conclusion

In this chapter we studied the role of multiple-pathways in defining effective distances on top of networks. These effective distances provide an analytical description of the hidden geometry derived from the interplay of network structure and non-linear interactions between nodes. We define a multi-pathways distance for a generic dynamical model on networks derived from universal scaling laws, of which knowledge allows the estimation of the time of arrival of a signal propagated from a source node. We have also discussed how a simple distance D_{ID} derived from the random-walk Laplacian effectively captures the time of arrival of a signal in a reaction-diffusion process mediated by collective mobility flows. Finally, we have shown how by leveraging data from human movements and exploiting network-based effective distances we can obtain valuable information that can inform a general pipeline for pandemic intelligence.
In the next chapter (Chapter 4) we will focus on shorter scales where the definition of an effective temporal distance is used to optimize transportation networks. Moreover in the context of urban science we exploit collective mobility models to recover spatial features of transportation networks similar to empirical complex topologies.

Chapter 4:
Optimal topologies in flow-weighted transportation networks

4.0.1 Chapter Overview

Transportation and distribution networks represent the arteries for human mobility. These networks are characterized by complex spatial topologies: from tree-like topologies, regular graphs with lattice-like structures, to more complex ones such as networks with a connected core paired with mono-dimensional peripheral branches. The latter in particular can be observed in both natural and man-made systems like subway and railway networks, especially in urban environments. In this chapter we devise a minimal and parsimonious framework where the optimization of a temporal effective distance on a network planar substrate leads to the emergence of these complex topological patterns. By optimizing the networks' edges velocities via a simulated annealing schedule, we study the traffic conditions under which these topologies represent optimal structures. We investigate how the interplay between spatial traffic flow probabilities and travel costs shapes the resulting optimal network. In the context of urban transportation, we also simulate flows using human mobility models to derive Origin-Destination matrices. This chapter stems from our work in Ref. [3].

In Section 4.2 we first introduce a simple framework for generating spatial networks on a lattice planar substrate. Here network's weights act as velocities which in turn define an effective temporal metric reflecting the efficiency of the transportation network in moving traffic flows across the urban space. On this framework, In Section 4.3 we show how introducing simple probabilities biasing the optimal efficiency which only depend on the distance, force a transition between a tree-like topology and a network resembling a simple lattice. Moreover, in Section 4.4 we discuss the effect of introducing travelers' behaviors in choosing the route between origin and destination, following concepts such as User Equilibrium (UE) and congestion dynamics [41, 113, 114]. There we route traffic demand under UE in the optimization of the substrate's velocities, uncovering a transition from the tree structure to a multiple-pathways alternative structure necessary to alleviate the temporal delays introduced by congestion effects. In Section 4.5 we show also that the modeling of flows which reproduce patterns typical of collective human

mobility in urban areas leads to the presence of preferential network paths that mediate the traffic between the largest OD pairs. The optimal topologies defined by these paths also showcase features seen in real systems. Features such as a bi-modality in the edges' velocity distribution, characteristic of multi-layered transportation, and branches coupled with large loop formations typical of subway systems [115, 30] are recovered, reflecting a balance of heterogeneous mobility patterns such as short-range travels and longer-range commutes. Finally in Section 4.6 we show an application of the model within the Greater London Area by exploiting Census and OSM data, finding scaling similarities of the optimal synthetic network generated by our framework with the London Underground network.

4.1 Introduction

Transportation networks have a fundamental role in mediating human activity and mobility [14, 23] at different scales. Extensively studied over the last two decades, [75, 116, 117, 118] the analysis of this class of spatial networks often revolves around optimizing and/or minimizing cost-based metrics with the aim of unveiling the networks features that determine their efficiency or that can be used as benchmarks for policy making [24, 67], as we discussed in Chapter 2. The concept of optimal networks, grounded in energy-like minimization principles, finds a natural intersection with concepts from physics [69]. Systems that minimize functional trade-offs, such as the minimization of a free energy, often reflect the most probable states observed in real-world systems. Although deriving these physical variables directly from first principles may be challenging in complex systems like cities, analogies and concepts from physics provide valuable perspectives for understanding and analyzing these systems' structure and dynamics. Recent studies have explored simple laws that can drive the emergence of hierarchy and better highlighted the influence of traffic on optimal network states [119, 67]. Additionally, both global and local optimization criteria play a significant role in shaping the development of man-made transportation systems. Hence, a network science perspective can offer insights that policymakers and planners can leverage in their urban planning efforts.

Complex topologies characterise the spatial organization of transportation infrastructures across different scales and modalities. As an example, a sparse, scale-free network topology characterises the World Airtransportation Networks seen in Chapter 3. Other modes of transportation, as a consequence of having the link infrastructure embedded in the geographical space such as rail networks, exhibit different network topologies. Among alia, branches paired with loops represent some of the characteristic structures of man made systems such as railways and subways, but also seen in natural systems [120, 30, 43, 115, 54, 30]. These structural elements play a pivotal role in facilitating effective urban transportation [30].

Emergence of complex topologies via optimization of a network substrate. We introduce a framework based on the definition of a lattice substrate which acts as an effective planar discretization of space, on which network topological features may emerge from an optimization process. We aim to uncover the minimum requirements to replicate these non-trivial structures as outcomes of the optimization process, especially in the context of human transportation in cities [121].
At variance with the recent works on network efficiency and transportation topologies, we adopt some fundamentally different modeling choices. We evaluate the efficiency in terms of time necessary to explore the network, where edges' weights w_e act as travel speeds. We optimize these speeds in a continuous interval, instead of using a multi-layer constraint or specific topologies [47, 122]. It is worth noting that we refer to "optimal network" topology when the functional form computed on the network properties is minimized, while definition of optimality in applied transport planning may differ based on the specific target [123].

The underlying network lattice (as represented in its simplest form by the triangular lattice in the next sections) acts as a planar substrate that allows the network to evolve [36, 122] and possibly exhibit the network topological features typical to real world systems. Specifically, we seek to reconstruct these topologies by means of an optimal configuration of a network. This is done assuming a fixed total amount of edges' weights (e.g., a constraint in the expenditure available on infrastructure). The optimal configuration is the assignation of edges' velocities, such that the cumulative amount of travel time between origin-destinations is minimized for all pairs of nodes [124, 47]. Additionally, as these networks represent the arteries in urban navigation via public transportation, we investigate the role of traffic flow between these OD pairs [125, 119] in weighting the importance of routes. Moreover, the modeling of urban spatial features, allow us to generate OD demands which simulate human mobility in urban areas, biasing these optimal networks to converge towards specific topological structures [23].
While this framework represents a theoretical abstraction which allows us to embed mechanisms such as human mobility [23] and congestion [41], it does not consider more complex elements such as mode choice and trip frequency, which are crucial in urban planning [126]. We remark that our objective is to investigate optimal network features using parsimonious modeling choices, rather than directly informing urban planning [127, 119, 124, 47, 23, 41, 126].

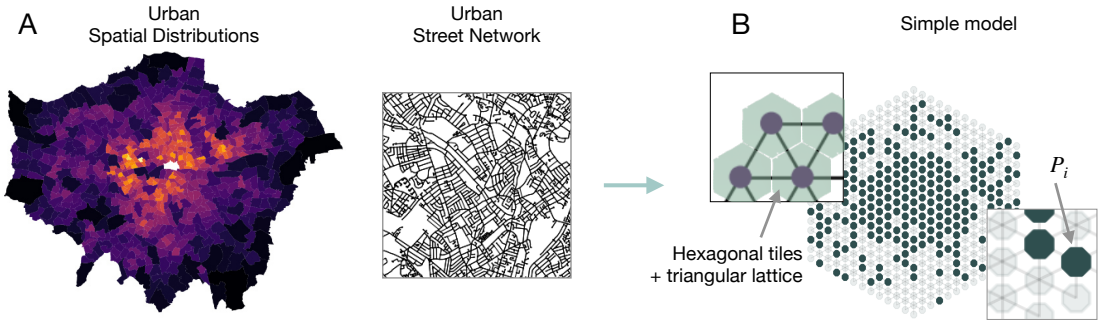

Figure 4.1: Planar network framework and urban structure mapping. A) Population distribution in Greater London Area wards from Census 2014, and sample from Open Street Map (OSM) urban street networks. **B)** These spatial distributions and the urban transportation network can be mapped to a planar lattice substrate. There, nodes encode urban features (e.g. P_i as population in node i) and edges can be optimized. Example with hexagonal tiling and its dual, the triangular lattice.

4.2 Simple framework for urban spatial structure

We devise a simple framework to study and optimize a transportation network between a set of areas. A lattice substrate serves as a network-based planar discretization of the urban space between these areas (nodes). Nodes can embed urban attributes such as population density and/or distribution of points of interest (POIs). This ultimately can

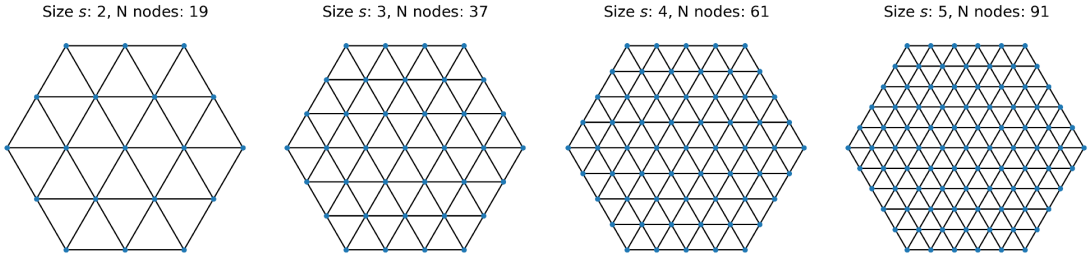

Figure 4.2: Triangular Lattice sizes s. Triangular lattice substrates for different sizes s ($s = 2, 3, 4, 5$) with associated N of nodes. Each lattice size s can also be seen as the number of s layers of nodes that are added from the center. Therefore given a size s the number of nodes N in the lattice can be computed as $N = 1 + \sum_{k=1}^{s+1}(k-1) \cdot 6$. Figure adapted from [3].

be generalized as a simple modeling of a spatial urban structure.
The most elementary example of this substrate can be found in the two-dimensional hexagonal tiling (based on the hexagonal lattice) [128] and its corresponding planar dual, the 2-dimensional triangular lattice [14, 129]. A triangular lattice is a spatial network embedded in 2 dimensions, where nodes have degree $k \geq 3$ and each edge has fixed length d. The concept of a dual representation also allows to directly connect the spatial tile (node) to its adjacent nodes in the substrate (refer to Fig. 4.1). The physical adjacency between nodes or tiles i and j is expressed through the adjacency matrix A, where $A_{ij} = 1$ denotes adjacency between regions in the lattice. We denote different lattices sizes as s. Size s refers to a triangular lattice with s-layers of additional nodes form the central node. Samples of different triangular lattices s is shown in Fig. 4.2. We remark that discretization of space via hexagonal tiling allows symmetry in the connectivity between nearby tiles [128]. Being isotropic, it represents an ideal test-bed as it presents fewer equivalent degenerate paths compared to a rectangular lattice.

Subsequently, network metrics and effective distances are computed and optimized on this network substrate following a suitable definition of efficiency or cost [36]. The network distance that we aim to minimize (optimize) is an effective temporal distance computed on the planar substrate. This approach relies on the fundamental assumption that in urban mobility the cost of traveling is not defined by the geographical distance between an origin and destination point, but rather on the amount of travel time required. Examples such as public transit networks exhibit hierarchical organization [43, 75], with distinct characteristic speeds across links in their multi-layer structure [72]. Therefore to improve the accessibility and efficiency, the optimization of an effective temporal distance on top of a transportation network is essential.

4.2.1 Flow-weighted temporal Efficiency

We denote e an edge in the substrate network, w_e as its weight representing its velocity within the transportation network and d_e as length of edge e (euclidean distance between its vertices). Edge weights are visually represented as the widths (as in Fig. 4.3). Definition of d_e is relevant in cases where the substrate is not represented by regular structures such as triangular lattice but by random spatial networks where edges may vary in length. For the generalization to non-spatial networks (non embedded in euclidean space) the model can be adapted by using topological distances, and $d_e = A_{ij}$. Given these definitions, the free-flow travel time of an edge e is therefore d_e/w_e. Refer to Fig. 4.3 for a visual representation.
Transportation efficiency between nodes i and j is quantified in terms of time cost [44], and our objective is to maximize this efficiency (or minimize travel time). Hence given the set of weights $\{w_e\}$, we compute the travel cost c_{ij} to go from source node i and terminating at destination node j, as the path having the minimum cumulative total travel time:

$$c_{ij}(\{w_e\}) = \min_{\Pi \in \Omega_{\Pi_{ij}}} \left[\sum_{e \in \Pi_{ij}} \frac{d_e}{w_e} \right], \tag{4.1}$$

where $\Omega_{\Pi_{ij}}$ is defined as the set of paths linking two nodes. Therefore the route assignment for each pair $i-j$ is based on a shortest-path paradigm, and does not account for congestion dynamics [41]. In practical terms, travel time of an edge is not influenced by the traffic flow on that edge. This is done under the assumption that in public transit networks, such as subways, congestion effects are less relevant (at variance with road traffic). In Section 4.4 we extend this framework by introducing the effect of traffic routing following the principles of User Equilibrium [41, 113] in the optimization process.

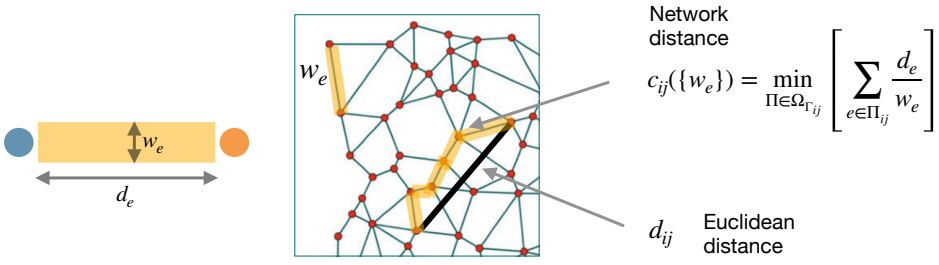

Figure 4.3: Effective temporal distance on a spatial disordered lattice. The edge's weight (depicted as width) is a proxy for the link velocity in a transit system. Here the effective network distance c_{ij} is computed as the path having the smallest temporal travel time from origin i to destination node j on the substrate. c_{ij} is visually compared to the standard euclidean distance d_{ij}. Figure adapted from [3].

Moreover, we weight the importance of a route $i-j$ via a travel probability T_{ij} between node pairs or the traffic demand between two points from a data-derived Origin-

Destination (OD) matrix. T_{ij} may represent the likelihood of an individual traveling from node i to node j or can be extended to collective human mobility flows. Here we adopt the general notation T_{ij} to refer to general "flows", and these either encode travel probabilities or actual OD traffic demands, throughout the following sections. In Section 4.5 and Section 4.6 we specify the quantity modeled. Consequently, the flow-weighted transportation cost (as a measure of efficiency) of network G given the weights assignation $\{w_e\}$ is expressed as:

$$E(G(\{w_e\})) = \frac{1}{N(N-1)} \sum_{i}^{N} \sum_{j \neq i}^{N} T_{ij} \cdot c_{ij}(\{w_e\}). \tag{4.2}$$

Here, $G(\{w_e\})$ denotes the network configuration along with the corresponding set of edge weights $\{w_e\}$. Given that paths linking different node pairs may share common edges within the network substrate, intricate topologies arise from the optimization of these edges that affect multiple routes.
We aim to find the assignation of weights $\{w_e\}$ on the edges which maximizes the efficiency (minimizes Eq. 4.2 of $E(\{w_e\})$) while requiring that the total network infrastructure cost remains constant C_G:

$$C_G = \sum_{e \in G} d_e w_e, \tag{4.3}$$

which is the sum of edge weights per unit length multiplied by edge distance. We do this using a Simulated Annealing procedure which we describe more in detail in Appendix 7.2.

4.3 Optimization of network substrates biased by simple probabilities

Before optimizing the efficiency with traffic-like ODs, we start by investigating the impact of distance in T_{ij} in Eq. 4.2. To this aim we use simple travel-probabilities that depend only on the distance between nodes: $T_{ij} \propto e^{-\beta d_{ij}}$. Here distance is used as the variable that defines the typical spatial range of flows on top of the substrate. The parameter β in Eq. 4.3 acts as a penalizing factor, determining how relevant is the distance d_{ij} between pairs $i - j$ in diminishing the probability. β can be interpreted as the inverse of a characteristic travel distance d_0 for an agent within the network, where $\beta \sim \frac{1}{d_0}$. Other distance-dependent probabilities (e.g., power-laws $T_{ij} \propto d_{ij}^{-\gamma}$) may be considered. These formulations are grounded on the empirical evidence of the power-law properties of jump length distributions in human mobility [23], which we also observe in origin-destination data discussed in Chapter 5. Here we focus on the exponential dependence due to its association with the maximum entropy derivation of gravity flows [74], which will be discussed in Section 4.5 and was introduced in Chapter 2. Future investigations may focus on the network optimization with flows having power laws dependencies. Here

instead of using T_{ij} we use the notation p_{ij} to differentiate simple flows depending on distances from more complex ones which include the nodes' attractiveness. Hence p_{ij} is:

$$p_{ij} = \frac{e^{-\beta d_{ij}/\langle d \rangle}}{\sum_{k \neq i}^{N-1} e^{-\beta d_{ik}/\langle d \rangle}}. \tag{4.4}$$

Here, $\langle d \rangle = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$ represents the average euclidean distance between points in the network and serves as a normalization factor. Therefore, p_{ij} in this simple scenario determines the range of space explored by an agent started from source node. Fig. 4.4 illustrates the spatial dependency of target probabilities. Here we first adopt the "all-or-nothing" paradigm where each OD demand is routed on the shortest path, and congestion dynamics is not present.

Experiments on HEX and ER substrates.

We perform experiments on simple generative models of both spatial and non-spatial networks. The first substrate we study here is the triangular regular lattice (HEX of size $s = 5$ in Fig. 4.4 and n Fig. 4.2 for visualization). Furthermore, we extend experiments to the case of network topologies which are not embedded in a metric space such as the Erdős-Rényi (ER) network. In this case we use the topological shortest path L_{ij} instead of d_{ij} [57] in Eq. 4.4.
Given these substrates, we generate an ensemble of optimal configurations in a wide range of β (see Fig. 4.5). We first observe that in case of $\beta = 0$ we have that the tree structure represents the optimal topology, aligning with the literature of spatial networks [14]. As we generate optimal substrates for values of $\beta > 0$, we study the emergence of loops in $G(\{w_e\})$ breaking the tree symmetry and generating complex topologies [130]. To this aim we analyze the minimum cycle basis of optimal $G(\{w_e\})$ realizations for each range of flows T_{ij}, β. A loop is defined here as a set of connected edges that defines a closed path in the network. The minimum set is the subset of loops such that any closed path in the network can be obtained via a combination of this basis. Specifically we compute the average loop size (length of the loops) in the cycle basis, and the dimension of this basis (as total number of loops present in the basis) versus β. These two metrics will allow to distinguish different topologies from the reference tree structure.

Results. Results for HEX and ER substrates are present in Fig. 4.5, while for the ER topology are shown in Appendix Fig. 7.3. When the probability to reach a target node in the network is uniform and independent of the distance (when $\beta = 0$ in Eq. 4.4) we obtain that the tree topology is the optimal structure (Fig. 4.5). This is in line with recent results [14]. As farther nodes become less likely to be travelled to from a source node ($\beta > 1$), loops in the optimal topology emerge in configurations where the effective temporal distance is optimized for close range trips. Notably, the average loop size undergoes a transition from zero to large values. In this β regime, the tree structure no longer represents the most efficient topology for peripheral leaf nodes: based on Eq. 4.4, peripheral nodes for β large enough will have most travel probabilities targeting

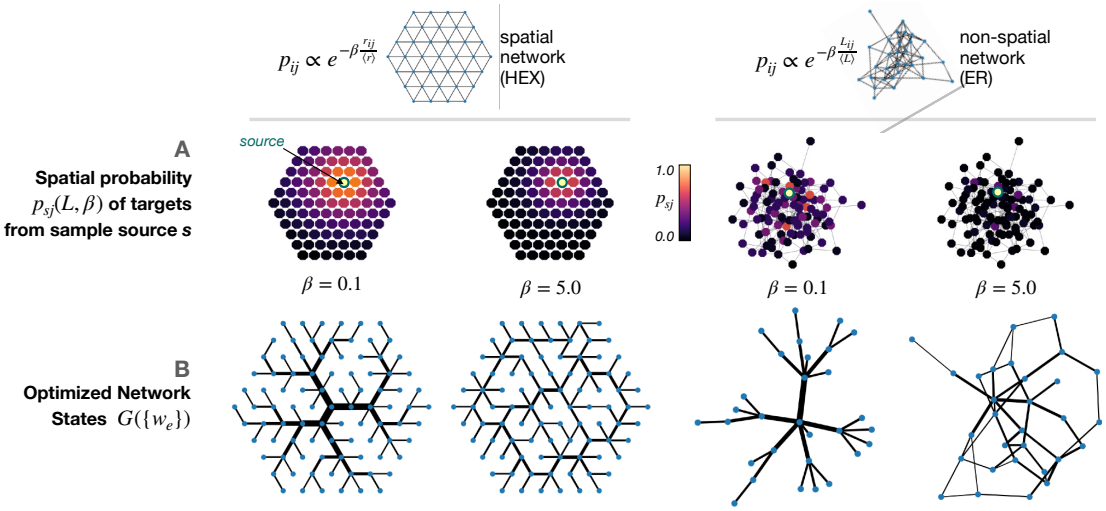

Figure 4.4: Optimization of simple synthetic networks. The role of β is studied for two substrate networks: the triangular lattice (HEX) and the non-spatial (ER) network, where distances are encoded in the topological distance L_{ij}. **A)** Heatmap of target nodes probabilities p_{ij} from the source (yellow) in the case of two β values on a HEX of size $s = 5$. As β grows, distance becomes more relevant and farther nodes are more penalized: flows tend to stay close to the source. **B)** Samples of the associated optimized network states: when flows are less affected by distances ($\beta = 0.1$) source nodes target all the other nodes in the network with approximately equal probability, and the optimal network converges to a tree-like structure. With larger β ($\beta = 5.0$), trip probabilities are more localized and loops appear as optimal structures. Figure taken from [3].

close neighbors. Consequently, as β grows, a transition point will be reached during the optimization where a connection between nodes in different branches will appear, breaking the tree symmetry. This process justifies the emergence of large scale loops in this framework (which can be seen in Fig. 4.5 panel B). Finally when $\beta >> 1.0$, the effective distance is optimized only for direct neighbors, and a lattice-like structure with many small loops is obtained. These observations are also present in the non-spatial network (Fig. 7.3).

Finally we also provide an illustrative application focusing on a scenario where a single target node is positioned at the lattice's perimeter. Through this analysis, our model effectively replicates the intricate patterns observed in leaf venation, as documented in previous studies [68]. Overall we have observed how we can transition between different topologies when distance-dependent probabilities bias the effective distances on simple substrates. As this was done following the "all-or-nothing" assignment paradigm, in the next section we investigate the effects of congestion dynamics in breaking the tree topology.

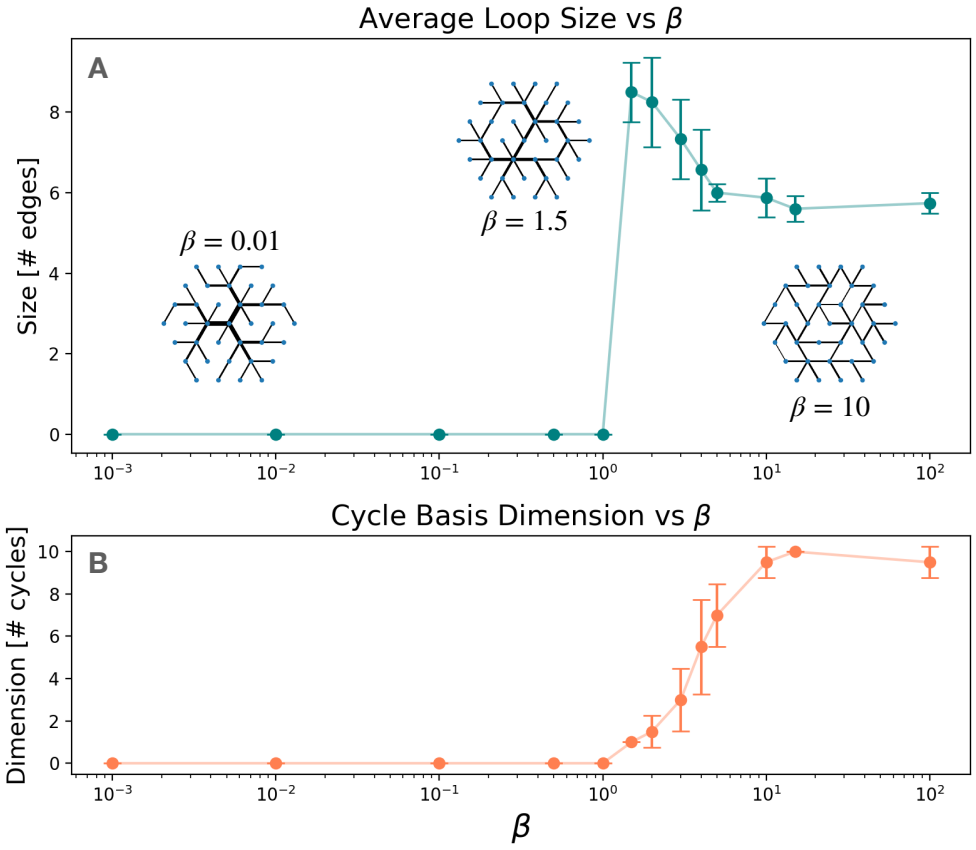

Figure 4.5: Loop Dimension vs β on HEX: We adopt the minimum cycle basis to study the presence of loops in the optimal configurations. For each β the median and its absolute deviation computed on the ensemble of optimal networks from a HEX size $s = 3$ substrate. **A)** Average number of edges (dimension) of the loops present in the cycle basis. **B)** Number of loops that constitute the minimum cycle basis. The optimal network transitions from tree structure to a lattice-like with many small loops, where a set of network realization is plotted for $\beta = 0.01, 1.5, 10.0$. Figure taken from [3].

4.4 Routing traffic under User Equilibrium: the effect on optimal topologies

Here we discuss the effect of routing the flows T_{ij} in Eq. 4.2 following conditions that in the context of urban mobility model fundamental properties of users' traveling behavior. While in previous section a flow OD T_{ij} was routed on the substrate path having the smallest cumulative travel time c_{ij}, here we route the traffic following principles of User Equilibrium [131, 41] during the generation of $G(\{w_e\})$ via simulated annealing.
Traffic Assignment (TA) consists in routing an OD demand among the set of pathways $\Omega_{\Pi_{ij}}$ in a transportation network, following some criteria. In Section 4.3 we used the "all or nothing" criterion [41]. Here instead we implement the effect of congestion, a

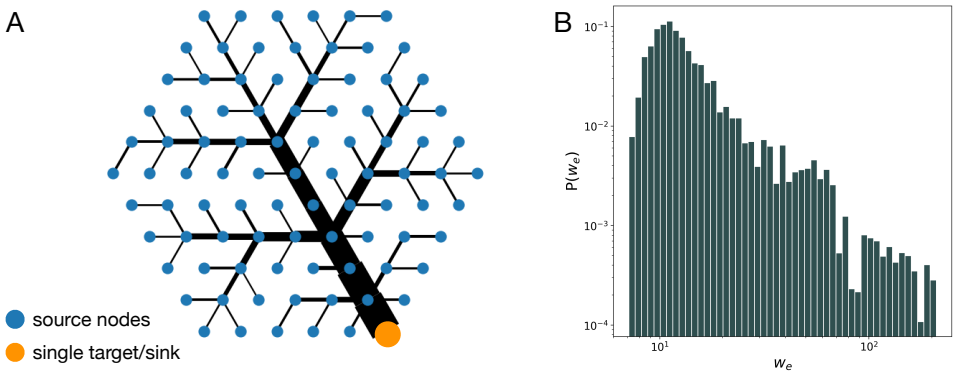

Figure 4.6: Optimal networks resembling leaves' veins patterns. Optimal state when a single sink (orange node) in the extremity of the substrate is considered. Efficiency is optimized for all nodes in space to reach the target, as source nodes. The resulting optimal state resembles tree-like patterns found in leaves (A), while the distribution of edges weights is shown in (B). Figure adapted from [3].

phenomenon typically represented through "latency functions" where the presence of traffic introduces delays in an edge's travel time, requiring traffic routing adjustments to effectively redistribute T_{ij} across other pathways and mitigate congestion effects. The Wardrop's principles [113, 41], introduce User Equilibrium (UE) and system optimum criteria that can be implemented during TA. According to the User Equilibrium principle, individual travelers select routes aiming to minimize their own travel time. Once equilibrium is reached, travelers from i to j experience identical travel times c_{ij} regardless of the route. On the other hand, the system optimum principle directs traffic routing to minimize collective total travel within the system. Hence solving traffic assignment requires optimization algorithm (such as the Frank-Wolfe or Method of Successive Averages [113]) designed to distribute this traffic respecting these conditions.

The conventional latency function to model traffic-induced temporal delays on a route is the Bureau of Public Roads (BPR) [41, 113]. This function defines a relationship between the time taken to traverse a link or edge e and the total traffic routed through e:

$$t_e = t_0 \cdot \left[1 + \alpha \left(\frac{T_e}{c_e}\right)^{\lambda}\right]. \tag{4.5}$$

Here, t_0 represents the travel time without congestion (free-flow), T_e denotes the traffic volume, and c_e stands for the capacity of the road or edge. The exponent λ governs the non-linearity in the ration between traffic and capacity, while α determines its intensity. Since we optimize edges w_e as proxies of velocities, the free-flow time on an edge in the substrate can be computed as w_e/d_e (as in Eq. 4.2). Moreover, we treat edge capacity as a constant, denoted by $c_e = 1.0$. Finally, we apply this User Equilibrium based traffic assignment during the annealing process to generate the optimal topology. Specifically, for each proposed configuration of weights $\{w_e\}$, we apply UE to route traffic along edges

(T_e) and compute $E(\{w_e\})$. We then accept $\{w_e\}$ based on the Metropolis rule [132]. In the next step we discuss in detail an analytical approach to quantify a threshold of parameters in Eq. 4.5 on the simple tree topology.

4.4.1 Emergence of alternative paths in a tree structure.

In the context of simple substrates presented in Section 4.3, we focus on the tree-like topologies (obtained in Eq. 4.4 with $\beta = 0$) these will serve as benchmark. In a tree structure, only one path between each OD pair $i-j$ is present [57]: as congestion becomes more relevant, we expect a transition from this tree structure to a new topology where additional pathways emerge to alleviate the congestion on the single path. We compute the analytical condition in which the tree topology, given the congestion parameters α, λ, represents a more efficient structure than one with additional n edges. To this aim we study the critical situation of the highest-load edge T_e in the tree (which in Eq. 4.5 is the one suffering from largest delays). We can map this to a mathematical condition where the efficiency on a single edge ($n = 1$) with velocity w_e, accommodating the traffic T_e, is compared to the efficiency (Eq. 4.2) of a set of n edges ($\mathcal{N}_e$) on which we instead distribute the traffic.

To this aim, we first show that we can perform analytical computations on the efficiency of a single edge, after the OD demands T_{ij} $\forall i, j$ has been assigned on the routes of $G(\{w_e\})$. We can decompose Eq. 4.2 as a sum of edges E_e which allows us analytical treatment. We first decompose c_{ij} as the sum of the edges travel times t_e on path Π_{ij}, therefore $c_{ij} = \sum_{e\in\Pi_{ij}} t_e$ and we have that:

$$\sum_{i,j} T_{ij}(d_{i,j}, \beta) \cdot c_{ij}(\{w_e\}) = \sum_{i,j} T_{ij}(d_{i,j}, \beta) \cdot \sum_{e\in\Pi_{ij}} t_e. \tag{4.6}$$

Now if we introduce $\mathcal{K}_{e,(i,j)} = \text{card}(\{e\} \cap \{\Pi_{ij}\})$ a set operator which is equal to 1 if the edge $e \in \Pi_{ij}$, otherwise equals to 0 (Π_{ij} is a set of unique elements) we can extend the sum on all $e \in G$. Finally by switching summation on edges we have:

$$\sum_{i,j} T_{ij} \cdot c_{ij}(\{w_e\}) = \sum_{i,j} T_{ij} \sum_{e\in G} t_e \cdot \mathcal{K}_{e,(i,j)} \tag{4.7}$$

$$= \sum_{e\in G} t_e \sum_{i,j} T_{ij} \cdot \mathcal{K}_{e,(i,j)} \tag{4.8}$$

$$= \sum_{e\in G} t_e \cdot T_e. \tag{4.9}$$

Where the term $\sum_{i,j} T_{ij} \cdot \mathcal{K}_{e,(i,j)}$ represents the total traffic routed through edge e assigned by the TA algorithm, which we define T_e. Therefore we see that if we define $E_e = T_e \cdot t_e$ after TA, we can decompose Eq. 4.2 as $\sum_{i,j} T_{ij} \cdot c_{ij} = \sum_{e\in G} E_e$.

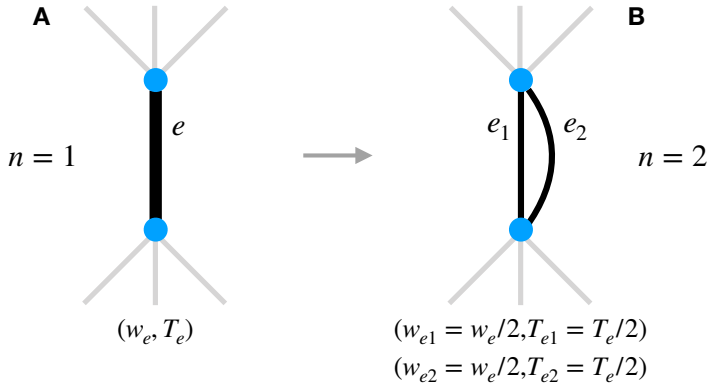

Figure 4.7: Transition from a single edge to two edges. Simple situation where the traffic T_e in the assignment process is either routed through a single edge versus a 2-pathways ($n = 2$) alternative structure. The optimality of the two scenarios is evaluated via traffic-weighted temporal efficiency after traffic assignment. We remark that the total edges weight and traffic has to be conserved between the two configurations. Figure adapted from [3].

Analytical condition: derivation of $\mathcal{F}$

From Eq. 4.9 we can focus on specific edges of interest, as the one having the maximum traffic volume in the tree structure and being subject of largest travel delays. Knowing that in case of $\alpha = 0$ (in Eq. 4.5) the TA converges to the "all-or-nothing" scenario, obtaining the tree topology (setting $\beta = 0$ in Eq. 4.4). Eq. 4.9 can be rewritten after the TA process (hence T_e is assigned on each edge and is a known scalar value) under User Equilibrium using Eq. 4.5:

$$\sum_{e \in G} t_e \cdot T_e = \sum_{e \in G} T_e \cdot \frac{d_e}{w_e} \cdot \left[1 + \alpha \left(\frac{T_e}{c_e} \right)^{\lambda} \right]. \tag{4.10}$$

We now find the condition under which a single edge is more efficient than a set $\mathcal{N}_e$ of n edges configurations under the same traffic load and total w_e availability. In Fig. 4.7 we also visually present these two configurations in which (w, T) are decomposed on a set of $n = 1$ and $n = 2$ edges. Now, we aim to determine analytically when a single edge is more efficient (lower traffic-weighted travel time E_e) than a set of n edges, when the same amount of traffic T_e and availability of edges speed w_e is conserved in the two cases. We express this via an inequality between $E_{edges}(T_e, w_e)$ of the two configurations:

$$E_e \left[T_e, w_e, (\lambda, \alpha) \right] \leq \sum_{e \in \mathcal{N}_e} E_e \left[\frac{T_e}{n}, \frac{w_e}{n}, (\lambda, \alpha) \right]. \tag{4.11}$$

We then impose that w_e and T_e are equally distributed on the n edges in the set $\mathcal{N}_e$:

$$\frac{T_e}{w_e} + \alpha \frac{T_e^{\lambda+1}}{w_e} \leq \sum_{e \in \mathcal{N}_e} \left(\frac{T_e/n}{w_e/n} + \alpha \frac{(T_e/n)^{\lambda+1}}{w_e/n} \right) \tag{4.12}$$

$$\leq n \frac{T_e}{w_e} + \alpha \frac{T_e^{\lambda+1}}{w_e} \sum_{e \in \mathcal{N}_e} \frac{1}{n^{\lambda}} \tag{4.13}$$

$$\leq n \frac{T_e}{w_e} + \alpha \frac{T_e^{\lambda+1}}{w_e} \cdot \frac{1}{n^{\lambda-1}} \tag{4.14}$$

where the summation across the n edges was carried out. Calculation of $t_0 = d_e/w_e$ was simplified with $d_e = 1$ and capacity was taken as constant $c_e = 1$. We can now re-write the inequality in a functional form of which the sign (given congestion parameters) will determine if the inequality is respected:

$$\frac{T_e}{w_e} \cdot \left[(1-n) + \alpha T_e^{\lambda} \left(1 - \frac{1}{n^{\lambda-1}} \right) \right] \leq 0 \tag{4.15}$$

$$\frac{T_e}{w_e} \cdot \mathcal{F}\left(n, \alpha, T_e, \lambda\right) \leq 0 \tag{4.16}$$

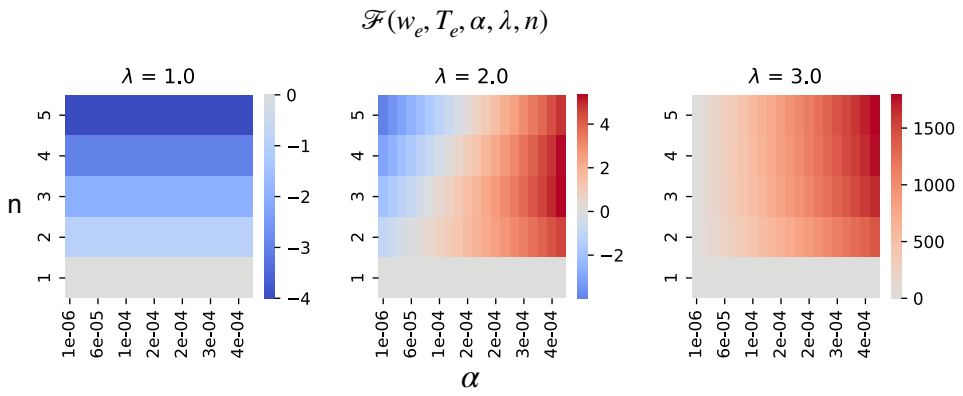

Figure 4.8: Function $\mathcal{F}$ values. $\mathcal{F}$ function $\mathcal{F}$ values, whose sign determines which of configuration is more favorable under different congestion parameters. The values of $\mathcal{F}$ identify different regimes at varying the non-linearity in the congestion term given by parameter λ. Interestingly the optimization of travel velocities (free-flow times) leads to the tree being the optimal structure in any scenario, even with $\lambda = 1$, a scenario where congestion mechanics is present, but linearly dependent on T_e. For $\lambda = 2$ we observe a transition in the optimality of n-pathways as congestion intensity α increases. Figure adapted from [3].

Hence the sign of functional $\mathcal{F}\left(n, \alpha, T_e, \lambda\right)$ given the values of parameters T_e, α, λ determines if the tree structure is more favorable to a $n > 1$ configuration with multiple routes. Traffic T_e and velocity w_e are to be considered always positive quantities. Specifically, the inequality is satisfied when $E(w_e, T_e)$ is lower for the single edge case (as flow-weighted travel time) and therefore $\mathcal{F} < 0$. Otherwise if for the set of parameters there

is a value of $n > 1$ in which $\mathcal{F} > 0$, a multi-pathways structure is therefore more favorable.

Given the OD demand and the benchmark tree structure obtained setting $\beta = 0$, we first assign the amount of traffic on the edges adjacent to the root node of the tree T_e (using $\alpha = 0$). Then for a range of congestion parameters α and λ we can evaluate the value of $\mathcal{F}$. In Fig. 4.8 we evaluate the functional $\mathcal{F}(n, \alpha, T_e, \lambda)$ by setting a value of $T_e = 170.0$, which corresponds to the largest traffic volume in the tree topology of a HEX ($s = 2$). In the next section we test the predictions based on this functional by generating optimal networks under UE and study the transition from a tree to multi-pathways topology.

4.4.2 Optimal topologies and phase transition from tree to multi-pathways structures under congestion.

In this section we investigate the properties of optimal topologies $G(\{w_e\}$ generated via traffic routing following User Equilibrium. Specifically, we test the transition predicted by the functional $\mathcal{F}$ on different triangular lattice HEX of sizes $s = 2$ and $s = 3$.
Here, to work with traffic volumes in all edges that respect the condition $T_e > 1$ and simplify analytical calculations, we impose $\sum_j T_{ij} = N$ where N is the number of OD nodes in the substrate. The number of nodes in the lattice (related to the size via the Equation in Fig. 4.2) hence determines the total traffic in Eq. 4.4 and therefore defines the maximum amount of traffic T_e that is routed on the central edge of the tree. In the tree topology, the edges connected to the root are also the ones having the largest betweenness centrality (and traffic volume when $\beta = 0$).
In each substrate we first obtain the associated tree topology $G^{tree}(\{w_e\})$ which represents the optimal solution when $\beta = 0$ and $\alpha = 0$. Consequently we can compute the value of T_e on central edges, which we then use to compute $\mathcal{F}$. Then we can generate the structures via simulated annealing for each combination of parameters, where $\alpha > 0$.
Along the set of parameters, $\mathcal{F}(n, \alpha, T_e, \lambda)$ transitions from negative to positive values, and we compute the number of additional edges N_{edges} from the tree topology that emerges in the predicted transition point. We compute N_{edges} as the number of edges of optimal network $G(\{w_e\})$ under UE minus the number of edges of $G^{tree}(\{w_e\})$. Therefore it represents a proxy observable of the additional pathways that emerge from the tree topology.

Results on size $s = 2$ and $s = 3$. Results for size $s = 2$ network are presented in Fig. 4.9 . When $\beta = 0.0$ in Eq. 4.4 the resulting maximum traffic flow is $T_e = 170.0$. In panel A the value of $\mathcal{F}(n, \alpha, T_e, \lambda)$ is plotted by separating the cases for different congestion intensity α, where the heatmaps axes are λ versus n additional paths. In panel B we present the generated optimal network topologies G. Most notably, $\lambda = 2.0$ and $\alpha = 0.0001$ represents a scenario where the value of $\mathcal{F}$ is the largest for $n = 2$ than other values of n. This predicts that in this congestion regime, N_{edges} will be different than zero but still limited to few additional edges.
This is indeed observed in the associated optimal network of Fig. 4.9B which results in a non-trivial topology where only edges close to the center appear to sustain congestion,

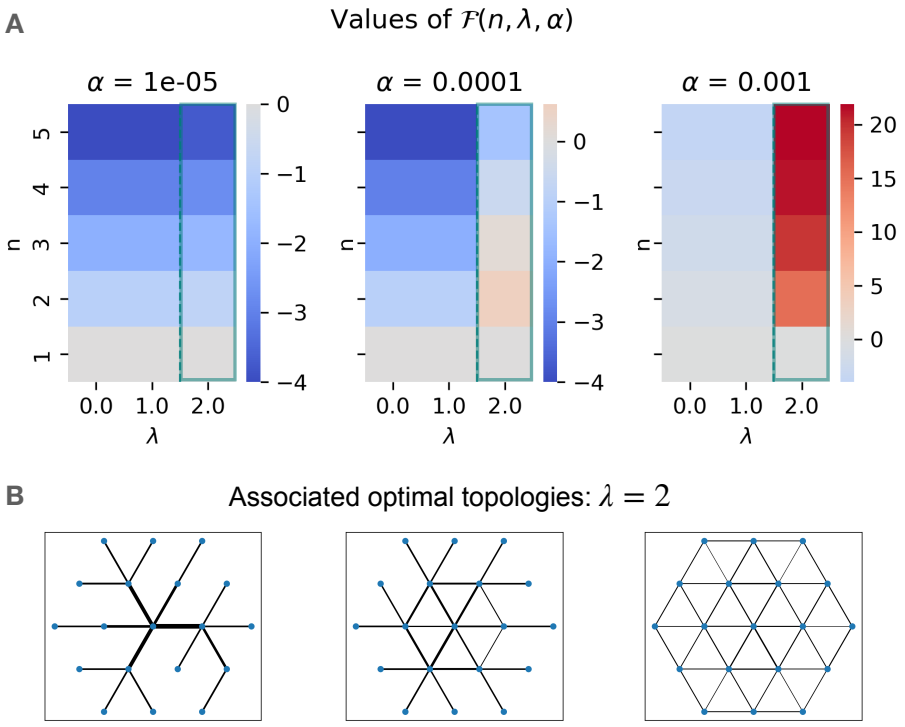

Figure 4.9: Value of $\mathcal{F}(n, \alpha, T_e, \lambda)$ and generated optimal topologies for s=2 and λ = 2. A) Heatmap of the sign of $\mathcal{F}(n, \alpha, T_e, \lambda)$ as function of n and λ are shown for three different values of α. This is to better highlight (in the teal quadrant) the different optimality of the number of additional edges n when $\lambda = 2$ for different congestion intensity α. Interestingly, the value of $\mathcal{F}$ when $\alpha = 0.0001$ for $n = 2$ is larger than $n = 3$ and in the associated topology in panel B this results in only the central edges of the lattice appearing in the optimal topology other than the tree structure (rather than all edges). This network also resembles a core paired with branches topology, which therefore can also appear in situations when $\beta = 0.0$ (tree topology) but emerge in the presence of congestion dynamics [47]. With $\lambda = 3$ the strong non-linearity forces all edges to be present, as in the third optimal topology. Figure adapted from [3].

and not farther edges. Which instead would appear if $\mathcal{F}(n > 2) > \mathcal{F}(n = 2)$. Remarkably, these particular patterns can be therefore predicted just from the sign of functional $\mathcal{F}$, despite the strong approximations and the simple scenario from which $\mathcal{F}$ was derived.

In Fig. 4.10 we report the results on the generated topologies of $N_{edges} = N_{edges}(G) - N_{edges}(G^{tree})$ as a function of the parameters in Eq. 4.5. These parameters are present in the x-axis as re-scaled control parameter αT_e^{λ}. We compare N_{edges} versus the curve of the functional $\mathcal{F}$ for different values of n and different sizes s and parameter λ. We observe the transition from the tree-like structure to a multi-pathways structure when $\mathcal{F}(n > 1) > 0$ (in Fig. 4.10 computed with $n = 1.5$), ultimately converging to a complete lattice when $\alpha T_e^{\lambda} >> 1.0$. $\mathcal{F}$ approximates the transition at $\alpha T_e^{\lambda} = 1.0$ for different lattice sizes and a set of congestion parameters which span several orders of magnitudes.

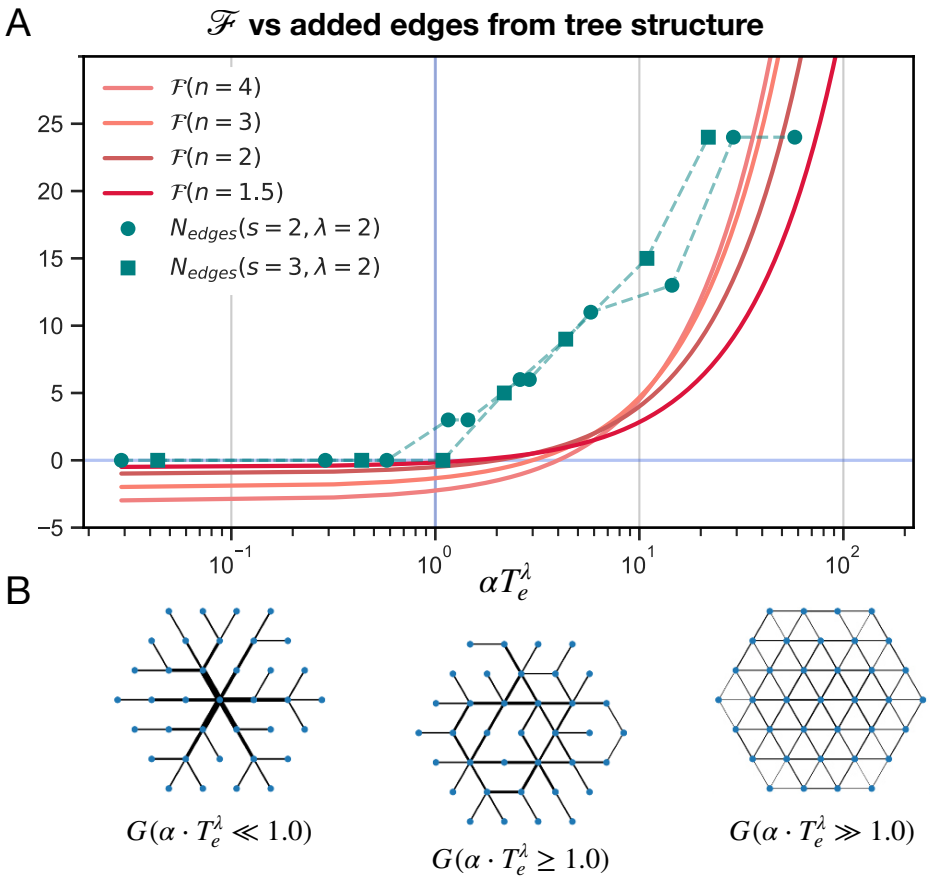

Figure 4.10: Function F values. A) Plot of the emergence of additional pathways encoded in n from the tree topology G^{tree} versus the value of $\mathcal{F}$. We plot $\mathcal{F}(\lambda = 2)$ for different n and HEX sizes s and highlight the threshold $\alpha T_e^{\lambda} = 1.0$ after which $\mathcal{F}(\lambda = 2, n > 1) > 0$. Therefore these curves predict the emergence of additional edges from the tree structure, computed as N_{edges}. N_{edges}, computed on generated optimal topologies, transitions to non-zero values after the threshold $\alpha T_e^{\lambda} = 1.0$. **B)** Associated topologies in the case of $s = 3$, $\lambda = 2$. Specifically for $\alpha T_e^{\lambda} \ll 1.0, \geq 1.0, \gg 1.0$. We can observe the situation in which the tree structure transitions to a complete lattice. This is a consequence of $\mathcal{F}(n, \alpha, T_e, \lambda)$ that for large non-linearity λ and intensity of congestion α favors large n and therefore the routing of traffic in as many pathways as possible in the lattice. Figure taken from [3].

4.5 Adding node attractiveness W_j

In this section we study the addition of nodes features in the lattice substrate. Human mobility in urban systems is characterized by traffic flows [125] towards areas of interest, with high density of commercial and/or business activities and points of interests (POIs). Here we extend this framework to the case where we have flows which reproduce typical patterns of human mobility in urban contexts, by adding the presence of nodes with high attractiveness (POIs density W_j) that bias these flows. Spatial-interaction models can

be adopted to this aim as discussed in Chapter 2. Moreover, the validity of gravity-like models in urban contexts has been further studied by recent works [81]. Here a gravity-like equation was derived as the best closed-form in a Bayesian model fitting, using density of POIs as an additional measure of attractiveness [81]. More specifically, in a spatial interaction model [133, 74] nodes with attractiveness W_j compete as possible destinations of population nodes P_i, and here the distances d_{ij} are used as a cost of traveling between $i-j$ and determine the spatial dependency of the flows. Although we remark that a formulation with power-laws represents also a valuable venue to investigate. This leads to the equation:

$$T_{ij} \propto \frac{1}{Z} P_i W_j e^{-\beta d_{ij}}, \tag{4.17}$$

where Z is the normalization factor $Z = \sum_k W_k e^{-\beta d_{ik}}$ on the set of all target nodes. P_i represents the population in node i and W_j represents the attractiveness of node j [133] quantified in Section 4.6 using the number of POIs from OSM [134]. As a result, we can see T_{ij} as the fraction of population in node i travelling to node j. While in Section 4.6 we study the case of London as a more realistic distribution of population, we first start in this section by only modeling the attractiveness and working with the simple assumption of nodes having all the same population: $P_i = 1.0\ \forall i$ to better understand the role of W_j in generating $G(\{w_e\})$. We apply this modeling on the HEX lattice $s = 4$ to explore the optimal features that emerge when T_{ij} are biased towards specific areas/nodes with high attractiveness.

3-Points polycentric structure W_j.

We examine a spatial arrangement where three nodes (POIs) at the vertices of an equilateral triangle, are assigned high values of W_j, similar to a Steiner tree problem. This 3-points configuration mirrors a typical polycentric arrangement of city centers. Moreover, it is associated to the Euclidean Steiner Tree problem [135, 136], which represents a class of problems where, given a set of N points in a plane, the objective is to connect these points with the minimum total length. In this 3-points arrangement the central node in the lattice would represent the Fermat point [136], and the Steiner node, which links the three vertices of the high W_j triangle, as depicted in Fig. 4.11 panel A. In Fig. 4.11A we observe the influence of W_j in polarizing the traffic flows, where $P(T_{ij})$ can ideally be divided into close-range trips and longer-range flows towards POIs. Fig. 4.11 illustrates optimal solutions corresponding to $\beta = 0.1$ and 4.0. In both cases, a star with three branches from the central node to the high W_j nodes represents the solution of the Steiner tree problem.
When $\beta = 4.0$ nodes with large W_j are penalized by the distance, and flows are more localized with $P(T_{ij})$ being non-negligible especially between non-POIs nodes. In this scenario we observe the emergence of large-scale loops. Additionally, due to the heterogeneity in T_{ij} we observe a second mode in the distribution of optimal speeds $P(w_e)$ (refer to Fig. 4.11), which can also be interpreted as the optimization process offering as a solution a network having two different speed layers, similar to a multi-layer transportation network.

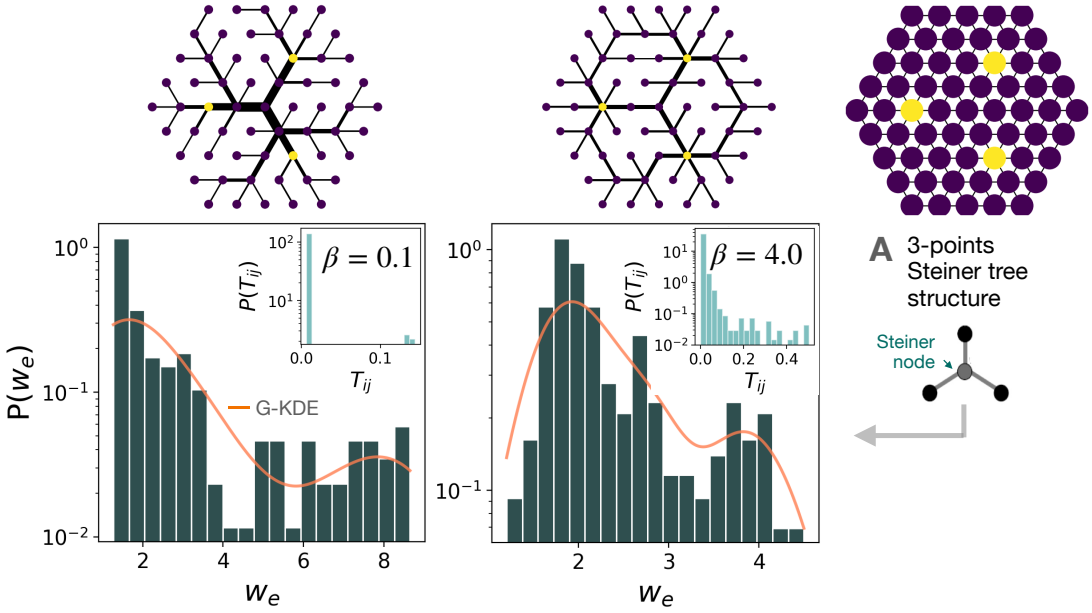

Figure 4.11: Simple models of urban attractiveness distributions under study (3-Points): Spatial distribution of attractiveness W_j as number of POIs in a node, mapped with color intensity (yellow higher). Appearance of bi-modality in the optimized velocity distributions $P(\{w_e\})$ that points to the multi-layered structure of the optimal transportation networks when close-range flows are paired with long-range traffic towards city centers (Insets $P(T_{ij})$ with peaks biased by POIs). Gaussian KDE is shown in orange as a visual aid. 3-points polycentric distribution of POIs, resembling the Steiner Tree problem [135, 136] in the 3-points configurations. The cases of $\beta = 0.1$ and $\beta = 4.0$ as exemplary cases, where branches connecting the POIs are also paired with loops when the number of close range flows ($\beta = 4.0$) is more relevant. Figure adapted from [3].

Therefore we can map this process to one in which the optimal structure, to accommodate heterogeneous flows with a fixed amount (C_G) of resources w_e, is mapped to a topology with multi-layer speeds. In this topology, fast connections emerge as optimal to provide fast travel time to long-range or commuting trajectories, while a lower speed layer for shorter-range travels with a lower amount of traffic.

4.6 Case study: Greater London Area

We finally show an extension of this framework to a realistic case in which we integrate data from a real urban structure. Specifically, we model the Greater London Area (GLA) urban features on top of the triangular lattice substrate and optimize Eq. 4.2 given the OD demand in Eq. 4.17. The aim of this analysis is to understand if a the optimization of a travel time effective distance on the spatial substrate biased by realistic urban collective flows are sufficient conditions to reproduce similar topological features as the London

subway system (such as a central core paired with peripheral branches [115]).

4.6.1 Datasets and spatial interaction model

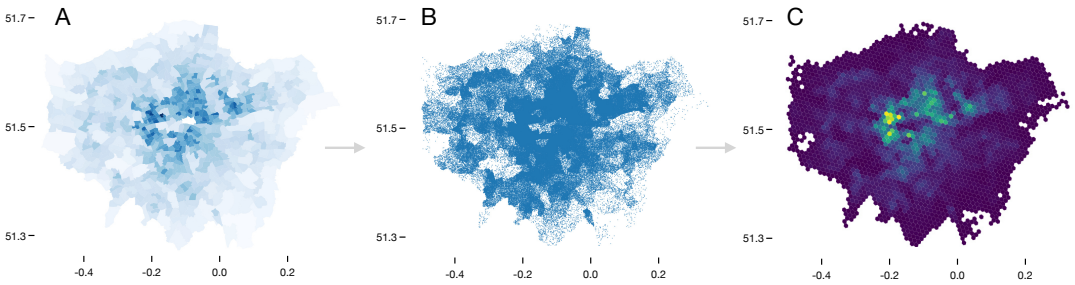

Figure 4.12: Census data retrieval for Greater London Area. A) London's wards data from Census 2014 (https://data.london.gov.uk/dataset/ward-profiles-and-atlas). **B)** Points are sampled in each Ward's area with a density proportional to Census population. **C)** The number of points are mapped to HEX3 tiles as P_i. Finally, a restriction to a disc is used to enforce symmetry in the distribution of nodes and to ease the computational load. Figure adapted from [3].

We first briefly discuss the datasets employed as P_i and W_j. We recover the distribution of amenities [137] from OpenStreetMap (OSM) [134] as a set of geo-taggeed points of interests (POIs) which we can use as proxy to estimate the attractiveness W_j of a urban area. We report the dictionary of amenities extracted from OSM in Appendix 7.3.7, following the same procedure employed in Chapter 5. Census data for Greater London Area wards from 2014 is used as urban feature P_i. Finally, to uniquely define the areas and obtain the spatial discretization in hexagonal tiles i we use Uber's H3 tiling [138] of GLA at resolution 8 [138]. OSM points W_j and Census population data are therefore mapped as densities on these H3 tiles areas. We restrict the GLA area to a disc centered in London City coordinates, with a radius of 14 Km, allowing us to work with an isotropic spatial distribution of nodes. The process of mapping the Census wards densities and point-based data to H3 tiles is presented in Fig. 4.12. With this information we can simulate the spatial interaction OD flow T_{ij} as defined in Eq. 4.17. In Fig. 4.13 we provide a view of a sample OD matrix that emerges from the Eq. 4.17. Moreover, we impose an upper limit on edge weight w_e this also translates in a distribution during the optimization process: $w_e \in (0, w^*)$, where $w^* = 7.0$. This constraint better simulates the upper bound in speed of real multilayer systems and allows to better recover the bi-modality of real multi-layer networks.

In Fig. 4.13C we observe that the velocity $\{w_e\}$ distribution displays a bi-modal shape, with a a fraction of edges being concentrated close to the upper bound w^* where most of the traffic is routed. This aggregation allows to limit the analysis of the optimal network in the sub-graph defined by the set of edges in this second mode (where $w_e > w_{threshold}$, highlighted in Fig. 4.13C). In Fig. 4.13, panel C, we present a visual representation

of this sub-graph for $\beta = 0.35$. We can observe that the optimal configuration of the network constitutes of a central core paired with peripheral branches.

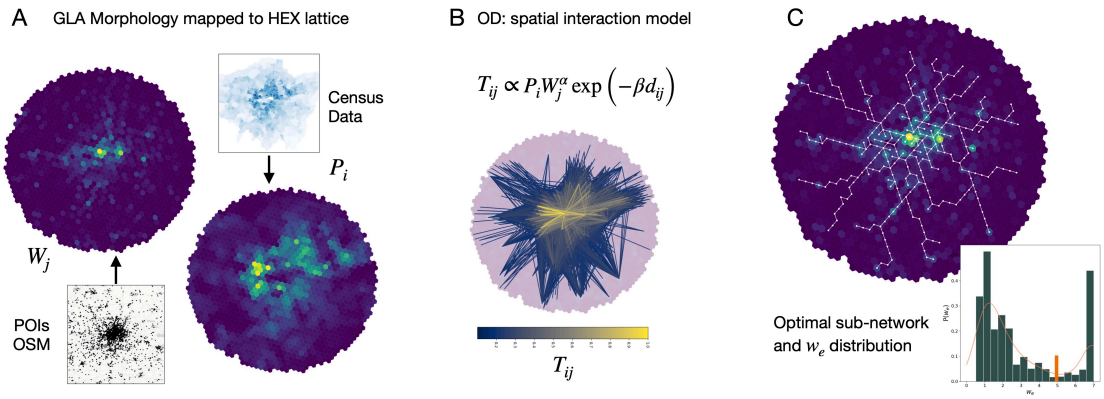

Figure 4.13: Optimal network model for Greater London Area subway network: Application of the efficiency optimization with traffic flows simulating collective human mobility on the urban structure of the Greater London Area. **A)** Population data from Census and POIs densities from OSM are mapped to the H3 tiling [138]. **B)** Data on H3 tiles are mapped to the triangular lattice, with nodes having features that allow simulation of traffic-like flows. A visual representation of the OD matrix is shown where T_{ij}, computed with $\beta = 0.35$. **C)** Optimal network state for the London model, where only fast edges having weight larger than a threshold ($w_e > 5$) are kept to isolate the sub-graph constituted by a high velocity set of edges, such as a subway system, are shown. Core network structure with loops paired with peripheral branches can be visually seen. Figure adapted from [3].

Comparison with London Subway network: Scaling and statistical properties

The model's subgraph of high speed edges is compared the Greater London Area [43] subway system to assess statistical similarities between the optimized substrate and the real tube system. We obtain data of the real tube network in the GLA [43] to quantify the similarities between the optimal structure emerging from our parsimonious framework and the real subway system. We exploit known spatial scaling laws [115] to quantify this similarity, these are convenient to highlight the presence of the central core structure with dense loops, paired with quasi mono-dimensional lines branching from it.
We analyze the spatial arrangement of stations/nodes using a profile function $N(r)$ which represents the number of stations at a distance r from the network barycenter. The barycenter is computed as the average position of all station nodes (in our subgraph $\bar{G}$ we can consider the nodes as "stations")[115]. We perform this scaling analysis for both the real network and our sub-graph simulated and the results are presented in Fig. 4.14. We observe the two characteristics scaling regimes [115] which indicate the core plus branches topological separation. First the scaling of r^2 of the spatial distribution of nodes of the core subgraph up to a radius r_C (radius of the core structure). Then a second trend $r > r_C$ due to mono-dimensional branches can be observed. This curve can be approximated analytically via an integral curve for $N(r > r_C)$ which itself can be

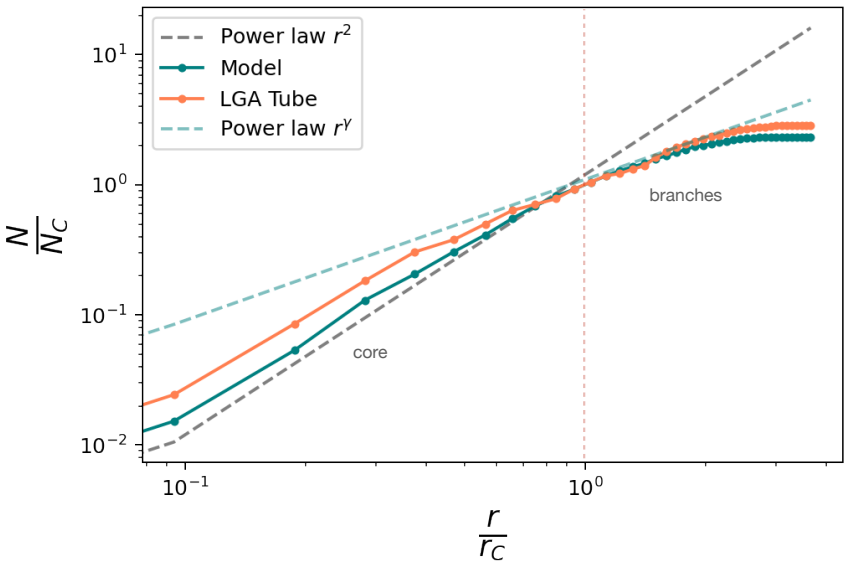

Figure 4.14: Scaling properties of GLA subway network: Profile of the number of stations (nodes in the optimal subgraph network) against distance from the barycenter. $N(r)$ profile scaling properties are compared between the simulated and the real network system. Scaling properties predicted in [115] are verified, finding the two different scaling regimes separated at approximately $\frac{r}{r_C} \sim 1$ for core paired with branches systems. Where r_C is the radius of the core structure, in which $N(r)$ is characterized by r^2 scaling, and N_C is the number of stations inside the core $r < r_C$. The scaling exponent $\gamma = 1.25 \pm 0.02$ is obtained as a linear fit of the integral curve [115] for $r > r_C$. Figure taken from [3].

further approximated by a power law r^γ ($\gamma = 1.25 \pm 0.02$), as in Ref. [115]. The curve of $N(r)$ is consistent with the real network and confirms scaling laws prediction from [115].

4.7 Conclusion

Transportation networks mediate human mobility in urban areas, exhibiting complex spatial topologies that are crucial for efficient mobility. In this chapter, we have developed a simple framework to generate networks in urban areas, focusing on the emergence of optimal topological patterns. By exploring the interplay between spatial traffic flow probabilities, travel costs, and congestion dynamics, we have uncovered insights into the optimal features of efficient transportation structures.

Our study has revealed that by weighting the metric used to evaluate the temporal efficiency via simple flow probabilities influenced by distance, we can induce transitions from tree-like to lattice-like network structures. Moreover, incorporating travelers' behaviors, such as User Equilibrium criteria in traffic routing and congestion dynamics, has led to the emergence of additional edges from a tree reference structure. These edges allow routing traffic on alternative paths which are essential for mitigating congestion-induced delays, and we provide an analytical derivation to predict the intensity of congestion that drives this transition. Additionally, by modeling flows resembling collective human mobility patterns, we have generated network features that facilitate flows between major Origin-Destination pairs, reminiscent of real-world transportation systems. In applying our model to the Greater London Area, we have observed scaling similarities with the London Underground network. Therefore, with the addition of traffic flows that simulate the fundamental patterns of collective mobility in urban areas (with attractiveness modeled as POIs), we were able to recover non-trivial network topologies typical of efficient public transit systems. In the next chapter, we will instead focus on the statistical properties of collective mobility from a data-driven perspective. We will further establish its relationship with the presence of POIs. Specifically, we leverage its predictability to inform individual next-location prediction models in the context of out-of-routine mobility.

Chapter 5:
Mixing individual and collective behaviors to predict out-of-routine mobility.

5.0.1 Chapter Overview

Across diverse domains, human mobility emerges as a critical factor and driving process. Traffic forecasting, crime prevention strategies, disease propagation models or social segregation dynamics represent examples of areas where accurate predictions of individuals' future whereabouts can be essential for an effective modeling. In recent years sophisticated prediction models ranging from pattern-based to deep-learning approaches have been developed, although they often struggle to generalize beyond recurrent mobility patterns. Particularly they are limited when dealing with non-routinary behaviors, which are crucial for addressing critical situations and large-scale collective events such as epidemics of natural disasters. This chapter studies a Markov-based model designed to bridge this gap by dynamically integrating both individual and collective mobility behaviors, leveraging the predictability of individual movements. This chapter is based on the work we did in Ref. [4].

In Section 5.2 we introduce the Markov model and the process of mixing the individual and collective behaviors via an individual's entropy. We then introduce the dataset of trajectories of millions of individuals in Boston, Seattle and New York City over an eight-month period in 2020. On this dataset we provide a comprehensive evaluation the model's performance and compare it against models relying solely on individual or collective data. Moreover, we focus on the test overlap problem, presenting the concept of LCST. We show how the proposed model demonstrates considerable generalization capabilities, particularly in out-of-routine mobility behaviors, surpassing even sophisticated deep learning methods. Then in Section 5.3 we perform a spatial analysis which highlights that out-of-routing mobility is effectively captured by collective behaviors particularly in areas with high attractiveness. In this section we also investigate in detail the properties of collective mobility that justify these results, analyzing the statistical patterns of travel

distances and entropies of collective mobility in proximity of POIs. We also observe that collective behaviors remain predictive in these areas even when the accessible information is the same of other less populated areas, highlighting a behavioral aspect. In Section 5.4 we show that the model exhibits resilience to disruptions in recurrent mobility patterns, such as those experienced during the COVID-19 pandemic, due to its dynamic integration of individual and collective behaviors. Overall, this chapter underscores the significance of integrating individual and collective mobility information for enhancing next-location prediction accuracy, offering insights into addressing challenges associated with non-routinary behaviors and disruptive events in mobility forecasting.

5.1 Introduction

Understanding and modeling the dynamics of individual mobility is a relevant task in many societal challenges [19]. Especially in the context of urban systems, both individual and collective mobility represent a fundamental dimension in the design of sustainable and livable cities [139, 140], in avoiding traffic congestion [141, 142], and fundamental to epidemics spread mitigation and public health monitoring as also shown in Chapter 3 [143, 144, 145, 146].
Thanks to the large availability of extensive mobility data and the development of advanced statistical techniques, the prediction of individuals' future whereabouts has seen continuous growth in the last years. In particular the task next-location prediction [147, 148, 32]. While deep learning solutions have obtained substantial attention and dedicated research due to the capacity to model complex patterns from extensive datasets [149, 150, 151], they often lack interpretability and/or generalization power, acting as black boxes which are unable to map these patterns into explicit mechanistic and generalizable rules [19, 147]. Conversely, simple models such as the mobility Markov chains are intrinsically explainable and allow the analysis of the mechanisms behind the predictions. This although comes at the cost of a lower accuracy in forecasting future movements [152, 153, 154]. Both deep learning and Markov models are mainly based on individual mobility trajectories. Individual trajectories are characterized by regular patterns such as visits to previously visited locations at regular times, hence rendering them inherently predictable [31, 77, 155, 23]. However, in some scenarios, individuals may have a marked preference for exploring new destinations [77] or be forced to alter their routine due to external driving factors such as job loss, natural disasters, or epidemics [156, 157, 158, 146].
The mobility patterns and individual trajectories that emerge in these contexts are defined as out-of-routine mobility, and represent a challenge for statistical models. In fact, these models are designed to capture regular patterns in individual trajectories, and hence they often memorize these regular patterns from data rather than learning generalised mobility behaviours [147, 159]. A large body of literature on human behaviour across various contexts, such as social networks [160], financial networks [161], voting and political polarization [162, 163, 164], indicates that an individual's decisions can be influenced or intrinsically similar to the behavioral patterns of the group they are exposed to [165]. This suggests that information about collective mobility behaviours also holds predictive power for individual behaviours, offering insights into out-of-routine mobility. However, the potential of combining individual and collective behaviour to enhance human mobility prediction also represents an unexplored field of interest. In this chapter we tackle this challenge by mixing individual transition probabilities with collective behaviors in three cities, exploiting the definition of an individual predictability from a location to weight these two probabilities. This (i) allows us to generalize and accurately predict out-of-routine behaviors, as quantified via test overlaps. And (ii) by studying the spatial statistical properties of collective mobility we uncover the spatial correlation with the presence of attractive areas in the urban space, highlighting a dependency of individual movements to follow collective mobility patterns in these areas.

5.2 Mixing individual and collective Markov models

In this section we formalize individual trajectories as set of spatio-temporal points, defining the task of next-location prediction and the Mobility Markov Chain (MMC) models. We present the Cuebiq dataset of individual trajectories in the cities of Boston, New York and Seattle, on which we test the generalization capabilities of these models when the test trajectories are characterized by different amount of novel transitions never before seen in training.

Individual trajectories and Origin-Destination matrices

A spatio-temporal point is defined by a pair $p = (i, t)$, where i represents a location (defined by geographical coordinates) and t the timestamp of the point. We define a trajectory, of length n, $P = \{p_1, p_2, \ldots, p_n\}$ as a daily time-ordered sequence of n spatio-temporal points. Each individual user u in our datasets has a set of N historical trajectories $\mathcal{H}^{(u)} = \{P_1, \ldots, P_N\}$ from which we compute $I_i^{(u)}$, representing the set of transition probabilities of user u starting from location i (see Figure 5.1B). Trajectories of all individuals are then aggregated to obtain the collective origin-destination matrix and then compute C, where C_i represents the probability distribution of all transitions made by any individual starting from i (see Figure 5.1A). Specifically, for each user u in the datasets we compute the individual origin-destination matrix which captures the transition probabilities, $T_{ij}^{(u)}/T_i^{(u)}$, between each pair of locations visited by u. Where $T_{ij}^{(u)}$ is the total number of transitions in $\mathcal{H}^{(u)}$ from location i to location j, and $T_i^{(u)} = \sum_{j \in L} T_{ij}^{(u)}$ accounts for the total transitions from location i to any other location in the set $L^{(u)}$ of locations visited by u [78, 166].
Given a trajectory $P \in \mathcal{H}^{(u)}$ of an individual, a next-location task is the problem of predicting the next point $p_{n+1} \in P$ [147, 159]. In the literature, $I_i^{(u)}$ and C_i are typically used separately as MMC [159] models for next-location predictions.

Individual-Collective Model

In this chapter we introduce a novel model $M_i^{(u)}$ that dynamically combines $I_i^{(u)}$ and C_i information based on the predictability of user u from origin i. We remark that in this Chapter, at variance with other works [167], we use the term predictability referring to the information of historical movements of the individual user from a specific location i, and not the predictability defined by the previous history of a trajectory. When the next location is highly predictable based on the historical information $\mathcal{H}^{(u)}$, the model relies more on individual information in $I_i^{(u)}$ for the prediction. Conversely, more collective information in C_i is embedded in the model probabilities from location i. We quantify here the predictability of u's next location from i via a normalised Shannon's entropy,

computed from the probability $I_i^{(u)}$ [31, 168]:

$$S_i^{(u)} = -\frac{\sum_{k \in L^{(u)}} I_{i,k}^{(u)} \cdot \log\left(I_{i,k}^{(u)}\right)}{\log |L^{(u)}|}, \tag{5.1}$$

where $L^{(u)}$ is the set of distinct locations visited by u, while $\log |L^{(u)}|$ is a normalisation factor so that $S_i^{(u)} \in [0, 1]$, and $I_{i,k}^{(u)}$ is the probability learned from $\mathcal{H}^{(u)}$ of individual u moving from an origin location i to a destination location k. We then use $S_i^{(u)}$ to combine probabilities from $I_i^{(u)}$ and C_i as follows:

$$M_i^{(u)} = (1 - S_i^{(u)}) I_i^{(u)} + S_i^{(u)} C_i. \tag{5.2}$$

Where $1 - S_i^{(u)}$ is therefore the confidence of model $M_i^{(u)}$ in relying on individual information. To derive Markov transition probabilities, we normalise $M_i^{(u)}, \forall i$ using the softmax function:

$$\text{softmax}(l)_i = \frac{e^{l_i}}{\sum_{j=1}^{n} e^{l_j}}, \tag{5.3}$$

where $l_1, l_2, \ldots l_n$ are the transition probabilities of $M_i^{(u)}$ in Equation 5.2. In instances where location i is not represented as an origin location in $\mathcal{H}^{(u)}$, the probability distribution $I_i^{(u)}$ is empty, and we set $S_i^{(u)} = 1$ to indicate maximum uncertainty of the model. In such instances, no historical trajectories data is available for location i, and the model prediction relies solely on collective information in C_i. Figure 5.1C-E illustrates the combination of collective and individual information in $M_i^{(u)}$.

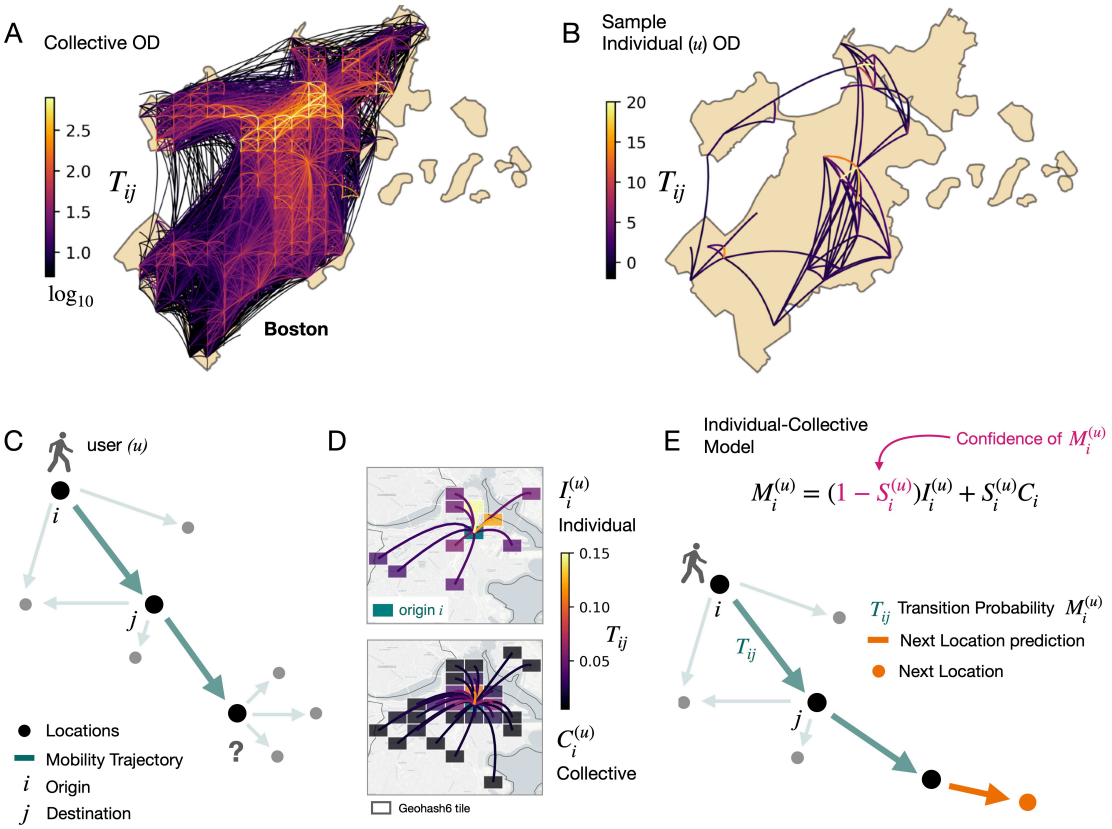

Figure 5.1: Interplay of individual and collective mobility in next location prediction. **(A)** The collective Origin-Destination (OD) matrix of aggregated flows computed for the city of Boston using GPS trajectories. **(B)** An individual origin-destination matrix for a synthetic individual u. **(C)** A mock individual trajectory for a user u starting from location i. We aim to predict u's next visited location. **(D)** The set of u's trajectories in the training set $\mathcal{H}^{(u)}$ (panel B) is used to compute the transition probabilities $I_i^{(u)}$ from location i. C_i represents the probability distribution of collective flows from location i, generated from the collective OD flows (panel A). Destinations' tiles are coloured based on the probability T_{ij} of visiting them from origin i. **(E)** Model $M_i^{(u)}$'s prediction of individual u's next location is devised by combining $I_i^{(u)}$ and C_i, based on the predictability of u from location i, computed as a normalised Shannon entropy $S_i^{(u)}$ from $\mathcal{H}^{(u)}$. Figure taken from [4]. Maps: Stamen Maps, Icons: Fontawesome.

5.2.1 Trajectories datasets: Boston, Seattle and NYC

We now introduce the Cuebiq dataset [146] used to perform experiments and list the pre-processing procedures. We adopt here notations I, C, and M to refer to the individual, collective and the model mixing the two information.

Dataset and Pre-processing information.

To obtain I and C for real individuals and to test our models, we use privacy-enhanced GPS trajectories provided by Cuebiq as part of the Data for Good COVID-19 Collaborative program. This dataset uses privacy-enhanced GPS locations spanning nine months (from January to August, 2020) in the cities of New York City, Seattle, and Boston in the United States. The dataset originated from approximately 2.5 million users who willingly opted to share their information anonymously for research purposes, adhering to the guidelines of the CCPA (California Consumer Privacy Act) compliant framework. Moreover, sensitive points of interest are removed, and inferred home locations are obfuscated to the Census Block Group level by the data provider. Individual user stop-locations and trajectories are extracted from the dataset through the following procedure. Initially, each temporal sequence of GPS coordinates within a 65-meter radius is identified, where a user stayed for a minimum of 5 minutes [169]. Subsequently, we apply the DBSCAN algorithm [170] to identify dense clusters of points within a distance of $\epsilon = \Delta_s - 5$. We define these dense clusters as stop locations. For an extended description of this procedure of the GPS data processing, refer to [146].

We perform a tessellation of the cities into GeoHashes at level 6 (locations) which refer to tiles of 1.2 Km $\times$ 609.4 m. These tiles map the locations of user stops and of users' trajectories. To address potential under- and over-representation issues, we filter out trajectories with less than four points ($|P^{(u)}| < 4$), and users with less than two trajectories ($|\mathcal{H}^{(u)}| < 2$), and remove the top 95-th percentile of the most represented users. We remark that our analysis uses trajectories consisting of at least four spatio-temporal points, defined at Geohash level 6 for spatial resolution, and all temporal points occurring within the same day. Figure 5.1A-B shows samples of I and C (OD) from the trajectories of GPS Boston dataset.

Accuracy on full test

We first train and test our model in a period of three months (from January 3rd to March 1st, 2020), prior to the COVID-19 pandemic, which introduces shifts in mobility patterns which will be modeled and discussed in Section 5.4.
For each individual u, we allocate 80% of their least recent trajectories for model training, while the 20% most recent trajectories form the test set (which we refer to as "Full test"). This chronological-based train-test split can be understood as a scenario in which the movements of an hypothetical individual u are observed for 2 months (as an example), and then his future trajectories are then predicted using these observations for the following 2 weeks. Both recurrent and novel mobility patterns will characterize this prediction phase. In the training phase, for each location i and user u, we compute the distributions $I_i^{(u)}$, C_i, and $S_i^{(u)}$.
Subsequently, we perform next-location (NL) predictions on the test set and assess the models' accuracy using the top-5 accuracy metric (ACC@5). ACC@5 computes the percentage of transitions (from location i to location j) where the correct next location

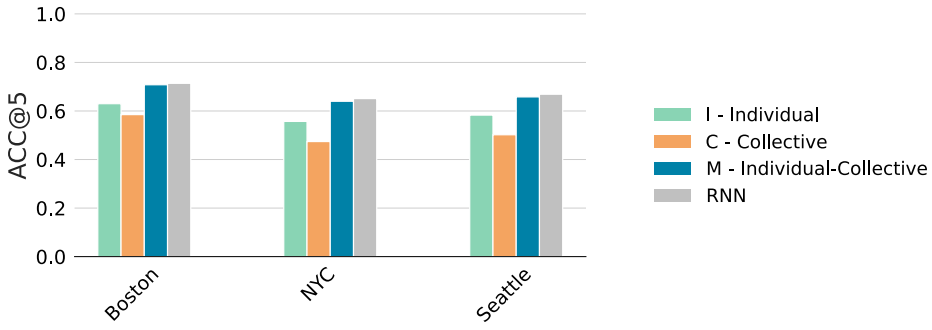

Figure 5.2: Top-5 accuracies (ACC@5) of the models: Accuracies (ACC@5) for Boston, New York City and Seattle using Individual (I), Collective (C), and Mixed (M) models. Full test set ACC@5, where M shows comparable results to the RNN. Figure adapted from [4].

is among the top five predicted destinations, and in NL tasks represents a common metric [159, 147]. We also compare I, C, and M with a deep learning Recurrent Neural Network-based baseline RNN (see Appendix 7.3.5 for details). Table 5.1 and Figure 5.4A show that on the full set, M obtains higher accuracies than I and C (eg +15% in NYC versus I), and comparable performances to RNNs [171, 147]. Despite relying on a large number of parameters, RNNs exhibit only marginal relative improvements over M: for example +2% in NYC.

5.2.2 Models generalisation capability

Human mobility can also be characterized by recurrent patterns, which may result in significant similarities among an individual's trajectories [23, 78, 159]. As a consequence, in test set a significant portion of trajectories from the training set may be present. Considering trajectory overlap between the training and test sets is a relevant step to evaluate of a model's generalisation capability [159]. We quantify trajectory train-test overlap with the Longest Common Sub-Trajectory (LCST) [159], which evaluates shared sub-sequences (accounting for order and frequency of visits) between two trajectories.

LCST. Specifically, the overlap between two trajectories of an individual user u , $P^{(u)} = \{p_1, p_2, \ldots, p_n\}$ and $R^{(u)} = \{r_1, r_2, \ldots, r_m\}$ can be defined as follows. We introduce the prefix P_i of $P^{(u)}$ as the list of the first i-th locations in $P^{(u)}$, i.e., $P_i = \{p_1, \ldots, p_i\}$ (dropping index $^{(u)}$ for simplicity). This definition extends similarly to R trajectory. The size of the LCST for two prefixes P_i and R_j is defined as follows:

$$LCST(P_i, R_j) = \begin{cases} 0, & \text{if } i = 0 \text{ or } j = 0 \\ LCST(P_{i-1}, R_{j-1}) + 1, & \text{if } i, j > 0 \text{ and } p_i = r_j \\ \max(LCST(P_{i-1}, R_j), LCST(P_i, R_{j-1})), & \text{if } i, j > 0 \text{ and } p_i \neq r_j. \end{cases} \quad (5.4)$$

The overlap between a test trajectory R and the training set is quantified as the maximum LCST over all the trajectories in the training set:

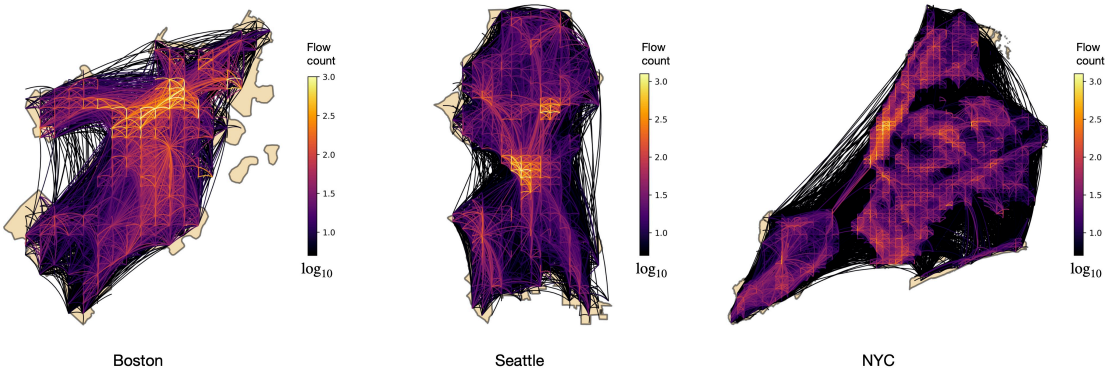

Figure 5.3: Collective ODs: Collective T_{ij} flows between Geo Hashes tiles from the Cuebiq dataset in the cities of Boston, Seattle and New York City. The stop-location detection described in Section 5.2 results in trajectories constituted by short-range movements between tiles. This can be also appreciated on the $P(r)$ visible in Figure 5.7 and Appendix 7.3.2. Figure adapted from [4].

$$\max_{P \in \mathcal{H}^{(u)}} \text{LCST}(P, R). \tag{5.5}$$

We normalise the LCST score within the range $[0, 1]$, and we assign each trajectory to one of the following five bins: 0-20%, 20-40%, 40-60%, 60-80%, and 80-100%, based on their LCST score [159]. Each of these bins will represent a stratification of test trajectories, which we refer to as Test Overlap (BIN %). For example, if a test trajectory has an overlap with any training trajectory with an LCST in $[0, 0.2]$ will be assigned to the bin of trajectories with 0-20% LCST.

Accuracies on mobility overlaps

In practise, the LCST measures how much of a trajectory in test has already been observed in training by the model. A high LCST indicates the presence of recurrent mobility patterns (present in training), while low LCST indicates individual's out-of-routine mobility behaviours. We test the I, C, M and RNN models on each of the five overlap test sets. Table 5.1 and Figure 5.4B provide accuracies for test trajectories within each overlap bin. M significantly outperforms I and C for intermediate levels of overlap, with improvements across cities up to +16% (20-40% overlap), +13% (40-60% overlap). I and RNNs outperforms M by only in high overlaps 80-100% where most transitions have already been seen in training.

Figure 5.4C also shows that the distribution of $1 - S_i^{(u)}$ and we observe that the median value increases with trajectory overlap. Hence the higher the trajectory overlap, the higher the confidence of model M on I and, consequently, the lower the weight given on C information on the test set of trajectories. When the current trajectory resembles those

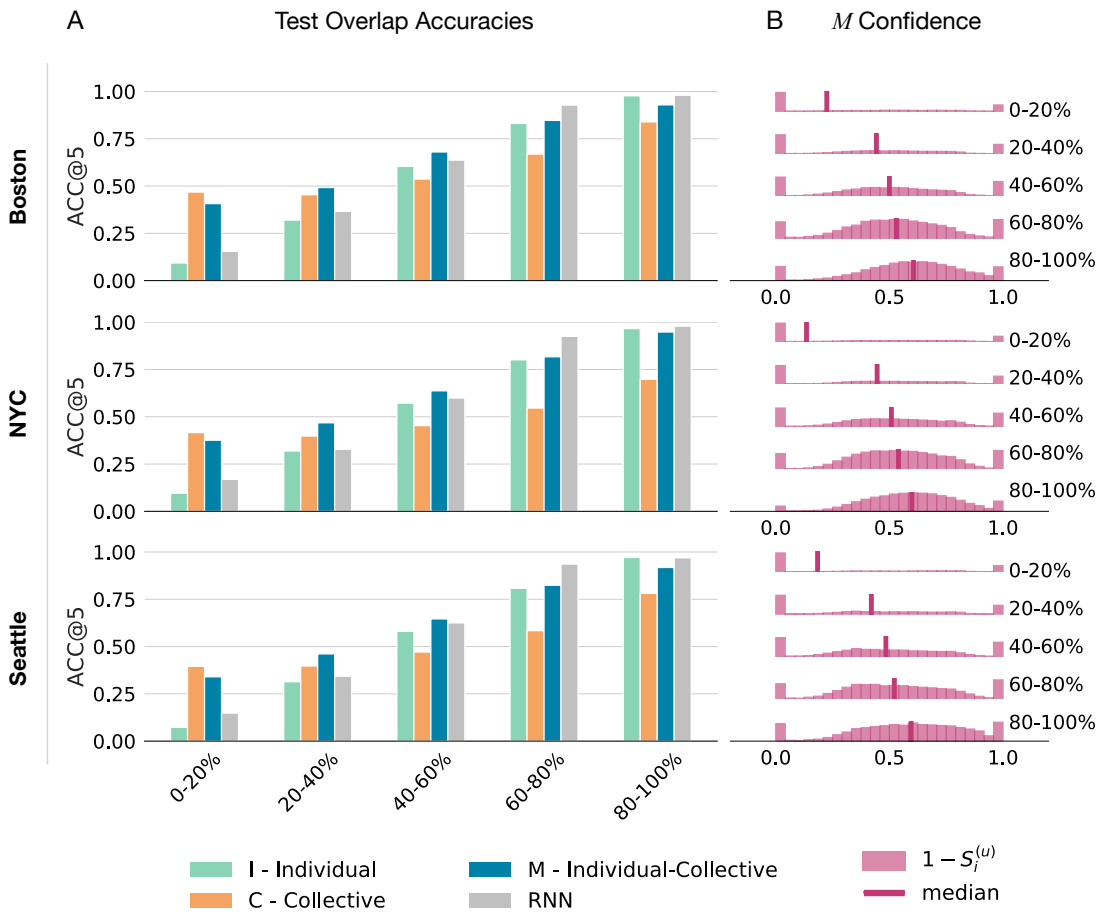

Figure 5.4: Top-5 accuracies (ACC@5) on overlaps test sets for Boston, NYC and Seattle: Accuracies (ACC@5) for Boston, New York City and Seattle using Individual (I), Collective (C), and Mixed (M) models. **(A)** Models' performance against different train-test overlap scenarios, with 0-20% being characterized by more novel mobility behaviour and 80-100% being constituted by repeated patterns behaviour. Model M shows improvements in accuracy over I and especially also against RNN in smaller overlaps, where test trajectories mostly consist of novel transitions never observed during training and exploits C. **(B)** Distributions of Model M's confidence in exploiting individual information I, computed as by $1 - S_i^{(u)}$. In the case of out-of-routine behaviours the lower median value of $1 - S_i^{(u)}$ indicates less reliance of M on individual I. In this scenario collective behaviours C instead increase the predictive power of M. The peaks observed around $1 - S_i^{(u)} = 0$ result from instances of transitions from a location i that is not represented in the training trajectories of user u. In such case of maximum uncertainty, we set $S_i^{(u)} = 1$, with M relying only on C. Figure taken from [4].

in the training set (e.g. a recurrent patterns with high trajectory overlap), it is instead best to rely on individual information. We observe that RNNs achieve a low accuracy on low trajectory overlap bins (e.g. 20-40%) and high accuracy on recurrent patterns

	full set			0-20			20-40			40-60			60-80			80-100		
	I	*C*	*M*	*I*	*C*	*M*	*I*	*C*	*M*	*I*	*C*	*M*	*I*	*C*	*M*	*I*	*C*	*M*
NYC	0.61	0.50	**0.68**	0.10	**0.42**	0.38	0.32	0.40	**0.47**	0.57	0.45	**0.64**	0.80	0.55	**0.82**	**0.97**	0.70	0.95
Boston	0.71	0.64	**0.75**	0.09	**0.47**	0.41	0.32	0.45	**0.49**	0.60	0.54	**0.68**	0.83	0.67	**0.85**	**0.98**	0.84	0.93
Seattle	0.65	0.55	**0.70**	0.07	**0.40**	0.34	0.31	0.40	**0.46**	0.58	0.47	**0.65**	0.81	0.58	**0.82**	**0.97**	0.78	0.92

Table 5.1: Accuracy of models on Full and Overlap based test sets. Performance of models I, C, M on unstratified trajectories in the test set (full set) and the performance on trajectories stratified based on their overlap (and respective overlap bin) with training trajectories using the longest common sub-trajectory (LCST) [159].

(80-100%) as we can see in Figure 5.4B. Therefore confirming that deep learning models struggle to predict out-of-routine mobility. In Appendix 7.3.6, we report the accuracy results with GeoHashes of level 7 (i.e., tiles of 153×153 meters) to verify robustness of these results on fine-grained tessellations.

5.3 Spatial properties of models' accuracy

We study here the models' accuracy given the location from which the individual next-location is predicted. To this aim we compute top-5 accuracies specifically for each subset of test transitions originating from a specific location i. Hence the accuracy of the model in computing a transition from i is ACC@5_i. To assess spatial dependencies, we compute the spatial-autocorrelation properties of this accuracy, to assess if some areas are better predicted than others.

In Fig. 5.5, we present the results of this spatial analysis for the city of Boston. When tested on trajectories constituted mostly by novel transitions (overlap 0-40%), the ACC@5_i distributions for M and C are consistently shifted to higher values compared to I (see Figure 5.5A).

Moreover, in Figure 5.5C we show the heat map of ACC@5_i, where each Geo Hash tile is plotted with the models accuracies ACC@5_i. We note that when out-of-routine is predominant (0-40% overlap), the distribution of ACC@5_i appears to have high values in specific locations i. This is particularly pronounced when using collective information for predictions. We remark that in this regime of low overlap M model mostly exploits collective information. This spatial property of clustered accuracy is quantified using Moran's Index [173, 172], and we detect that C and M exhibit a significant and positive spatial autocorrelation, indicating that locations where the models are accurate are spatially close (see Figure 5.5B-C and Appendix 7.3.1). In particular, we observe that locations where models C and M perform better are clustered in the proximity of critical urban areas. For instance, in Boston around downtown and Logan International Airport (see Figure 5.5C), in NYC in Manhattan, and in Seattle around downtown (Appendix 7.3.1). This is more relevant in low overlaps, indicating that novel mobility is better predicted in these areas, and also indicating that most novel transitions happen close to these areas (Moran's I is larger in low overlaps).

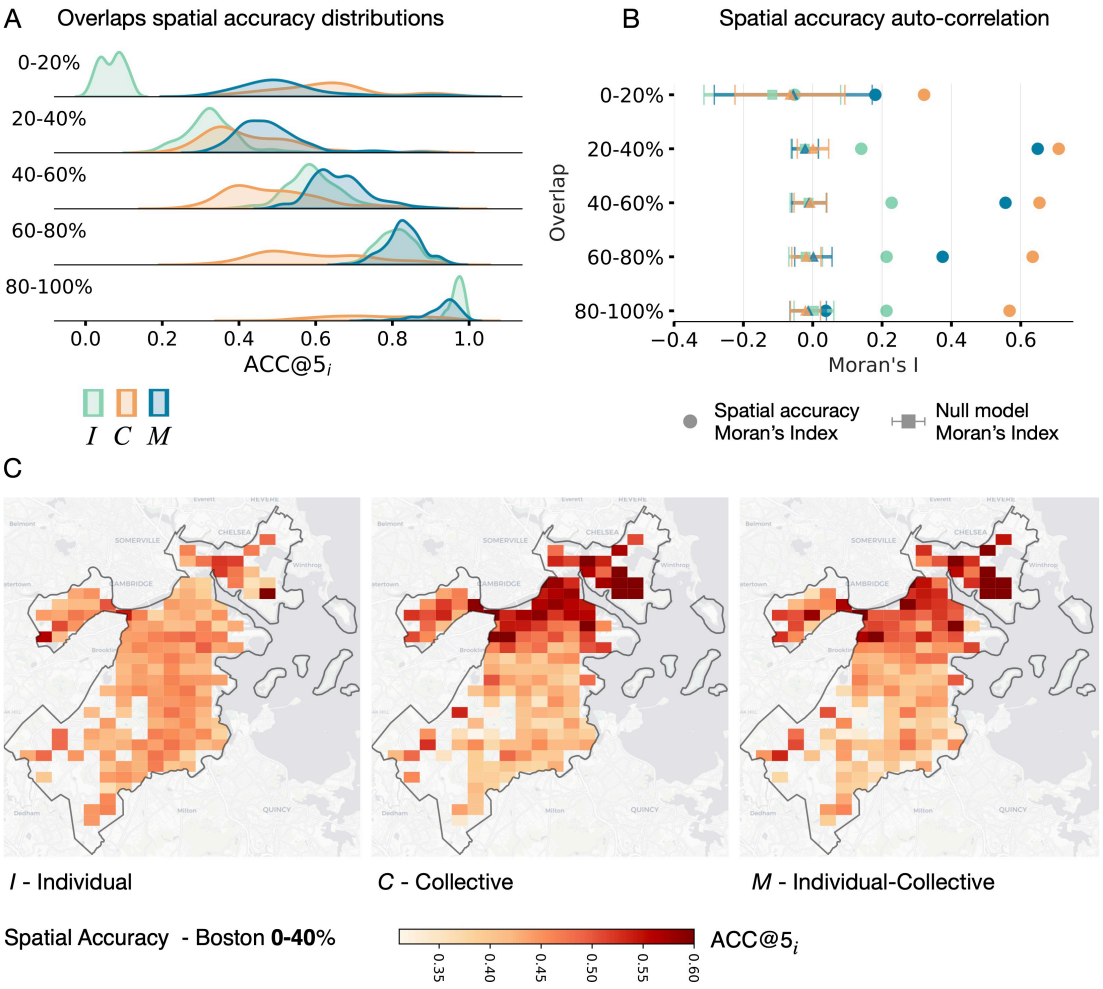

Figure 5.5: Spatial distributions of accuracies: Distribution of accuracies for I, C and M (ACC@5$_i$ in predicting a transition from an origin location i. **(A)** Distributions of spatial accuracies across different overlaps: as the test set includes more novel mobility patterns (e.g., 0-40 overlap), model M, which exploits collective information C, performs better than I. While the individual model I provides better predictions when is tested on recurrent individual patterns. **(B)** The autocorrelation properties of the spatial arrangement of ACC@5$_i$ in the corresponding overlaps is quantified via the Moran's Index [172]. When test set consists of out-of-routine behaviours, such as 0-40% overlaps, model C exhibits clustered accuracy (large Moran's Index), a property that also characterises model M since in these overlaps it mostly exploits C information. **(C)** The map of spatial accuracies ACC@5$_i$ in Boston for models I,C and M in the 0-40 overlap, where test transitions are mostly out-of-routine. Notably, for C and M models, locations with higher accuracies are concentrated in proximity of downtown areas in Boston. Interestingly, the area corresponding to the Boston Logan International Airport (upper right) is also well predicted by collective behaviours. Figure taken from [4]. Maps: Stamen Maps

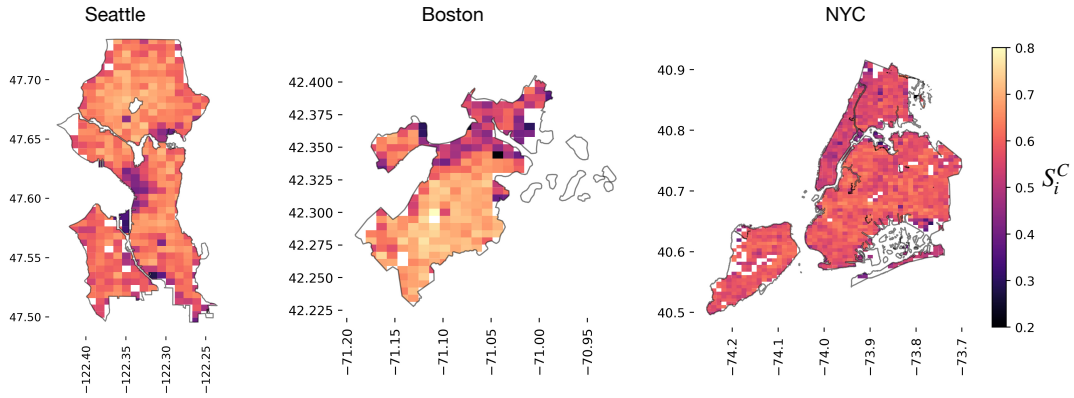

Figure 5.6: Spatial map of collective entropy S_i^C: Spatial distribution of normalized entropy S_i^C of the Collective OD transition probabilities. Areas in blue are characterized by lower entropy, thus dominated by a collective behavior which is more local and more predictable. This also results in a larger predictability and a stronger improvement of predictions of out-of-routine mobility. Also refer to Fig. 5.7 where in these areas we detect an earlier cutoff and people generally moving shorter distances (staying more in the area and thus resulting more predictable). Figure adapted from [4].

5.3.1 Predictive capabilities of collective behaviours in proximity of Points of Interest

This positive spatial autocorrelation for C and M suggests that an underling environmental factors might exist as driver for the accuracy of collective mobility behaviours and spatially-clustered urban factors. To investigate this hypothesis, we first measure the entropy of collective behaviours from a location i to assess their predictability:

$$S_i^{(C)} = -\frac{\sum_{k \in L^{(C)}} C_{ik} \cdot \log(C_{ik})}{\log |L^{(C)}|} \in [0, 1], \tag{5.6}$$

where $L^{(C)}$ is the set of unique locations in C. In Figure 5.6 we can observe the spatial map of $S_i^{(C)}$ for the three cities. We quantify the relationship between the entropy of collective behaviors from a location i and its model C accuracy via a direct Pearson correlation. We find that $S_i^{(C)}$ is strongly anti-correlated with ACC@5$_i$ (Pearson correlation of $\rho = -0.85$), indicating that locations from which C is the most accurate are the ones with the lowest entropy $S_i^{(C)}$. See Figure 5.7A for the scatter plot.
Also here we find that locations with low $S_i^{(C)}$ are clustered in proximity to specific urban areas. We hypothesise to be locations hosting key commercial, financial, and cultural venues. Therefore points of interests.

Proximity with Points of Interest. To verify this hypothesis, we collect from OpenStreetMap (OSM) the number of points of interest (POIs) in each location i, W_i (see Appendix 7.3.7). We then split each city into two different areas: one comprising

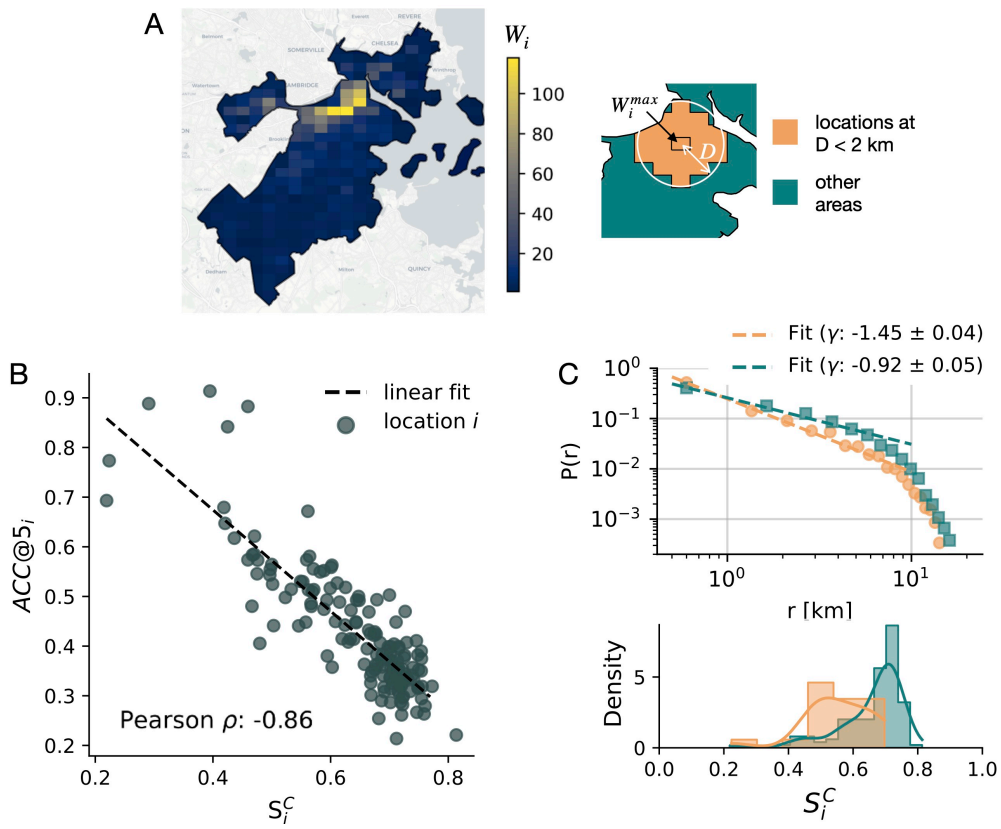

Figure 5.7: Statistical properties of Collective mobility in Boston. Statistical properties of C_i analyzed based on the proximity of Points of Interests (POIs) W_i within Boston are analysed. **(A)** Map of the number of POIs (W_i) in a Geo hash 6 tile i, extracted from OSM using the set of amenities described in Appendix 7.3.7. **(Inset)** Geo hash locations within a distance $D = 2$ Km from the tile with largest number of POIs (W_i^{max}) shaded in orange, indicative of downtown Boston and its most attractive areas. Other areas at distance $D > 2$ Km are indicated in teal. Albeit this separation of the city is solely based on proximity to high POIs areas, it allows to separate human mobility in two regimes. **(B)** Spatial accuracy of C model ($ACC@5_i$) to predict mobility originating from tile i is compared with the collective entropy S_i^C from that location. We observe anti correlation in the Pearson $\rho = -0.86$, indicating that areas from which C is more predictive are also characterized by a lower S_i^C. We report in **(C)** the distribution of travel distances ($P(r)$), distinguishing between origins within the two aforementioned areas and fitted with a power-law function [23] in the interval of 0 to 10 Km. The exponent of $\gamma = -1.45 \pm 0.04$ underscores the prevalence of localised mobility when individuals are in proximity to W_i^{max}, while in other areas we have an exponent of $\gamma = -0.92 \pm 0.04$. This observation aligns with the notion that mobility near POIs tends to be more concentrated and less spatially dispersed towards specific destinations. This behaviour is further corroborated by the entropy S_i^C distribution, which skews towards lower values and indicates mobility directed towards specific tiles. Figure adapted from [4]. Maps: Stamen Maps.

locations within a geographical distance D from W_i^{max}, and the other consisting of locations at a distance greater than D from W_i^{max} (see Figure 5.7B for $D = 2$ Km in

Boston). These locations with W_i^{max} correspond to Boston downtown, Seattle downtown and Lower Manhattan in NYC, respectively. This separation allows us to study the properties of human mobility whether they are in proximity of the city centre, or not. We compute the distribution of travel distances, $P(r)$ of jumps having the origin location within distance $D = 2$ km from W_i^{max} (in orange) and those originating in other areas (in teal). We show the results for Boston in Figure 5.7C. $D = 2$ Km stands as a trade-off between having a sufficient number of transitions close to W_i^{max} and avoiding the inclusion of too-distant tiles. We show a sensitivity analysis for other proximity thresholds, and results hold for other values of D (see Appendix 7.3.2). For the two distributions P(r), we fit a power-law $P(r) \sim r^{-\gamma}$ [23] in the interval of 0 to 10 Km (Fig. 5.7C) to better quantify the dependency with distance. When individuals are close to the city centre, the exponent $\gamma = -1.45 \pm 0.04$ in Boston highlights the predominance of more localised mobility (orange points) with shorter jump length with respect to areas far from POIs which are characterised by longer jumps.
Moreover we also focus on the entropy of collective movements within these areas. We find that the distribution of $S_i^{(C)}$ skews towards lower values when location i is within distance D to W_i^{max} (orange line in Figure 5.7D). Additionally, individuals moving in locations in the proximity of POIs tend to travel shorter distances compared to those originating in other areas. The distribution for Boston is present in Figure 5.7C. These results support the idea that mobility in proximity to POIs areas is less spatially dispersed, with the movement being more likely to a destination concentrated in a small subset of locations. We see similar results in Seattle and NYC in Appendix 7.3.1.

Overall, this spatial analysis of mobility patterns in proximity to POIs reveals the statistical properties of C that can justify its increased predictive power, and better explain why collective mobility better captures individual out-of-routine behaviours in these areas.

5.3.2 Robustness to Dataset Pruning

It's important to highlight that the original sample size of flows T_i utilised for estimating C_i probabilities presents significant differences across tiles i. Geo Hash tiles in areas with high density of POIs have more user check-ins in the dataset. To assess that the ACC@5_i improvements in these urban areas are not a consequence of probabilities C_i which simply may benefit from a larger sample size, we conducted a pruning test. We reduce the collective information that can be accessed by our model M by pruning the collective OD flows from a tile i. In what follows we described the stochastic sub sampling approach to estimate pruned C_i.

Pruning of Collective OD in over-represented locations i. We begin from the cumulative distribution function of the number of transitions T_i in the training set from location i used to compute C_i. If in a tile T_i exceeds the median ($50-th$ percentile) of all locations (which we refer to as T_i^{max}), we uniformly and randomly pruned the subset of transitions from location i until this threshold value. As a consequence, probabilities C_i

are computed with a comparable amount of information across all tiles i. This pruning process is presented in Appendix 7.3.4 with Algorithm 2.
In Figure 5.8 we present a visualization of the distribution of T_i and the effect of the pruning process for the urban area of Seattle. We observe in Figure 5.8A that the flows T_i spans different orders of magnitude for different Geo Hash tiles in the original dataset. The cumulative distribution function and percentiles (Figure 5.8B) also highlight this strong heterogeneity in T_i. We prune the set of training transitions following algorithm 2. The resulting set of flows T_i from the pruned Collective OD is presented in Figure 5.8C, where the maximum flow is $T_i^{max} = 418$ for Seattle, representing the 50-th percentile across locations i int the original dataset. This random sub-sampling process was repeated for a number of samples $N_{samples} = 10$, generating pruned C, from which $N_{samples}$ pruned Markov models are computed. These are then used to generate different models M, which are then tested. Finally, for each tile i we compute the accuracy ACC@5$_i$ as the average accuracy of the $N_{samples}$ of pruned M from that tile.

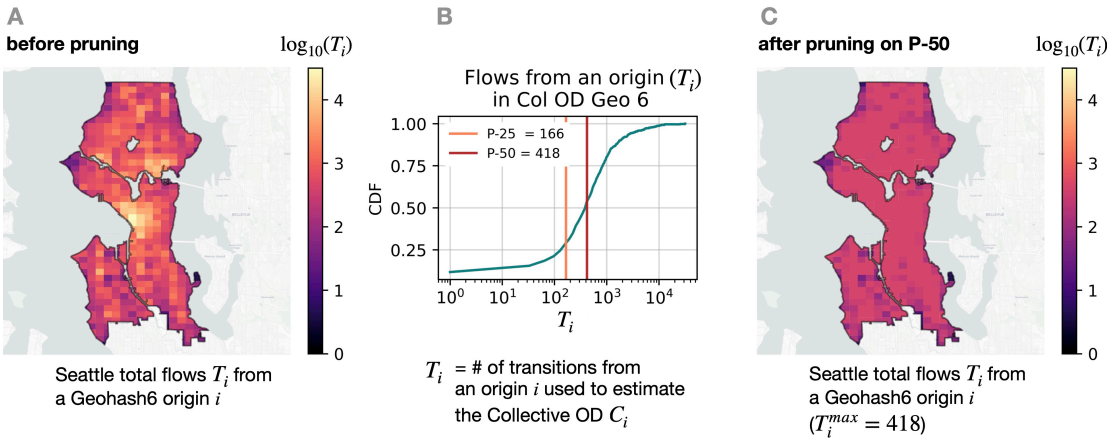

Figure 5.8: Sub-sampling of Collective OD flows and reduction of sample-size effects. A) The number for total transitions in the Cuebiq dataset from a Geo hash tile i for Seattle, computed as $T_i = \sum_j T_{ij}$ where T_{ij} is the flow count of transitions from location i to location j. B) The cumulative distribution function of T_i highlights that several orders of magnitude separate few locations from the remaining areas. The 25-th and 50-th percentiles are presented in the legend. We use the 50-th percentile, defined as T_i^{max}, as threshold to pruned the over-represented origins and generated $\bar{C}$. C) After the pruning process described in Algorithm 2 the number of flows T_i in all tiles is now comparable and does not suffer from strong heterogeneity (Here T_i^{max} = 418). Figure adapted from [4].

Results with pruned collective OD C. In Figure 5.9, we report the average ACC@5$_i$ of model M in Boston. Here we adopt exclusively novel transitions extracted from the overlap $0 - 40\%$, i.e. transitions instances in which the correct destination j from a location i in the test is not present in $I_i^{(u)}$ for user u. The rationale of this choice is

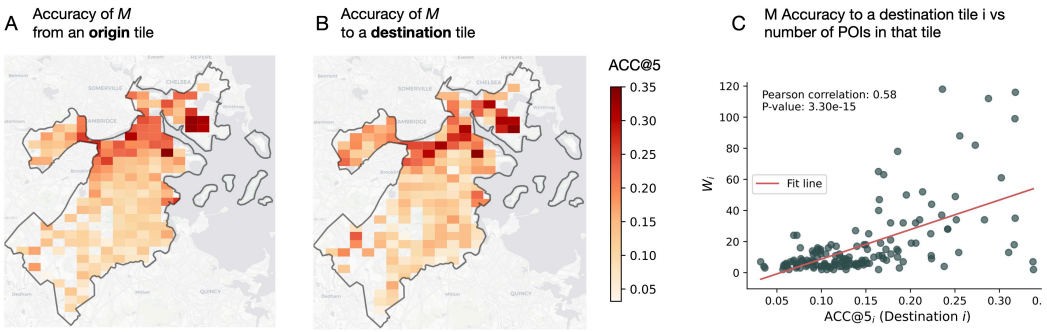

Figure 5.9: Spatial accuracy of M model on novel mobility with pruned $\bar{C}$ in Boston. The accuracy of M model in predicting a novel transition (never seen in a individual u's historical trajectories $\mathcal{H}^{(u)}$) from an origin in a Geo Hash i is presented in panel **(A)**. In panel **(B)** the accuracy ACC@5$_j$ in predicting a transition towards a destination j is depicted. We remark that the sample size T_i used to estimate C_i exhibits strong differences, with central areas being more densely sampled. To mitigate this potential bias, we employed a stochastic sub sampling process to estimate C_i. Thus, spatial heterogeneity that we observe in ACC@5$_i$ cannot be attributed to a richer dataset used to estimate C in these areas. Only locations i with at least 100 test transitions instances used to compute the ACC@5$_i$ are shown. **(C)** We show the Pearson correlation between the accuracy in predicting a destination and the number of POIs in that destination. Figure adapted from [4]. Maps: Stamen Maps

to investigate the spatial dependencies of pruned collective information $\bar{C}$ to inform the M model when individual patterns can not aid by definition because they lack the necessary information. Hence the accuracy of model M from a tile i in this test set can be attributable exclusively to collective pruned behaviors. In Figure 5.9A we present the accuracy ACC@5$_i$. Moreover, we also compute the accuracy in estimating correctly a destination to a specific tile j. This spatial accuracy ACC@5$_j$ is computed as the fraction of instances correctly predicted that have as destination the tile j. The results in Boston are in Figure 5.9B, where the resulting accuracy is the average over the set of sub-sampled $\bar{M}$ built from Collective OD pruned $\bar{C}$, as described in Algorithm 2. We observe that the enhanced accuracy in of M and C ACC@5$_i$ in proximity to areas with high density of POIs is still present and compatible to distributions in Figure 5.5. Therefore, the spatial improvement in proximity of POIs observed in M due to C in the accuracy ACC@5$_i$ cannot be attributed to the training dataset used to estimate C_i probabilities being richer and more representative in specific areas.

Accuracy in predicting a destination. Moreover we observe in the map of ACC@5$_j$ in Figure 5.9 that the best predicted destinations also appear be clustered around high POIs areas. We compute the Pearson correlation between the accuracy of model $\bar{M}$ in predicting correctly a destination ACC@5$_j$ and the number of POIs W_j in that destination tile j. In Fig. 5.9 we show the scatter plot between these two observable for each location in Boston. In all cities we can appreciated that areas with large number of POIs tend to

	Mar. - Apr.			Apr. - May			May - Jun.			Jun. - Jul.			Jul. - Aug.		
	I	C	M	I	C	M	I	C	M	I	C	M	I	C	M
New York	0.484	0.541	0.655	0.390	0.541	0.617	0.358	0.536	0.602	0.325	0.524	0.585	0.290	0.515	0.570
Boston	0.491	0.632	0.696	0.394	0.609	0.654	0.350	0.620	0.651	0.307	0.619	0.642	0.255	0.615	0.637
Seattle	0.424	0.578	0.649	0.352	0.582	0.629	0.311	0.583	0.619	0.271	0.557	0.596	0.236	0.543	0.579

Table 5.2: Models accuracies with behavioral shifts (COVID-19 monthly windows). Accuracies (ACC@5) of I, C and M trained on trajectories collected before March 1st. Models are then evaluated on trajectories during and after lockdown every month after the pandemic declaration (March 11). Here, we study the COVID-19 pandemic s a disruptive event that introduces behavioral changes and alters mobility patterns of individuals. We notice how the individual model loses half of its accuracy up to the last testing period. Instead, the collective behaviors generalize even after months, allowing the M model to maintain pre-disruptive event accuracies.

be the ones also better predicted as destinations by the collective behaviours. This trend in particular is more pronounced for Boston and Seattle, with a Pearson $\rho = 0.58$ and $\rho = 0.62$ respectively, while in NYC $\rho = 0.29$ (see Appendix 7.3.4).

5.4 Distribution shift: Models reliability under COVID-19 restrictions

In this section we analyze the generalization of M model in the context of distribution shifts. Events that disrupt the mobility patterns of individuals limit the predictive power of models trained using trajectories collected precedent to these events. As disruptive event we use the COVID-19 pandemic. The pandemic significantly altered people's mobility patterns [174, 146, 158, 175, 176], with non-pharmaceutical interventions inherently inducing a shift in how people moved and visited locations [177, 178, 146]. In this context, we investigate the reliability of models during these behavioural shifts.
To this aim, we first train I, C, and M using trajectories recorded until March 11th, 2020. Then we evaluated their performance on a test set consisting of five months of data collected between March 11th, 2020 (the World Health Organization's pandemic declaration), and August 11th, 2020. For each month in this test set, we report models' ACC@5 in Table 5.2.
We observe that in each city, I exhibits a notable decline in performance as the test period (month) is farther from the training period. I loses up to 44.16% of its predictive power when tested on trajectories collected between July and August. C's performance also degrades over the testing months, but with a more moderate reduction in accuracy. Although M also experiences a drop in accuracy, it remains the most predictive model overall, with a limited drop in accuracy of 5.32%. Therefore models based solely on individual-level information are less resilient to behavioural shifts than models that exploit collective information. However, the combination of individual and collective information contributes to the model's resilience and good predictive power even in scenarios of behavioural changes.

5.5 Conclusion

Collective behaviors significantly influence individual decisions, a phenomenon extensively documented in studies on collective intelligence, social psychology, and behavioral economics.
In this chapter we have seen how collective mobility in urban scenarios can be leveraged to improve the prediction of out-of-routine behaviors. We propose a parameter-free approach to next-location prediction that dynamically integrates individual and collective information. Our model offers a potential solution to the limitations inherent in current state-of-the-art models, including sophisticated deep-learning approaches, in the challenge of predicting out-of-routine movements. Notably, our approach demonstrates strong performance in predicting out-of-routine movements, including those during disruptive events like the COVID-19 pandemic, even when the overlap between training and test trajectories is low. Integration of dynamic mechanisms exploiting an individual entropy to mix individual and collective Markov mobility chains allows to obtain predictions that are both interpretable and accurate in scenarios involving both routinary and out-of-routine movements. Furthermore, our study reveals that accurate predictions relying on collective information spatially cluster around urban areas characterized by a dense concentration of key commercial, financial, and cultural venues. These findings align with recent research on flow generation, indicating the significant role of points of interest in shaping human mobility flows, a concept that we also exploited in Chapter 4 to generate OD flows that simulated collective mobility patterns in the Greater London Area. Moreover, the predictive power of collective mobility in these areas was found to be robust even to the pruning of the accessible information in the dataset.

Discussion

Understanding the interplay between human mobility and transportation networks is fundamental in many challenges that society is currently facing: from the design of sustainable, efficient and inclusive cities up to the containment of epidemic spreading. This thesis aimed to explore this interplay, spanning from theoretical frameworks to data-driven analyses, with the ultimate goal of shedding light on emergent properties and predictive models in urban mobility. From the hidden geometries of diffusive processes mediated by mobility flows, the complex network topologies of transit systems, up to the statistical regularities in human trajectories in relationship with the urban functional areas.

In Chapter 3 we started from a more theoretical perspective on the hidden geometries induced by dynamical processes on top of complex networked systems. To this aim, network effective distances allow to capture the fundamental mechanisms that drive these processes, unraveling these geometries which in turn can be used to predict arrival times. In particularly, we focused on the application on human mobility viewed through the lens of diffusive flows. We (i) introduced a multi-pathways temporal distance, which captures the role of the ensemble of paths between source and target nodes in conveying signals. We demonstrated its efficacy in predicting arrival times of propagated signals, with applications ranging from protein-protein dynamics to epidemic spreading. With a focus on the probabilistic diffusive nature of collective mobility, we (ii) introduced the Information Distance metric, derived from the Laplacian matrix. Using global air transportation networks as a proxy for human mobility flows, we have shown how the Information Distance effectively computes the arrival time of infectious agents to target destinations. Finally, (iii) the predictive capabilities of effective distances provided valuable insights into the pandemic potential of SARS-CoV-2 Variants of Concern, by integrating genomic surveillance with global mobility data in a comprehensive pipeline. This overall allowed us to better understand the role of multiple pathways in the definition of effective network distances, establishing the relevance of these network based metrics in providing relevant insights.

We then discussed in Chapter 4 the optimization of an effective temporal travel distances in the context of transportation networks. Here we studied the emergence of complex topologies in planar networks when the edges velocities were optimized to provide the most efficient network topology, in terms of effective travel time. Through a simple

framework based on a planar substrate, we highlighted how spatial flow probabilities, congestion dynamics and urban features, interact to shape optimal network structures. Our framework first revealed transitions from tree-like to lattice-like network configurations based simply on the range of flows T_{ij}. When traffic is instead routed following realistic traveling behaviors and concepts such as User Equilibrium, congestion effects instead alter these optimal structures. We devise an analytical condition that we use to predict the amount of congestion necessary to disrupt the tree structure leading to the emergence of alternative pathways (and loops) to sustain and mitigate congestion. We show that by biasing these travelling probabilities adding spatial attractiveness W_j and adopting spatial interaction models simulating collective mobility patterns, we recover some complex topological features typical of human transportation networks. Finally, by modeling the real spatial structure of the Greater London Area, we recover the scaling properties similar to real-world transportation systems such as the London subway system.

Finally, in Chapter 5 we explored the predictive potential of collective human mobility behaviors, particularly in informing next-location predictors when individuals deviate from their routinary patterns. We studied a dynamical integration of individual and collective mobility behaviors exploiting the definition of an individual normalised entropy. Our data-driven approach showcased the significant influence of collective mobility on individual decisions, particularly during disruptive events like the COVID-19 pandemic. Our model offered a parameter-free solution to next-location prediction, surpassing the limitations of traditional deep-learning approaches in out-of-routine instances. Interestingly, accurate predictions relying on collective information clustered in areas with a high concentration of key commercial, financial, and cultural venues, with the best predicted flows pointing towards areas with large density of points of interests. This finding aligns with research on flow generation, highlighting the impact of points of interest on shaping human mobility patterns.

Collectively, these findings illustrate the complex interplay between human mobility, transportation networks, and urban spatial features, with implications for diverse fields ranging from epidemiology to urban planning. For example, the network effective distance serves as a paradigmatic illustration of complexity science's ability to offer a framework for cross-disciplinary integration and modeling. The resulting comprehensive modeling of networked dynamical systems via effective distances not only allows to map apparently disordered processes into coherent and predictable patterns [1], but also to inform and enrich analyses across diverse domains [2]. Moving further on tasks where network science can provide optimal benchmark features, our framework for the optimization of spatial networks in urban contexts may show further extensions to better accommodate concepts of multi-layer network and other aspects in urban planning research. Moreover, it has the potential to be adjusted to provide guidance on the optimal expansion of an existing transit network. Finally, future analyses may focus on the interdependence of collective mobility properties and urban space, further unravelling the complex regularities of human mobility in cities. As an example, mixtures of individual and collective pattern may incorporate an additional mechanism that accounts for the density of points of

interest in the area where the individual is currently located. Eventually, an intriguing research challenge lies also in the design of public transit infrastructure to optimize collective flows on multi-layer transportation modes and active micro mobility.

Appendix

7.1 Appendix A

7.1.1 Adapting $\mathcal{L}_{\mathcal{MP}}$ to the SIR model.

We discuss here the limitations of employing $\mathcal{L}_{MP}$ in the case of a SIR model described in Chapter 3 in Eq. 3.18. It is not possible to map the 3 variables (i,s,r) of SIR model to the set of $\mathbf{M}$ functions in Eq. 3.1. Therefore analytical results for the scaling properties (such as τ_i) can't be derived for Eq. 3.18. Moreover, the assumption of dealing with small perturbations around an equilibrium basin is unfeasible as introducing a small fraction of infected $i_n \approx 10^{-3}$ triggers exponential growth over time for $i_n(t)$. We can try to disregard the SIR self-dynamics which is a valid assumption in the regime of $i_m(t) < \epsilon$, where $\sigma(i/\epsilon)$ sets the infection threshold for node i, and therefore focus on the diffusion term in Eq. 3.18. Under this condition: $\frac{\partial i_n}{\partial t} = \gamma \sum_{m \neq n} P_{mn} (i_m - i_n)$. Here applying same conditions $i_n = \sum_{m \neq n} P_{mn} i_m$ and deriving it with respect to node m, i_m, yields $\partial x_i / \partial x_j = P_{mn}$.
Thus, P_{mn} serves as a substitute for R_{mn}. However, it should be noted that it lacks the steady-state normalization factor x_m/x_n.
Finally, we set as first approximation the exponent for the estimation of τ as $\theta = 0$ as for a simple diffusive dynamics. We note that a more complex derivation of θ can be employed via numerical approximations [17]. Despite these strong approximations the Spearman's rank coefficient is above 0.9 for $\mathcal{L}_{\mathcal{MP}}$ as reported in Fig. 3.8.

7.2 Appendix B

7.2.1 Optimization via simulated annealing

To generate the optimal topologies, i.e. the configuration of weights/velocities $\{w_e\}$ such that the flow weighted effective temporal distance E in Eq. 4.2 is minimized, we employed a Simulated Annealing (SA) scheme [179]. We start from the initial lattice substrate having all weights equal $\{w_e\}_{init} : w_e = 1.0, \forall e$ and perform (SA) to obtain $\{w_e\}_{optimal}$ (see Figure 7.1).
Since in the generation of optimal topologies we enforce the condition that the total sum $C_G = \sum_{e \in G} d_e w_e$ is conserved, this requires that also the Monte Carlo move [132] to propose a new configuration needs to guarantee the conservation of C_G. Here we provide the algorithm for the proposal of new configurations of edges' weights. Specifically, given a configuration $\{w_e\}_t$ at SA-step t, we propose a new configuration $\{w_e\}_{t+1}$ and accept it via Metropolis rule [132]. To propose a candidate while respecting the condition on the total cost C_G, we implement the algorithm described in Algo. 1.

Input : Configuration $\{w_e\}_t$ at step t, parameter α_{MC}
Output : Configuration $\{w_e\}_{t+1}$ at step t

- Select randomly two edges e_1 and e_2;
- Reduce the weight of e_1 by α_{MC}: $w_{e_1} \leftarrow w_{e_1} - \alpha_{MC} \cdot w_{e_1}$;
- Compute the total associated weight that will be assigned to e_2:
  $\delta_{MC} \leftarrow \alpha_{MC} \cdot w_{e_1} \cdot d_{e_1}/d_{e_2}$;
- Update e_2 weight using δ_{MC}: $w_{e_2} \leftarrow w_{e_2} + \delta_{MC}$;

return Configuration $\{w_e\}_{t+1}$ at step t;

Algorithm 1: Candidate edge's configuration proposal Algorithm

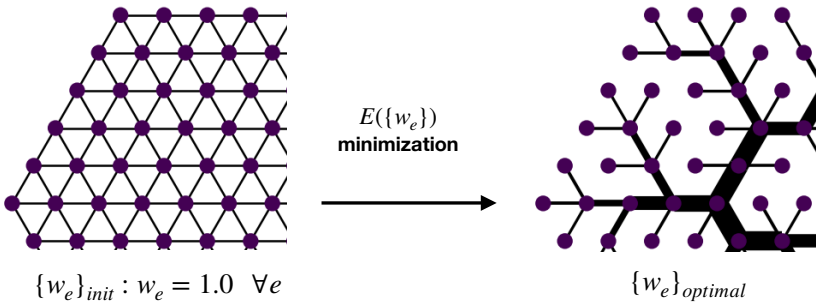

Figure 7.1: Optimization via Simulated Annealing: Initial weight configuration (the lattice substrate) is optimized via SA to obtain the optimal configuration.

7.2.2 Simple substrate optimization

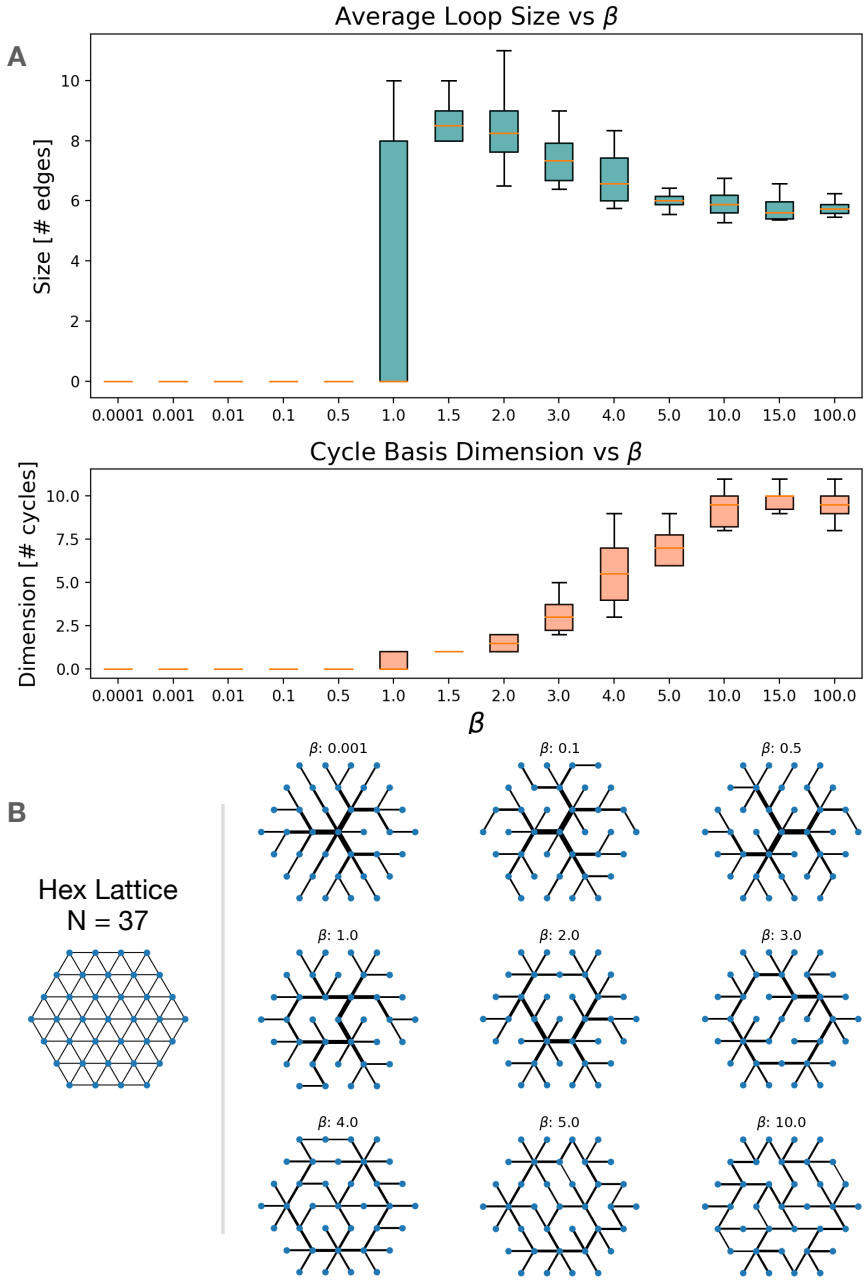

Figure 7.2: Loop Dimension vs β on HEX. We adopt the minimum cycle basis to study the presence of loops in the optimal configurations. For each β the median and its MAD computed on the ensemble of optimal networks from a HEX size $s = 3$ substrate. **A)** Average number of edges (length) of the loops present in the cycle basis. In the lower panel the number of loops in the basis is shown. **B)** The optimal network transitions from tree structure to a lattice-like with many small loops. Figure taken from [3].

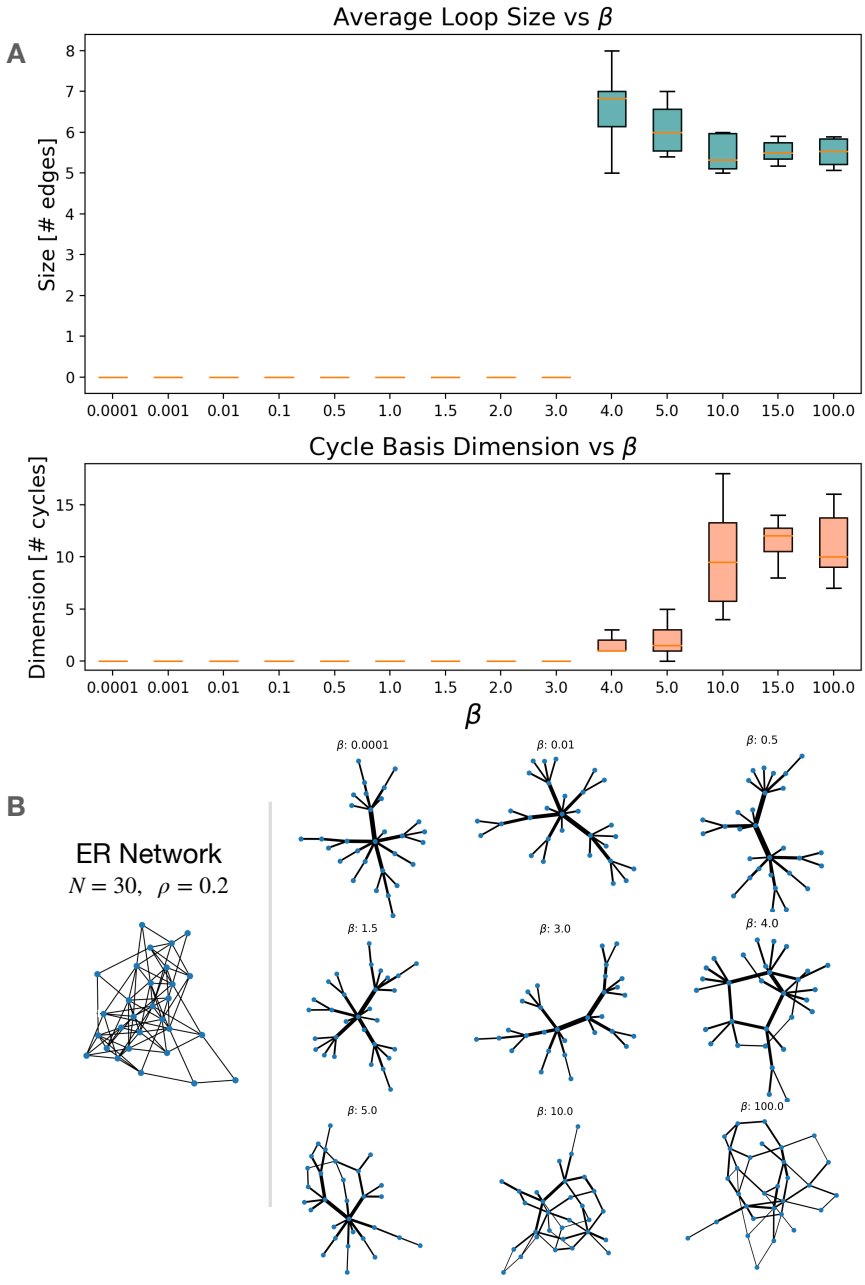

Figure 7.3: Results on the ER network. A) Average loop size and Cycle basis dimensions exhibit the similar transitions as those presented for the HEX lattice. B) Sample networks for different β in the case of ER with N=30 and edge probability $\rho_{ER} = 0.2$. Figure taken from [3].

7.3 Appendix C

7.3.1 Spatial accuracies

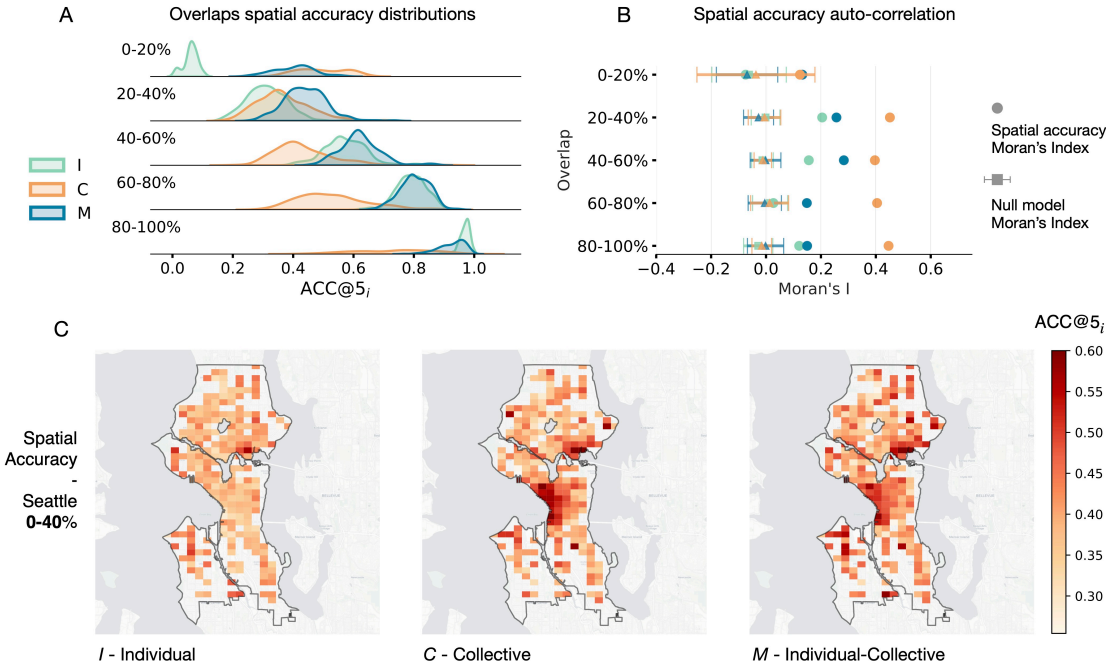

Figure 7.4: Spatial distributions of accuracies - Seattle. Distribution of accuracies for the individual I,the collective C and model M which combines individual and collective information, in predicting a transition from a Geo Hash 6 tile. **(A)** Spatial accuracies across different overlaps. **(B)** Spatial autocorrelation of the models' accuracies in corresponding overlaps quantified via the Moran Index. For larger out-of-routine behaviours like 0-20 and 20-40 overlaps, model C exhibits clustered accuracy (large Moran Index). This is a consequence of novel mobility patterns being better predicted by the collective behaviors in proximity of POIs **(C)** Map of spatial accuracies in Seattle for models I,C and M in the overlap 0-40% (test trajectories in overlaps 0-20% and 20-40% merged in one single overlap). Figure taken from [4]. Maps: Stamen Maps

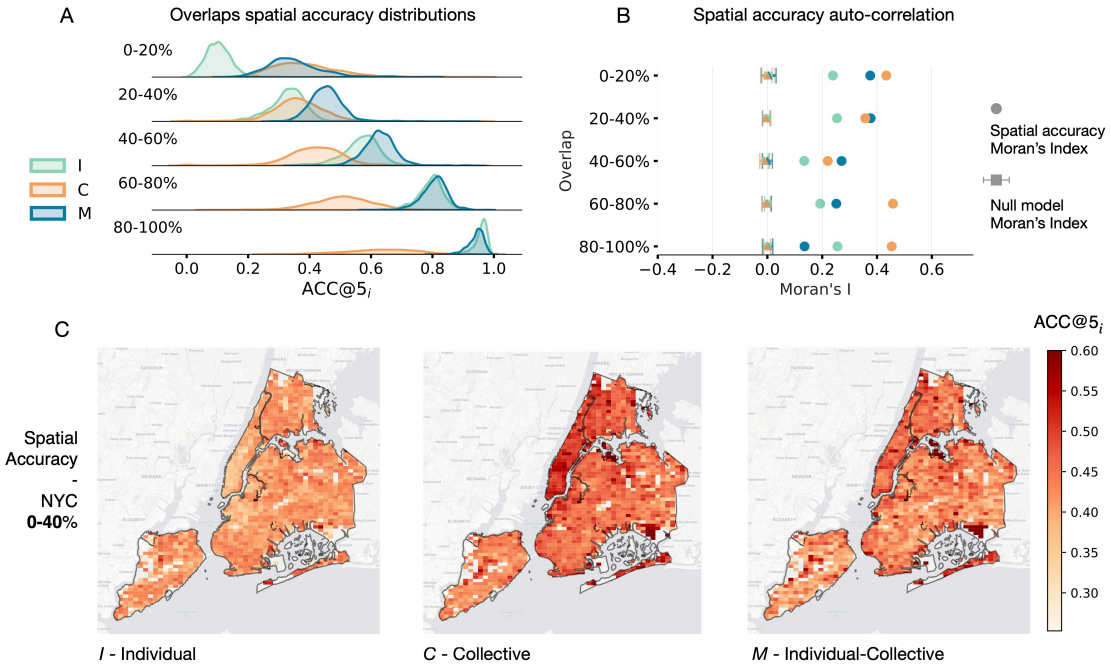

Figure 7.5: Spatial distributions of accuracies - NYC. Distribution of accuracies for the individual I,the collective C and model M which combines individual and collective information, in predicting a transition from a Geo Hash 6 tile. **(A)** Spatial accuracies across different overlaps. **(B)** Spatial autocorrelation of the models' accuracies in corresponding overlaps quantified via the Moran Index. For larger out-of-routine behaviours like 0-20 and 20-40 overlaps, model C exhibits clustered accuracy (large Moran Index). This is a consequence of novel mobility patterns being better predicted by the collective behaviors in proximity of POIs **(C)** Map of spatial accuracies in New York City for models I,C and M in the overlap 0-40% (test trajectories in overlaps 0-20% and 20-40% merged in one single overlap). Figure taken from [4]. Maps: Stamen Maps

7.3.2 Collective mobility statistical properties

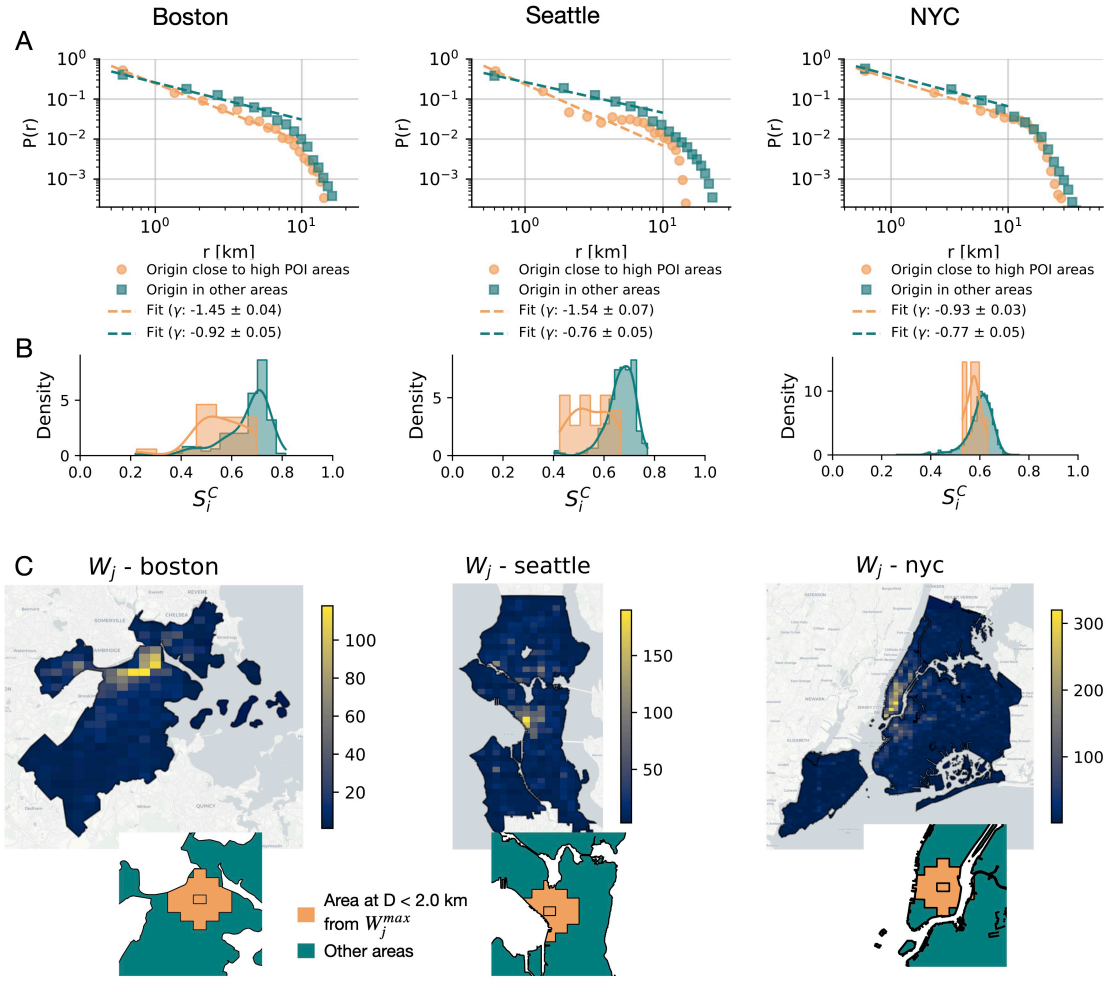

Figure 7.6: Statistical properties of Collective mobility and proximity of Points of Interests. We first define the area with highest number of POIs as the attractiveness centre W_j^{max} (corresponding to downtown areas in the three cities). Then we split the cities into tiles within a distance D from this tile and other areas with distance greater than D. We report in **(A)** the distribution of travel distances ($P(r)$), distinguishing between origins within 2 km from W_j^{max}. Distribution are fitted with a power-law function $P(r) \sim r^{-\gamma}$ [23] in the interval of 0 to 10 Km. The different characteristic exponents γ for the three cities highlight a more localised mobility in proximity of POIs. **B)** Entropy S_i^C distributions of collective mobility originating from tiles in the two regions. **C)** Map of the number of POIs (W_i) in a Geo hash 6 tile i, extracted from OSM following the procedure in Appendix 7.3.7. **(Inset)** Areas within a distance $D = 2$ Km from the tile with largest number of POIs (W_i^{max}) shaded in orange. While in teal other areas. Albeit this separation of the city is solely based on proximity to high POIs areas, it separates collective human mobility in two different predictability regimes. Figure taken from [4]. Maps: Stamen Maps

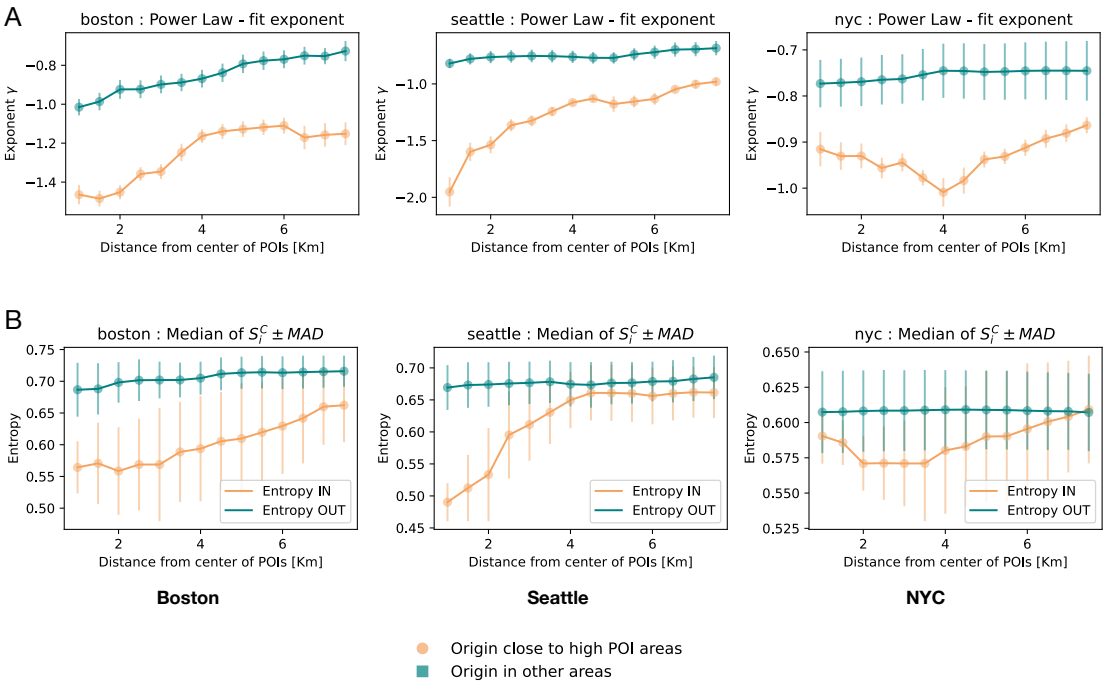

Figure 7.7: Sensitivity analysis of travel distances $P(r)$ and collective S_i^C in proximity of POIs and in other areas.(A) Exponents γ of power law fits (in the range 0 to 10 km) of the travel distributions $P(r)$. The exponents are presented for the two separate distributions: $P(r)$ for transitions having location origin at distance D from the tile W_j^{max}, and $P(r)$ for origin in other areas. **B)** The median value and its median absolute deviation (MAD) of entropy S_i^C distributions for origins within D or in other areas. A sensitivity analysis of these exponents is proposed for a range of different separation distances D. For Seattle and New York City in particular we observe a range of distances in which we can appreciate larger differences between the exponent γ and median value of S_i^C between the two proximity-based regions. In particular D = 2 Km represents a proximity distance where in all cities we observe different predictability and mobility properties. Figure taken from [4].

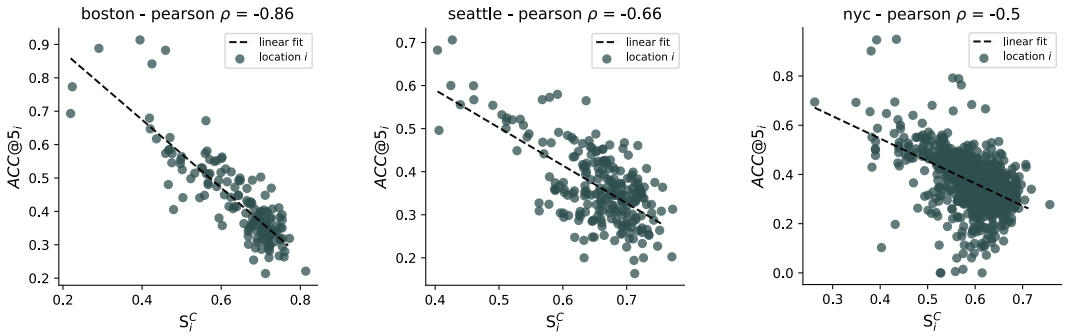

Figure 7.8: Pearson correlation of S_i^C versus ACC@5$_i$. We report the Pearson correlation ρ of accuracy of C model (ACC@5$_i$) in a tile (in the 0-40% overlap) versus its entropy S_i^C for all the three cities in the dataset. We observe anti correlation between the two variables, specifically for Boston $\rho = -0.86$, Seattle $\rho = -0.66$ and NYC $\rho = -0.5$. Figure taken from [4].

7.3.3 Sub-sampling collective information

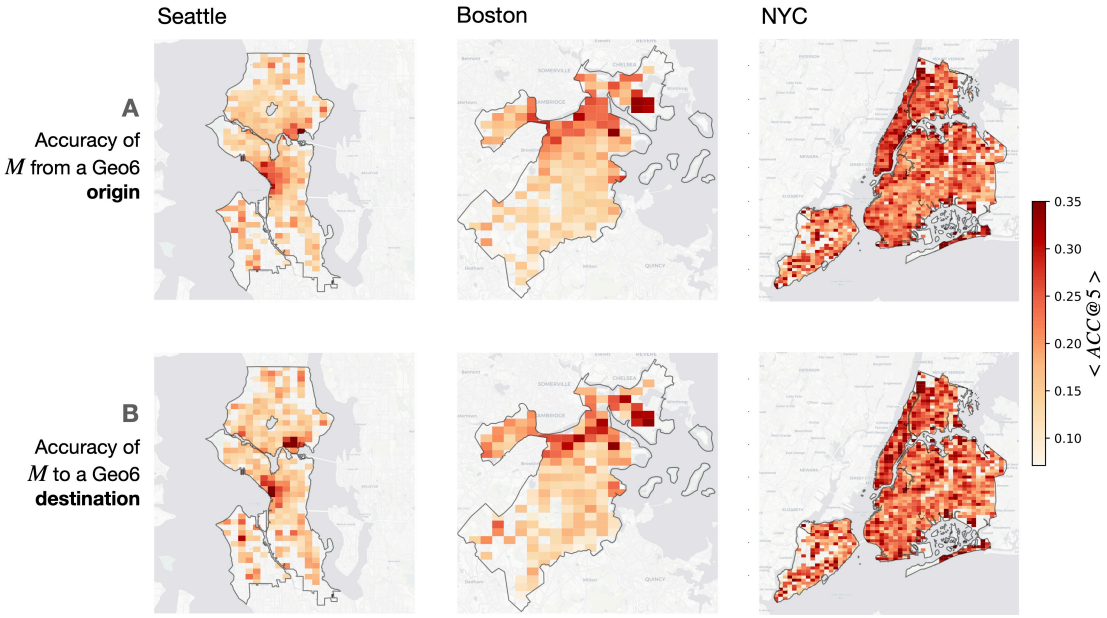

Figure 7.9: Spatial accuracy of M model on novel mobility with pruned C in Seattle, Boston and NYC. The accuracy of M model in predicting a novel transition extracted in the 0-40% overlap (a specific instance never seen in a individual u's historical trajectories $\mathcal{H}^{(u)}$) from an origin in a Geo Hash i is presented in panel **(A)**. In panel **(B)** the accuracy in predicting a transition towards a destination in a Geo Hash tile is depicted. It's worth noting that in the original dataset the sample size T_i used to estimate C_i exhibits strong heterogeneity based on the locations i, with central areas being more densely sampled. To mitigate this potential bias, we employed a stochastic sub sampling process to estimate C_i. Thus, spatial heterogeneity observed in ACC@5_i cannot be attributed to richness of the dataset in specific areas to estimate C_i probabilities. Figure taken from [4]. Maps: Stamen Maps

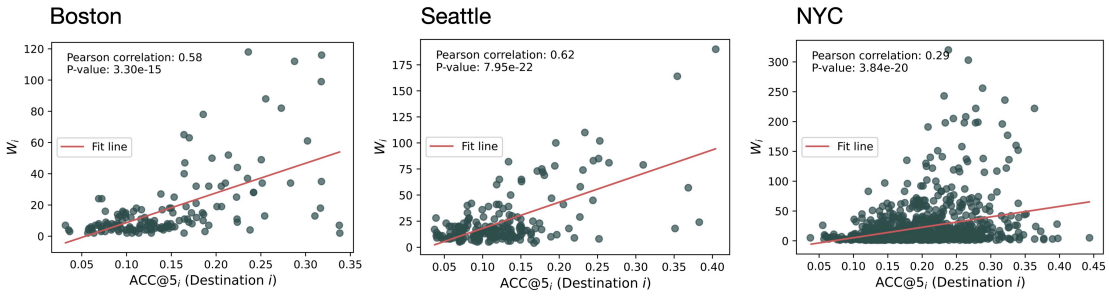

Figure 7.10: Pearson coefficient between $ACC@5_j$ to destination j, and W_i. Correlation between accuracy in predicting a destination location and the number of POIs in that location for Seattle, Boston and New York City. We observe that, specifically for Boston and Seattle, the best predicted destinations in the regime of full novel mobility (never seen transitions in the trajectories in $\mathcal{H}^{(u)}$ in the overlap 0-40%) tend to be the locations having the largest number of POIs. A lower Pearson is observed for New York City, where also locations with lower density of POIs are well predicted as destinations. Figure taken from [4].

7.3.4 Details on pruning process of Collective OD

Data: Set of training transitions D_i for collective OD estimation from origin i
Data: Size of the training set from location i: $T_i = |D_i| = \sum_j T_{ij}$
Data: Percentile X (in the analysis $X = 50\%$)
Result: Pruned collective OD $\bar{C}_i$

Compute CDF and determine percentile X transitions P_X;
foreach *origin i* **do**
  if *T_i is larger than X percentile transitions P_X (T^{max})* **then**
    Sample uniformly P_X transitions from D_i, as sub-sampled set $\bar{D}_i$;
    Estimate new pruned collective probabilities $\bar{C}_i$ using sub sample $\bar{D}_i$;
  end
end

Algorithm 2: Pruning Collective OD Estimation. Pruning algorithm to remove bias of collective OD C_i being estimated with a larger sample size.

7.3.5 Recurrent Neural Network Implementation

Recurrent Neural Networks (RNNs) are commonly used as baselines for tasks in which sequential information (such as time series) is involved. Given a sequence as input, an RNN performs the same task for each element and the output depends on the previous computation. Each computation involves three core parameters:

- x_i the input at the i^{th} step
- h_i the hidden layer at the i^{th} step
- y_i the output of the i^{th} step

There are many different types of recurrent neural networks and for this work, we leveraged an Elman RNN as implemented in [151]. The computation performed by the networks at with n gates are the following:

$$h_i = \sigma_h(W_{h_i} + U_{h_i} h_{i-1} + b_{h_i}), \text{for each } i \in \{1, \ldots, n-1\} \tag{7.1}$$

$$y_i = \sigma_y(W_{h_{n-1}} + b_{n-1}) \tag{7.2}$$

To train the RNN, we fine-tuned the following hyper parameters: learning rate (0.001 for New York, 0.005 for Seattle and Boston), hidden size (750), embedding size (400), and epochs (250 with early stop mechanism). We used Adam as optimizer.

7.3.6 Accuracies Tables

		Full Set				0-20				20-40				40-60				60-80				80-100			
		I	*C*	*M*	RNN	*I*	*C*	*M*	RNN	*I*	*C*	*M*	RNN	*I*	*C*	*M*	RNN	*I*	*C*	*M*	RNN	*I*	*C*	*M*	RNN
Geo 6	NYC	0.608	0.503	0.678	0.649	0.096	0.416	0.376	0.169	0.319	0.398	0.468	0.328	0.572	0.453	0.637	0.599	0.801	0.546	0.817	0.925	0.966	0.698	0.948	0.979
	Boston	0.706	0.643	0.753	0.746	0.093	0.468	0.407	0.155	0.32	0.454	0.492	0.366	0.604	0.537	0.68	0.637	0.831	0.669	0.847	0.928	0.977	0.839	0.929	0.98
	Seattle	0.645	0.549	0.697	0.694	0.073	0.395	0.34	0.148	0.314	0.397	0.461	0.343	0.581	0.471	0.646	0.625	0.808	0.584	0.824	0.936	0.971	0.781	0.918	0.969
Geo 7	NYC	0.475	0.329	0.530	0.494	0.063	0.221	0.222	0.108	0.244	0.229	0.333	0.26	0.508	0.305	0.547	0.529	0.757	0.424	0.768	0.77	0.951	0.616	0.939	0.969
	Boston	0.562	0.446	0.607	0.582	0.07	0.264	0.254	0.103	0.259	0.285	0.364	0.277	0.551	0.394	0.594	0.583	0.787	0.534	0.794	0.801	0.962	0.745	0.919	0.97
	Seattle	0.530	0.412	0.571	0.548	0.05	0.223	0.214	0.099	0.255	0.262	0.344	0.275	0.524	0.368	0.56	0.54	0.764	0.504	0.77	0.779	0.958	0.74	0.916	0.966

Table 7.1: Full accuracies table. Accuracies on full set and overlaps for all Markov models and RNNS for the three cities. Both results for the trajectories with locations mapped to Geo Hash tiles at level 6 and level 7 are reported.

7.3.7 OSM Amenities and POIs retrieval

OpenStreetMap is a collaborative mapping project that provides rich geographical data, including various amenities types and points of interest. The following list outlines the specific types of amenities extracted from OSM for our analyses:

```
amenities:['cafe','college','library','university', 'restaurant','pub','fast
food','bar','bank','pharmacy','arts centre','cinema','community centre','post
office','marketplace']
```

In Chapter 5 number of POIs extracted from OSM in each Geo Hash level 6 tile for each city is presented as a map in Fig. 5.7. Instead in Chapter 4 the number of POIs extracted is mapped to each H3 tile [138] (at resolution 8) in the Greater London Area as W_j.

7.3.8 Dataset details

city	users	# points
Seattle	270K	12M
Boston	375K	11M
NYC	1,5M	140M

Table 7.2: Dataset statistics. Total number of users and number of spatio-temporal points $p = (i, t)$ in the Cuebiq dataset for the cities of Seattle, Boston and New York City. Numbers have been approximated.

Bibliography

[1] Sebastiano Bontorin and Manlio De Domenico. "Multi pathways temporal distance unravels the hidden geometry of network-driven processes". In: *Communications Physics* 6.1 (2023), p. 129.

[2] Pascal P Klamser et al. "Enhancing global preparedness during an ongoing pandemic from partial and noisy data". In: *PNAS nexus* 2.6 (2023), pgad192.

[3] Sebastiano Bontorin et al. "Emergence of complex network topologies from flow-weighted optimization of network efficiency". In: *arXiv preprint arXiv:2301.08661* (2023).

[4] Sebastiano Bontorin et al. "Mixing Individual and Collective Behaviours to Predict Out-of-Routine Mobility". In: *arXiv preprint arXiv:2404.02740* (2024).

[5] Oriol Artime and Manlio De Domenico. *From the origin of life to pandemics: Emergent phenomena in complex systems*. 2022.

[6] *An Introduction to Systems Biology*. Chapman and Hall/CRC, July 2006. DOI: 10.1201/9781420011432. URL: https://doi.org/10.1201/9781420011432.

[7] Roger Guimerà and Luís A. Nunes Amaral. "Functional cartography of complex metabolic networks". In: *Nature* 433.7028 (Feb. 2005), pp. 895–900. DOI: 10.1038/nature03288. URL: https://doi.org/10.1038/nature03288.

[8] S. Maslov and I. Ispolatov. "Propagation of large concentration changes in reversible protein-binding networks". In: *Proceedings of the National Academy of Sciences* 104.34 (Aug. 2007), pp. 13655–13660. DOI: 10.1073/pnas.0702905104. URL: https://doi.org/10.1073/pnas.0702905104.

[9] M. Girvan and M. E. J. Newman. "Community structure in social and biological networks". In: *Proceedings of the National Academy of Sciences* 99.12 (June 2002), pp. 7821–7826. DOI: 10.1073/pnas.122653799. URL: https://doi.org/10.1073/pnas.122653799.

[10] Romualdo Pastor-Satorras and Alessandro Vespignani. *Evolution and Structure of the Internet*. Cambridge University Press, Feb. 2004. DOI: 10.1017/cbo9780511610905. URL: https://doi.org/10.1017/cbo9780511610905.

[11] M. E. J. Newman, Stephanie Forrest, and Justin Balthrop. "Email networks and the spread of computer viruses". In: *Physical Review E* 66.3 (Sept. 2002). DOI: 10.1103/physreve.66.035101. URL: https://doi.org/10.1103/physreve.66.035101.

[12] Riccardo Gallotti et al. "Assessing the risks of 'infodemics' in response to COVID-19 epidemics". In: *Nature human behaviour* 4.12 (2020), pp. 1285–1293.

[13] Christopher W. Lynn and Danielle S. Bassett. "The physics of brain network structure, function and control". In: *Nature Reviews Physics* 1.5 (Mar. 2019), pp. 318–332. DOI: 10.1038/s42254-019-0040-8. URL: https://doi.org/10.1038/s42254-019-0040-8.

[14] Marc Barthélemy. "Spatial networks". In: *Physics Reports* 499.1-3 (Feb. 2011), pp. 1–101. DOI: 10.1016/j.physrep.2010.11.002. URL: https://doi.org/10.1016%2Fj.physrep.2010.11.002.

[15] Oriol Artime et al. *Multilayer network science: from cells to societies*. Cambridge University Press, 2022.

[16] Baruch Barzel and Albert-László Barabási. "Universality in network dynamics". In: *Nature Physics* 9.10 (Sept. 2013), pp. 673–681. DOI: 10.1038/nphys2741. URL: https://doi.org/10.1038/nphys2741.

[17] Chittaranjan Hens et al. "Spatiotemporal signal propagation in complex networks". In: *Nature Physics* 15.4 (Jan. 2019), pp. 403–412. DOI: 10.1038/s41567-018-0409-0. URL: https://doi.org/10.1038/s41567-018-0409-0.

[18] G Caldarelli et al. "The role of complexity for digital twins of cities". In: *Nature Computational Science* 3.5 (2023), pp. 374–381.

[19] Luca Pappalardo et al. "Future directions in human mobility science". In: *Nature computational science* 3.7 (2023), pp. 588–600.

[20] Oriol Artime et al. "Robustness and resilience of complex networks". In: *Nature Reviews Physics* (2024), pp. 1–18.

[21] Chandrakala Meena et al. "Emergent stability in complex network dynamics". In: *arXiv preprint arXiv:2007.04890* (2020). URL: https://arxiv.org/abs/2007.04890.

[22] Arsham Ghavasieh et al. "Multiscale statistical physics of the pan-viral interactome unravels the systemic nature of SARS-CoV-2 infections". In: *Communications Physics* 4.1 (2021), p. 83.

[23] Hugo Barbosa et al. "Human mobility: Models and applications". In: *Physics Reports* 734 (Mar. 2018), pp. 1–74. DOI: 10.1016/j.physrep.2018.01.001. URL: https://doi.org/10.1016/j.physrep.2018.01.001.

[24] Marc Barthelemy. "The statistical physics of cities". In: *Nature Reviews Physics* 1.6 (2019), pp. 406–415.

[25] Armando Bazzani et al. "Statistical laws in urban mobility from microscopic GPS data in the area of Florence". In: *Journal of Statistical Mechanics: Theory and Experiment* 2010.05 (2010), P05001.

[26] Filippo Simini et al. "A universal model for mobility and migration patterns". In: *Nature* 484.7392 (2012), pp. 96–100.

[27] Luís MA Bettencourt and Geoffrey B West. "Bigger cities do more with less." In: *Scientific American* 305.3 (2011), pp. 52–53.

[28] Laura Alessandretti, Ulf Aslak, and Sune Lehmann. "The scales of human mobility". In: *Nature* 587.7834 (2020), pp. 402–407.

[29] D. Brockmann and D. Helbing. "The Hidden Geometry of Complex, Network-Driven Contagion Phenomena". In: *Science* 342.6164 (Dec. 2013), pp. 1337–1342. DOI: 10.1126/science.1245200. URL: https://doi.org/10.1126/science.1245200.

[30] Aihui Pei et al. "Efficiency in the evolution of metro networks". In: *Scientific Reports* 12.1 (May 2022), p. 8326. DOI: 10.1038/s41598-022-12053-3. URL: https://doi.org/10.1038%2Fs41598-022-12053-3.

[31] Chaoming Song et al. "Limits of predictability in human mobility". In: *Science* 327.5968 (2010), pp. 1018–1021.

[32] Markus Schläpfer et al. "The universal visitation law of human mobility". In: *Nature* 593.7860 (2021), pp. 522–527.

[33] Rafael Prieto-Curiel, Jorge E Patino, and Brilé Anderson. "Scaling of the morphology of African cities". In: *Proceedings of the National Academy of Sciences* 120.9 (2023), e2214254120.

[34] Department of Economic United Nations and Population Division Social Affairs. *World Urbanization Prospects: The 2018 Revision*. 2019.

[35] Rafael Prieto-Curiel and Juan P Ospina. "The ABC of mobility". In: *Environment International* 185 (2024), p. 108541.

[36] Michael Szell et al. "Growing urban bicycle networks". In: *Scientific Reports* 12.1 (Apr. 2022), p. 6765. DOI: 10.1038/s41598-022-10783-y. URL: https://doi.org/10.1038/s41598-022-10783-y.

[37] Rafael Prieto Curiel et al. "A paradox of traffic and extra cars in a city as a collective behaviour". In: *Royal Society open science* 8.6 (2021), p. 201808.

[38] Rafael Prieto Curiel, Humberto González Ramírez, and Steven Bishop. "A Ubiquitous Collective Tragedy in Transport". In: *Frontiers in Physics* 10 (2022), p. 882371.

[39] Lukas Ambühl, Monica Menendez, and Marta C González. "Understanding congestion propagation by combining percolation theory with the macroscopic fundamental diagram". In: *Communications Physics* 6.1 (2023), p. 26.

[40] Esteban Moro et al. "Mobility patterns are associated with experienced income segregation in large US cities". In: *Nature Communications* 12.1 (2021), p. 4633.

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

[42] Alan Wilson. *Entropy in Urban and Regional Modelling (Routledge Revivals)*. Routledge, 2013.

[43] Riccardo Gallotti and Marc Barthelemy. "The multilayer temporal network of public transport in Great Britain". In: *Scientific Data* 2.1 (Jan. 2015), pp. 1–8. DOI: 10.1038/sdata.2014.56. URL: https://doi.org/10.1038/sdata.2014.56.

[44] Yihui Ren et al. "Predicting commuter flows in spatial networks using a radiation model based on temporal ranges". In: *Nature Communications* 5.1 (Nov. 2014), p. 5347. DOI: 10.1038/ncomms6347. URL: https://doi.org/10.1038%2Fncomms6347.

[45] Luís MA Bettencourt et al. "Growth, innovation, scaling, and the pace of life in cities". In: *Proceedings of the national academy of sciences* 104.17 (2007), pp. 7301–7306.

[46] Ruiqi Li et al. "Simple spatial scaling rules behind complex cities". In: *Nature Communications* 8.1 (2017), p. 1841.

[47] Matthias Dahlmanns, Franz Kaiser, and Dirk Witthaut. "Optimizing the geometry of transportation networks in the presence of congestion". In: *Physical Review E* 108.4 (2023), p. 044302.

[48] Giorgio Parisi. "Order parameter for spin-glasses". In: *Physical Review Letters* 50.24 (1983), p. 1946.

[49] Daniel L Stein and Charles M Newman. *Spin glasses and complexity*. Vol. 4. Princeton University Press, 2013.

[50] Michele Ballerini et al. "Interaction ruling animal collective behavior depends on topological rather than metric distance: Evidence from a field study". In: *Proceedings of the national academy of sciences* 105.4 (2008), pp. 1232–1237.

[51] Francesco Ginelli. "The physics of the Vicsek model". In: *The European Physical Journal Special Topics* 225 (2016), pp. 2099–2117.

[52] Danielle Smith Bassett and ED Bullmore. "Small-world brain networks". In: *The neuroscientist* 12.6 (2006), pp. 512–523.

[53] Haiyuan Yu and Mark Gerstein. "Genomic analysis of the hierarchical structure of regulatory networks". In: *Proceedings of the National Academy of Sciences* 103.40 (2006), pp. 14724–14731.

[54] Atsushi Tero et al. "Rules for Biologically Inspired Adaptive Network Design". In: *Science* 327.5964 (Jan. 2010), pp. 439–442. DOI: 10.1126/science.1177894. URL: https://doi.org/10.1126/science.1177894.

[55] Junjie Jiang et al. "Predicting tipping points in mutualistic networks through dimension reduction". In: *Proceedings of the National Academy of Sciences* 115.4 (2018), E639–E647.

[56] Pietro Landi et al. "Complexity and stability of ecological networks: a review of the theory". In: *Population ecology* 60.4 (2018), pp. 319–345.

[57] Mark Newman. *Networks*. Oxford University Press, Mar. 2010. DOI: 10.1093/acprof:oso/9780199206650.001.0001. URL: https://doi.org/10.1093/acprof:oso/9780199206650.001.0001.

[58] Albert-László Barabási and Réka Albert. "Emergence of Scaling in Random Networks". In: *Science* 286.5439 (Oct. 1999), pp. 509–512. DOI: 10.1126/science.286.5439.509. URL: https://doi.org/10.1126/science.286.5439.509.

[59] Manlio De Domenico. "Diffusion Geometry Unravels the Emergence of Functional Clusters in Collective Phenomena". In: *Physical Review Letters* 118.16 (Apr. 2017). DOI: 10.1103/physrevlett.118.168301. URL: https://doi.org/10.1103/physrevlett.118.168301.

[60] Arsham Ghavasieh and Manlio De Domenico. "Diversity of information pathways drives sparsity in real-world networks". In: *Nature Physics* (2024), pp. 1–8.

[61] Anna D. Broido and Aaron Clauset. "Scale-free networks are rare". In: *Nature Communications* 10.1 (Mar. 2019). DOI: 10.1038/s41467-019-08746-5. URL: https://doi.org/10.1038/s41467-019-08746-5.

[62] Leto Peel, Tiago P Peixoto, and Manlio De Domenico. "Statistical inference links data and theory in network science". In: *Nature Communications* 13.1 (2022), p. 6794.

[63] Alex Arenas et al. "Synchronization in complex networks". In: *Physics Reports* 469.3 (Dec. 2008), pp. 93–153. DOI: 10.1016/j.physrep.2008.09.002. URL: https://doi.org/10.1016/j.physrep.2008.09.002.

[64] Marián Boguñá et al. "Network geometry". In: *Nature Reviews Physics* 3.2 (Jan. 2021), pp. 114–135. DOI: 10.1038/s42254-020-00264-4. URL: https://doi.org/10.1038/s42254-020-00264-4.

[65] Flavio Iannelli et al. "Effective distances for epidemics spreading on complex networks". In: *Physical Review E* 95.1 (2017), p. 012313.

[66] E. N. Gilbert. "Random Graphs". In: *The Annals of Mathematical Statistics* 30.4 (Dec. 1959), pp. 1141–1144. DOI: 10.1214/aoms/1177706098. URL: https://doi.org/10.1214/aoms/1177706098.

[67] Michael T. Gastner and M. E. J. Newman. "Optimal design of spatial distribution networks". In: *Physical Review E* 74.1 (July 2006), p. 016117. DOI: 10.1103/physreve.74.016117. URL: https://doi.org/10.1103%2Fphysreve.74.016117.

[68] Eleni Katifori, Gergely J. Szöllősi, and Marcelo O. Magnasco. "Damage and Fluctuations Induce Loops in Optimal Transport Networks". In: *Physical Review Letters* 104.4 (Jan. 2010), p. 048704. DOI: 10.1103/physrevlett.104.048704. URL: https://doi.org/10.1103/physrevlett.104.048704.

[69] Marc Barthélemy and Alessandro Flammini. "Optimal traffic networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2006.07 (2006), p. L07002.

[70] Franz Kaiser, Henrik Ronellenfitsch, and Dirk Witthaut. "Discontinuous transition to loop formation in optimal supply networks". In: *Nature Communications* 11.1 (2020), p. 5796.

[71] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. "Understanding individual human mobility patterns". In: *Nature* 453.7196 (2008), pp. 779–782.

[72] Riccardo Gallotti et al. "A stochastic model of randomly accelerated walkers for human mobility". In: *Nature Communications* 7.1 (2016), p. 12600.

[73] Yanyan Xu et al. "Urban dynamics through the lens of human mobility". In: *Nature computational science* 3.7 (2023), pp. 611–620.

[74] A.G. Wilson. "Some new forms of spatial interaction model: A review". In: *Transportation Research* 9.2-3 (July 1975), pp. 167–179. DOI: 10.1016/0041-1647(75)90054-4. URL: https://doi.org/10.1016%2F0041-1647%2875%2990054-4.

[75] Laura Alessandretti et al. "Multimodal urban mobility and multilayer transport networks". In: *Environment and Planning B: Urban Analytics and City Science* 0.0 (July 2022), p. 239980832211081. DOI: 10.1177/23998083221108190. URL: https://doi.org/10.1177/23998083221108190.

[76] Dirk Brockmann, Lars Hufnagel, and Theo Geisel. "The scaling laws of human travel". In: *Nature* 439.7075 (2006), pp. 462–465.

[77] Luca Pappalardo et al. "Returners and explorers dichotomy in human mobility". In: *Nature Communications* 6.1 (2015), p. 8166.

[78] Christian M Schneider et al. "Unravelling daily human mobility motifs". In: *Journal of The Royal Society Interface* 10.84 (2013), p. 20130246.

[79] Chaoming Song et al. "Modelling the scaling properties of human mobility". In: *Nature physics* 6.10 (2010), pp. 818–823.

[80] Filippo Simini et al. "A deep gravity model for mobility flows generation". In: *Nature Communications* 12.1 (2021), p. 6576.

[81] Oriol Cabanas-Tirapu et al. "Human mobility is well described by closed-form gravity-like models learned automatically from data". In: *arXiv preprint arXiv:2312.11281* (2023).

[82] Maxime Lenormand et al. "Comparing and modelling land use organization in cities". In: *Royal Society open science* 2.12 (2015), p. 150449.

[83] Fabiano L Ribeiro and Diego Rybski. "Mathematical models to explain the origin of urban scaling laws". In: *Physics Reports* 1012 (2023), pp. 1–39.

[84] Joao Meirelles et al. "Evolution of urban scaling: Evidence from Brazil". In: *PloS one* 13.10 (2018), e0204574.

[85] Jonathan Norman, Heather L MacLean, and Christopher A Kennedy. "Comparing high and low residential density: life-cycle analysis of energy use and greenhouse gas emissions". In: *Journal of urban planning and development* 132.1 (2006), pp. 10–21.

[86] Elsa Arcaute et al. "Constructing cities, deconstructing scaling laws". In: *Journal of The Royal Society Interface* 12.102 (Jan. 2015), p. 20140745. DOI: 10.1098/rsif.2014.0745. URL: https://doi.org/10.1098/rsif.2014.0745.

[87] Rémi Louf and Marc Barthelemy. "Modeling the polycentric transition of cities". In: *Physical review letters* 111.19 (2013), p. 198702.

[88] Masahisa Fujita and Hideaki Ogawa. "Multiple equilibria and structural transition of non-monocentric urban configurations". In: *Regional science and urban economics* 12.2 (1982), pp. 161–196.

[89] Timur Abbiasov et al. "The 15-minute city quantified using human mobility data". In: *Nature Human Behaviour* (2024), pp. 1–11.

[90] WSN Wan Mohammad et al. "The implication of street network design for walkability: A review". In: *IOP Conference Series: Earth and Environmental Science*. Vol. 881. 1. IOP Publishing. 2021, p. 012058.

[91] Cécile Tannier et al. "Spatial accessibility to amenities in fractal and nonfractal urban patterns". In: *Environment and Planning B: Planning and Design* 39.5 (2012), pp. 801–819.

[92] Michael Batty and Paul A Longley. *Fractal cities: a geometry of form and function*. Academic press, 1994.

[93] Avishai Ceder. "Syncing sustainable urban mobility with public transit policy trends based on global data analysis". In: *Scientific reports* 11.1 (2021), p. 14597.

[94] Romualdo Pastor-Satorras et al. "Epidemic processes in complex networks". In: *Reviews of Modern Physics* 87.3 (Aug. 2015), pp. 925–979. DOI: 10.1103/revmodphys.87.925. URL: https://doi.org/10.1103/revmodphys.87.925.

[95] Juan A. Acebrón et al. "The Kuramoto model: A simple paradigm for synchronization phenomena". In: *Reviews of Modern Physics* 77.1 (Apr. 2005), pp. 137–185. DOI: 10.1103/revmodphys.77.137. URL: https://doi.org/10.1103/revmodphys.77.137.

[96] Xiaozhu Zhang, Dirk Witthaut, Marc Timme, et al. "Topological determinants of perturbation spreading in networks". In: *Physical Review Letters* 125.21 (2020), p. 218301. DOI: 10.1103/physrevlett.125.218301. URL: https://doi.org/10.1103/physrevlett.125.218301.

[97] Malte Schroder et al. "Dynamic Perturbation Spreading in Networks". In: *IEEE Transactions on Network Science and Engineering* 7.3 (July 2020), pp. 1019–1026. DOI: 10.1109/tnse.2019.2901582. URL: https://doi.org/10.1109/tnse.2019.2901582.

[98] Aurélien Gautreau, Alain Barrat, and Marc Barthélemy. "Arrival time statistics in global disease spread". In: *Journal of Statistical Mechanics: Theory and Experiment* 2007.09 (Sept. 2007), pp. L09001–L09001. DOI: 10.1088/1742-5468/2007/09/l09001. URL: https://doi.org/10.1088%2F1742-5468%2F2007%2F09%2Fl09001.

[99] Baruch Barzel and Ofer Biham. "Quantifying the connectivity of a network: The network correlation function method". In: *Physical Review E* 80.4 (Oct. 2009). DOI: 10.1103/physreve.80.046104. URL: https://doi.org/10.1103/physreve.80.046104.

[100] Uzi Harush and Baruch Barzel. "Dynamic patterns of information flow in complex networks". In: *Nature Communications* 8.1 (Dec. 2017). DOI: 10.1038/s41467-017-01916-3. URL: https://doi.org/10.1038/s41467-017-01916-3.

[101] Baruch Barzel, Yang-Yu Liu, and Albert-László Barabási. "Constructing minimal models for complex system dynamics". In: *Nature Communications* 6.1 (May 2015). DOI: 10.1038/ncomms8186. URL: https://doi.org/10.1038/ncomms8186.

[102] Alain Barrat, Marc Barthelemy, and Alessandro Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, 2008. DOI: 10.1017/cbo9780511791383. URL: https://doi.org/10.1017/cbo9780511791383.

[103] Ernesto Estrada and Naomichi Hatano. "Communicability in complex networks". In: *Physical Review E* 77.3 (2008), p. 036111. DOI: 10.1103/physreve.77.036111. URL: https://doi.org/10.1103/PhysRevE.77.036111.

[104] Ronald Aylmer Fisher. "The wave of advance of advantageous genes". In: *Annals of Eugenics* 7.4 (1937), pp. 355–369. DOI: 10.1111/j.1469-1809.1937.tb02153.x. URL: https://doi.org/10.1111/j.1469-1809.1937.tb02153.x.

[105] Michael Worobey et al. "The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic". In: *Science* 377.6609 (2022), pp. 951–959.

[106] Benjamin F Maier and Dirk Brockmann. "Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China". In: *Science* 368.6492 (2020), pp. 742–746.

[107] Manlio De Domenico. "Prevalence of long COVID decreases for increasing COVID-19 vaccine uptake". In: *PLOS Global Public Health* 3.6 (2023), e0001917.

[108] Oliver Eales et al. "Dynamics of competing SARS-CoV-2 variants during the Omicron epidemic in England". In: *Nature Communications* 13.1 (2022), p. 4375.

[109] Pascal P Klamser et al. "Inferring country-specific import risk of diseases from the world air transportation network". In: *PLOS Computational Biology* 20.1 (2024), e1011775.

[110] Stefan Elbe and Gemma Buckland-Merrett. "Data, disease and diplomacy: GISAID's innovative contribution to global health". In: *Global challenges* 1.1 (2017), pp. 33–46.

[111] Stephen W Attwood et al. "Phylogenetic and phylodynamic approaches to understanding and combating the early SARS-CoV-2 pandemic". In: *Nature Reviews Genetics* 23.9 (2022), pp. 547–562.

[112] Christophe Fraser. "Estimating individual and household reproduction numbers in an emerging epidemic". In: *PloS one* 2.8 (2007), e758.

[113] Valentina Morandi. "Bridging the user equilibrium and the system optimum in static traffic assignment: a review". In: *4OR* (2023), pp. 1–31.

[114] Chong Wei, Yasuo Asakura, and Takamasa Iryo. "Formulating the within-day dynamic stochastic traffic assignment problem from a Bayesian perspective". In: *Transportation Research Part B: Methodological* 59 (2014), pp. 45–57.

[115] Camille Roth et al. "A long-time limit for world subway networks". In: *Journal of The Royal Society Interface* 9.75 (May 2012), pp. 2540–2550. DOI: 10.1098/rsif.2012.0259. URL: https://doi.org/10.1098/rsif.2012.0259.

[116] Minjin Lee et al. "Morphology of travel routes and the organization of cities". In: *Nature Communications* 8.1 (Dec. 2017), p. 2229. DOI: 10.1038/s41467-017-02374-7. URL: https://doi.org/10.1038/s41467-017-02374-7.

[117] Riccardo Gallotti, Pierluigi Sacco, and Manlio De Domenico. "Complex Urban Systems: Challenges and Integrated Solutions for the Sustainability and Resilience of Cities". In: *Complexity* 2021 (Oct. 2021). Ed. by Rosa M. Benito, pp. 1–15. DOI: 10.1155/2021/1782354. URL: https://doi.org/10.1155/2021/1782354.

[118] R. G. Morris and M. Barthelemy. "Transport on coupled spatial networks". In: *Physical Review Letters* 109.12 (Sept. 2012), p. 128703. DOI: 10.1103/physrevlett.109.128703. URL: https://doi.org/10.1103/physrevlett.109.128703.

[119] Rémi Louf, Pablo Jensen, and Marc Barthelemy. "Emergence of hierarchy in cost-driven growth of spatial networks". In: *Proceedings of the National Academy of Sciences* 110.22 (May 2013), pp. 8824–8829. DOI: 10.1073/pnas.1222441110. URL: https://doi.org/10.1073%2Fpnas.1222441110.

[120] Jayanth R. Banavar, Amos Maritan, and Andrea Rinaldo. "Size and form in efficient transportation networks". In: *Nature* 399.6732 (May 1999), pp. 130–132. DOI: 10.1038/20144. URL: https://doi.org/10.1038/20144.

[121] Erwan Taillanter and Marc Barthelemy. "Evolution of road infrastructure in large urban areas". In: *Physical Review E* 107.3 (2023), p. 034304.

[122] Siddharth Patwardhan et al. "Symmetry breaking in optimal transport networks". In: *arXiv preprint arXiv:2311.05059* (2023).

[123] Mark J Nieuwenhuijsen. "Urban and transport planning pathways to carbon neutral, liveable and healthy cities; A review of the current evidence". In: *Environment international* 140 (2020), p. 105661.

[124] Abdullahi Adinoyi Ibrahim, Alessandro Lonardi, and Caterina De Bacco. "Optimal Transport in Multilayer Networks for Traffic Flow Optimization". In: *Algorithms* 14.7 (June 2021), p. 189. DOI: 10.3390/a14070189. URL: https://doi.org/10.3390/a14070189.

[125] Xiaoge Zhang et al. "A Biologically Inspired Network Design Model". In: *Scientific Reports* 5.1 (June 2015), pp. 1–14. DOI: 10.1038/srep10794. URL: https://doi.org/10.1038/srep10794.

[126] Reza Zanjirani Farahani et al. "A review of urban transportation network design problems". In: *European journal of operational research* 229.2 (2013), pp. 281–302.

[127] Rémi Louf and Marc Barthelemy. "How congestion shapes cities: from mobility patterns to scaling". In: *Scientific Reports* 4.1 (July 2014), p. 5561. DOI: 10.1038/srep05561. URL: https://doi.org/10.1038%2Fsrep05561.

[128] Colin P.D. Birch, Sander P. Oom, and Jonathan A. Beecham. "Rectangular and hexagonal grids used for observation, experiment and simulation in ecology". In: *Ecological Modelling* 206.3-4 (Aug. 2007), pp. 347–359. DOI: 10.1016/j.ecolmodel.2007.03.041. URL: https://doi.org/10.1016/j.ecolmodel.2007.03.041.

[129] Matheus P. Viana et al. "The simplicity of planar networks". In: *Scientific Reports* 3.1 (Dec. 2013), p. 3495. DOI: 10.1038/srep03495. URL: https://doi.org/10.1038/srep03495.

[130] Telikepalli Kavitha et al. "An Algorithm for Minimum Cycle Basis of Graphs". In: *Algorithmica* 52.3 (Oct. 2007), pp. 333–349. DOI: 10.1007/s00453-007-9064-z. URL: https://doi.org/10.1007/s00453-007-9064-z.

[131] Yosef Sheffi and Warren Powell. "A comparison of stochastic and deterministic traffic assignment over congested networks". In: *Transportation Research Part B: Methodological* 15.1 (1981), pp. 53–64.

[132] Malvin H Kalos and Paula A Whitlock. *Monte Carlo Methods*. John Wiley & Sons, 2009.

[133] Duccio Piovani, Carlos Molinero, and Alan Wilson. "Urban retail location: Insights from percolation theory and spatial interaction modeling". In: *PLOS ONE* 12.10 (Oct. 2017). Ed. by Li Daqing, e0185787. DOI: 10.1371/journal.pone.0185787. URL: https://doi.org/10.1371/journal.pone.0185787.

[134] OpenStreetMap contributors. *Planet dump retrieved from https://planet.osm.org*. https://www.openstreetmap.org. 2017.

[135] S. E. Dreyfus and R. A. Wagner. "The Steiner problem in graphs". In: *Networks* 1.3 (1971), pp. 195–207. DOI: 10.1002/net.3230010302. URL: https://doi.org/10.1002/net.3230010302.

[136] Marcus Brazil et al. "On the history of the Euclidean Steiner tree problem". In: *Archive for History of Exact Sciences* 68.3 (Aug. 2013), pp. 327–354. DOI: 10.1007/s00407-013-0127-z. URL: https://doi.org/10.1007/s00407-013-0127-z.

[137] César A. Hidalgo, Elisa Castañer, and Andres Sevtsuk. "The amenity mix of urban neighborhoods". In: *Habitat International* 106 (Dec. 2020), p. 102205. DOI: 10.1016/j.habitatint.2020.102205. URL: https://doi.org/10.1016%2Fj.habitatint.2020.102205.

[138] *Uber H3, 2024 Uber Technologies*. https://h3geo.org/.

[139] Minjin Lee and Petter Holme. "Relating land use and human intra-city mobility". In: *PloS one* 10.10 (2015), e0140152.

[140] Kasthuri Jayarajah, Andrew Tan, and Archan Misra. "Understanding the interdependency of land use and mobility for urban planning". In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. 2018, pp. 1079–1087.

[141] Yiannis Kamarianakis and Poulicos Prastacos. "Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches". In: *Transportation Research Record* 1857.1 (2003), pp. 74–84.

[142] Pu Wang et al. "Understanding road usage patterns in urban areas". In: *Scientific reports* 2.1 (2012), p. 1001.

[143] Amy Wesolowski et al. "Quantifying the impact of human mobility on malaria". In: *Science* 338.6104 (2012), pp. 267–270.

[144] Moritz UG Kraemer et al. "The effect of human mobility and control measures on the COVID-19 epidemic in China". In: *Science* 368.6490 (2020), pp. 493–497.

[145] Massimiliano Luca et al. "Crime, inequality and public health: a survey of emerging trends in urban data science". In: *Frontiers in Big Data* 6 (2023), p. 1124526.

[146] Lorenzo Lucchini et al. "Living in a pandemic: changes in mobility routines, social activity and adherence to COVID-19 protective measures". In: *Scientific reports* 11.1 (2021), pp. 1–12.

[147] Massimiliano Luca et al. "A survey on deep learning for human mobility". In: *ACM Computing Surveys (CSUR)* 55.1 (2021), pp. 1–44.

[148] Ayele Gobezie Chekol and Marta Sintayehu Fufa. "A survey on next location prediction techniques, applications, and challenges". In: *EURASIP Journal on Wireless Communications and Networking* 2022.1 (2022), p. 29.

[149] Alexandre De Brébisson et al. "Artificial neural networks applied to taxi destination prediction". In: *Proceedings of the 2015th International Conference on ECML PKDD Discovery Challenge*. 2015, pp. 40–51.

[150] Qiang Liu et al. "Predicting the next location: A recurrent model with spatial and temporal contexts". In: *Thirtieth AAAI conference on artificial intelligence*. 2016.

[151] Jie Feng et al. "Deepmove: Predicting human mobility with attentional recurrent networks". In: *Proceedings of the 2018 world wide web conference*. 2018, pp. 1459–1468.

[152] Sébastien Gambs, Marc-Olivier Killijian, and Miguel Núñez del Prado Cortez. "Show me how you move and I will tell you who you are". In: *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS*. 2010, pp. 34–41.

[153] Sébastien Gambs, Marc-Olivier Killijian, and Miguel Núñez del Prado Cortez. "Next place prediction using mobility Markov chains". In: *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*. 2012, pp. 1–6.

[154] Francesco Calabrese, Giusy Di Lorenzo, and Carlo Ratti. "Human mobility prediction based on individual and collective geographical preferences". In: *13th international IEEE conference on intelligent transportation systems*. IEEE. 2010, pp. 312–317.

[155] Kamil Smolak et al. "The impact of human mobility data scales and processing on movement predictability". In: *Scientific Reports* 11.1 (2021), pp. 1–10.

[156] Jameson L Toole et al. "Tracking employment shocks using mobile phone data". In: *Journal of The Royal Society Interface* 12.107 (2015), p. 20150185.

[157] Simone Centellegher et al. "The long-term and disparate impact of job loss on individual mobility behaviour". In: *arXiv preprint arXiv:2403.10276* (2024).

[158] Nils Haug et al. "Ranking the effectiveness of worldwide COVID-19 government interventions". In: *Nature Human Behaviour* 4.12 (2020), pp. 1303–1312.

[159] Massimiliano Luca et al. "Trajectory test-train overlap in next-location prediction datasets". In: *Machine Learning* 112.11 (2023), pp. 4597–4634.

[160] Winter Mason and Duncan J Watts. "Collaborative learning in networks". In: *Proceedings of the National Academy of Sciences* 109.3 (2012), pp. 764–769.

[161] Markku Kaustia and Samuli Knüpfer. "Peer performance and stock market entry". In: *Journal of Financial Economics* 104.2 (2012), pp. 321–338.

[162] Stephen Coleman. "The effect of social conformity on collective voting behavior". In: *Political analysis* 12.1 (2004), pp. 76–96.

[163] Joshua Becker, Ethan Porter, and Damon Centola. "The wisdom of partisan crowds". In: *Proceedings of the National Academy of Sciences* 116.22 (2019), pp. 10717–10722.

[164] Feng Shi et al. "The wisdom of polarized crowds". In: *Nature human behaviour* 3.4 (2019), pp. 329–336.

[165] James Surowiecki. *The wisdom of crowds*. Anchor, 2005.

[166] Salvatore Rinzivillo et al. "The purpose of motion: Learning activities from individual mobility networks". In: *2014 international conference on data science and advanced analytics (DSAA)*. IEEE. 2014, pp. 312–318.

[167] Zexun Chen et al. "Contrasting social and non-social sources of predictability in human mobility". In: *Nature Communications* 13.1 (2022), p. 1922.

[168] Luca Pappalardo et al. "An analytical framework to nowcast well-being using mobile phone data". In: *International Journal of Data Science and Analytics* 2 (2016), pp. 75–92.

[169] Ramaswamy Hariharan and Kentaro Toyama. "Project Lachesis: Parsing and Modeling Location Histories". In: *Geographic Information Science*. Ed. by Max J. Egenhofer, Christian Freksa, and Harvey J. Miller. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 106–124. ISBN: 978-3-540-30231-5. DOI: {10.1007/978-3-540-30231-5_8}.

[170] Martin Ester et al. "A density-based algorithm for discovering clusters in large spatial databases with noise." In: *kdd*. Vol. 96. 34. 1996, pp. 226–231.

[171] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[172] Yanguang Chen. "A new methodology of spatial cross-correlation analysis". In: *PLoS One* 10.5 (2015), e0126158.

[173] Luc Anselin. "Local indicators of spatial association—LISA". In: *Geographical analysis* 27.2 (1995), pp. 93–115.

[174] Juanjuan Zhang et al. "The impact of relaxing interventions on human contact patterns and SARS-CoV-2 transmission in China". In: *Science Advances* 7.19 (2021), eabe2584.

[175] S. Flaxman et al. "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe". In: *Nature* 584.7820 (2020), pp. 257–261.

[176] Matteo Chinazzi et al. "The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak". In: *Science* 368.6489 (2020), pp. 395–400.

[177] Takahiro Yabe et al. "Behavioral changes during the COVID-19 pandemic decreased income diversity of urban encounters". In: *Nature Communications* 14.1 (2023), p. 2310.

[178] Ludovico Napoli et al. "Socioeconomic reorganization of communication and mobility networks in response to external shocks". In: *Proceedings of the National Academy of Sciences* 120.50 (2023), e2305285120.

[179] Daniel Delahaye, Supatcha Chaimatanan, and Marcel Mongeau. "Simulated annealing: From basics to applications". In: *Handbook of metaheuristics* (2019), pp. 1–35.