Mesoscale Building Blocks of Pedestrian Mobility: a Discrete Vector Field Approach

Author: Robert Benassai Dalmau[1,2*]
[1] *Internet Interdisciplinary Institute, Universitat Oberta de Catalunya, Rambla del Poblenou, 156, 08018 Barcelona, Spain.*
[2] *Master en Física dels Sistemes Complexos i Biofísica. Facultat de Física, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain.*[†]

Advisors: Javier Borge-Holthoefer[1], Albert Solé-Ribalta[1] and Josep Perelló Palou[2]
[1] *Internet Interdisciplinary Institute, Universitat Oberta de Catalunya, Rambla del Poblenou, 156, 08018 Barcelona, Spain.*
[2] *Master en Física dels Sistemes Complexos i Biofísica. Facultat de Física, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain.*
(Dated: June 30, 2023)

Abstract: Understanding and characterising pedestrian mobility is crucial to develop sustainable cities. While classical statistical analysis and diffusion models are commonly used to analyze human trajectories either at the microscopic (e.g. sidewalk flows) or macroscopic scale (e.g. origin-destination matrices), they may not be suitable for capturing the nuances and intricacies of mobility patterns at the mesoscale. To overcome these limitations, the problem is approached by leveraging on vector field theory with the aim to describe how the urban geometry and structure of sidewalk networks affect pedestrian mobility flows. Considering the particularities of pedestrian movement (e.g. limited travel range) the discrete- (DTRW) and continuous-time (CTRW) random walk dynamics have been implemented to retrieve a baseline agent-based net flow along the edges of pedestrian networks with a temporal budget of mobility. These flows are subsequently interpreted as discrete vector fields. The Helmholtz-Hodge decomposition (HHD) allows the partition of vector fields into three well-defined patterns: cyclic (solenoidal and harmonic) and divergent/convergent (gradient) components. Results show that when mobility is agnostic to edge lengths (DTRW), that is, when the time budget is spent equivalently along the edges (steps), high-density regions with larger degree nodes show attractiveness, as existing literature already describes. However, when the time budget is spent proportionally to the edge lengths (CTRWs), the same regions show a repulsive effect. Intermediate regimes arise as well in the continuum between these two processes. An analytical description of both DTRWs and CTRWs has been developed to accurately estimate the gradient components of the vector fields. However, the presented deterministic developments do not predict the presence of the cyclic components as they seem to emerge from the stochasticity of the process. To validate this idea, the variance of the cyclic component, or its mean squared flow (MSF), has been analysed. Results show that the MSF grows linearly with the temporal budget of the walkers. This behaviour is similar to the characteristic linear temporal evolution of the Mean-Squared Displacement (MSD) in random walks and Brownian motion. Ultimately, this work contributes to the existing description and understanding of the behaviour of different random walk dynamics on spatially embedded graphs, providing a baseline to understand and analyse pedestrian mobility on sidewalk networks in future works.

I. INTRODUCTION

Cities evolved, and continue evolving, into different organizational patterns due to historical, political or financial circumstances and continuous optimization [1, 2]. As a paradigmatic example of complex systems, they exhibit many of their key properties, as they can be fundamentally characterized by a large number of interacting components or agents (citizens, cars, power networks, etc) that give rise to emergent properties (social interaction, congestion, robustness) that cannot be easily understood by studying their constituents at the individual level.

Many of the numerous elements that compose cities, such as buildings, infrastructure, people, and social institutions, rely on a physical o virtual substrate for their interaction. Usually, this substrate takes the mathematical form of a graph and the system ensembles a complex system. Among these many intertwined layouts, transportation networks stand out. They become fundamental for the movement of people, goods, and services within and between different parts of a city [3].

From the origin of cities and civilization there has always been a very tight binding between urban structure and their land use [4]. The location and capacity of roads, highways, and public transit systems can mold the distribution of residential, commercial, and industrial areas and vice-versa.

Understanding the binding between the underlying structure and the emerging dynamics of transport is crucial. In either data- or simulation-driven analyses, it is hard to disentangle whether the observed patterns are a consequence of pedestrian/driver behaviour, or they

* rbenasda7.alumnes@ub.edu
† master.complex.biophys@ub.edu

can be explained solely by the structure on which these pedestrians/drivers move.

In response to the growing threat of climate change and its impact on urban areas, cities across the globe have increasingly recognized the need to incorporate climate considerations into their policy and planning decisions. On the mobility front, increasing active travel, and walking in particular, is widely accepted as critical to achieving sustainable cities.

In order to reach this goal, pedestrian networks need to live up to the needs of the citizens, both in the quality of the sidewalks themselves and the structure of the pedestrian network. In this last regard, the analysis of this work is framed on how the urban infrastructure and sidewalk network topology and geometry affect pedestrian mobility and its dynamics.

Usually, mobility simulations in urban transport networks are often intricate and time-consuming due to the complexity of the network's rules and the diverse behaviour and destinations of the passengers. Many models have been presented over the years, from fluid dynamics [5], gravity [6] and radiation models [7] to agent-based models such as many variations of random walks [8–10].

In this work, the aim is to isolate the structural and geometrical effects of the network on pedestrian dynamics. For this reason, the chosen approach is to use an unbiased agent-based model for the walking agents, at the expense of realism, such that there are no intrinsical preferences and artefacts coming from the implemented dynamics. Consequently, Markovian discrete- and continuous-time random walk (DTRW and CTRW) models [11, 12] have been used to simulate pedestrian behaviour.

From the mobility data front, many works use vector field representations to study mobility patterns [13–15]. This analysis can be carried out using continuous, defined at each point in space, or discrete, restricted to an underlying structure such as a network, vector fields.

In this regard, the Helmholtz-Hodge Decomposition (HHD) has proven to be a useful approach to analyse these vector fields. This technique is able to decompose any flow, both in continuous or discrete spaces (for instance urban networks), into three orthogonal components that are informative of cyclic (solenoidal and harmonic components) and convergent/divergent (gradient component) potential flows. In mobility networks, the HHD may provide information about which regions of a city promote (and to which extent) either cyclical or convergent/divergent pedestrian behaviour.

This description of continuous and discrete vector fields is useful to simplify their analysis by directly working on the components that intrinsically describe properties such as incompressibility or vorticity. This technique has been traditionally applied to computational fluid dynamics and computational physics to visualize and analyse vector fields [16].

Other branches of science also use the HHD to determine the nature of different types of flows. For instance, [17] applied the decomposition to brain networks while [18] used the HHD to describe information flow on small-world networks. Both works characterize the importance of each component through the conservation of the $l^2 - norm$ of the total edge flow due to the orthogonality of the decomposition.

The HHD has already been applied in mobility networks via Origin-Destination (OD) matrices, which represent the flow of people, goods, or vehicles between different origins and destinations within a geographic area. In literature, [14] converted the OD matrix into a 2D vector field by averaging all trips starting from each location, getting an averaged field at each starting point. The resulting flow was found to be almost irrotational, and a scalar potential was introduced to generate human flows and reproduce the empirical data. However, this approach deletes the origin-destination information inherent in OD matrices due to the averaging of the trips at each point. A posterior article, [15], applied the Hodge decomposition to synthetic data to prove that, using the mentioned continuous average vector field from OD matrices, it was difficult to identify the number of city centres and their area. On the other hand, this article showed that by applying directly the discrete HHD to the OD matrix represented as a graph, a more detailed view of the city centre and its area could be obtained. Applying this framework to many US metropolitan areas it was seen that different regions showed gradient or curl dominances, but the driving mechanism of these changes of dominance is still an open question.

Despite their popularity, OD matrices only provide information on the mesoscale of the city, disregarding the complex connective structure of its streets and intersections. Consequently, they cannot provide a detailed description of the agent flows nor contain precise structural and geometrical correlations. The Hodge decomposition has barely been applied to urban networks that are not represented by OD matrices. This leads to the question of whether the results in [14] and [15] are particular to OD matrix flows or are indeed also applicable to a detailed urban network. Kan and López [19] applied an adapted version of the decomposition to the London Underground (LU) network to infer link flows using OD matrices. This is one of the few works applied to a physically existing network, although, this study still relied on OD matrices.

In this work, a discrete vector field has been obtained from simulating discrete- and continuous-time random walks. Under this framework, the net edge flow has been obtained by assigning a directionality to the links of the pedestrian network and making the difference between forward and backward traversals during the simulation time. This approach yields a vector field on a discrete domain representing the random walk flow across the network. Applying the HHD to the resulting flows, the aim is to observe if any of the components stand out due to the network's geometry and structure, revealing its core properties.

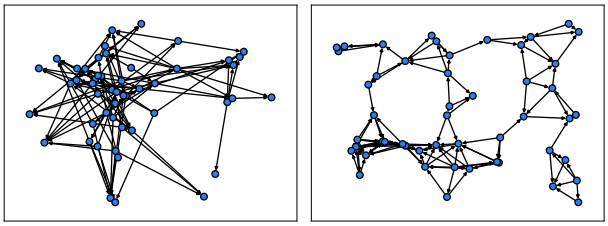
II. A BRIEF OVERVIEW ON GRAPH THEORY

A graph, $G(E, V)$, is a mathematical structure that consists of a set of vertices, V, (also known as nodes) and a set of edges, E, (also known links) that connect pairs of vertices [20]. Graphs are used to represent relationships or connections between objects or entities. Usually, graphs are described by the adjacency matrix $\boldsymbol{A}$, which captures the connections between vertices and is $a_{ij} = 1$ if nodes i and j are connected or $a_{ij} = 0$ otherwise. The number of links per node is quantified by the degree, denoted by $k_i = 0, 1, 2, \ldots$. Graphs can be weighted or unweighted. In weighted graphs, links can carry values representing given characteristics such as the width of a road or the importance of a link. In this work, the weight on the edges will represent the flow along that given link. The edges of a graph can be directed (one-way relationships) or undirected (symmetric relationships). In the case studied, dynamic processes will be running on undirected graphs. These concepts provide a foundation for analyzing complex systems in various domains, from computer science to social networks and beyond.

A. Synthetic Graphs and Toy Models

For validation purposes, the Random Geometric (RG) and the Erdös-Rényi (ER) models (Figure 1), [21], have been utilized, which are both random graphs. In the ER model, each node has a probability p to be connected to any other node on the graph. As the ER case does not have a spatial embedding, the spring-layout method [22, 23] has been used. This spatial embedding (Fruchterman-Reingold force-directed algorithm) simulates a force-directed representation of the network treating edges as springs holding nodes together. Conversely, nodes are treated as repelling objects. The simulation continues until the positions are close to equilibrium. In the case of the RG model, n nodes are placed uniformly at random in a unit square. Then, two nodes are joined by an edge if the distance between the nodes is at most r, a given radius.

In order to observe clear effects of the node degree and edge distances, a periodic boundary conditions (PBC) lattice has been modified such that one of the squares in the lattice contains a larger number of nodes with a higher degree. This model is useful because it contains a densely connected cluster of nodes, which can be seen as a simplified version of a dense city centre. The outer region is regular and every node has the same degree and average neighbouring link distances. The inner modified part contains nodes of different and higher degrees than the ones in the outer region, and the nodes are closer to each other.



(a) Erdös-Rényi graph with 50 nodes and $p = 0.1$.

(b) Random geometric graph with 50 nodes and $r = 0.2$.

FIG. 1: Examples of the two random graph models considered.

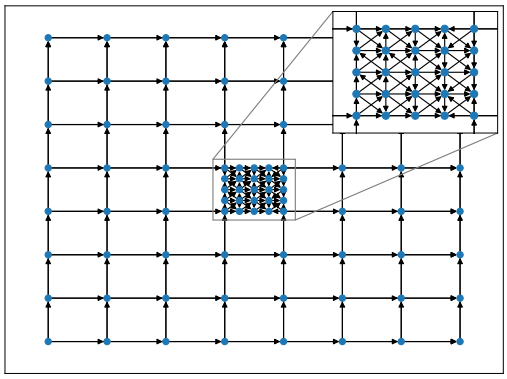


FIG. 2: Modified lattice with periodic boundary conditions (PBC). Edges are transitable in both directions, the arrows only set the positive orientation of the edge flow.

III. DECOMPOSING DISCRETE AND CONTINUOUS VECTOR FIELDS: THE HELMOLTZ-HODGE DECOMPOSITION

Typically, vector fields represent a magnitude that has a given direction at each point in space (fluid flows, electric and magnetic fields, etc.). This definition corresponds to continuous vector fields, which are the most common in nature because space is inherently continuous. Despite this, in some cases, vector fields can be defined in a discrete domain, where positions and directions are limited by an underlying structure, usually a network. In discrete vector fields, every link of a network is considered to be an element of the field and it is represented by an arrow. This arrow has a fixed direction, determined by the graph, and a magnitude which represents the value of the field.

Under the Helmholtz-Hodge decomposition (HHD), both continuous (Helmholtz decomposition) and discrete (Hodge decomposition) vector fields can be partitioned into two or three orthogonal components that represent their divergent (presence of sources or sinks) and rotational behaviour.

A. Decomposition of Continuous Vector Fields: The Helmholtz Decomposition

The Helmholtz decomposition, also known as the fundamental theorem of vector calculus [24], was the original decomposition for continuous vector fields, which was later generalized under the Hodge theory and the Hodge decomposition. This well-understood theorem of vector calculus states that any well-behaved fast-decaying three-dimensional vector field can be decomposed into a curl-free and a divergence-free component. This decomposition is illustrated in Figure 3.

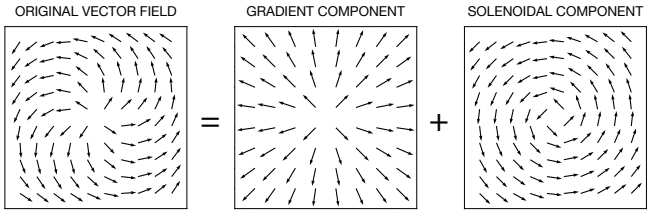


FIG. 3: Helmholtz decomposition of a vector field

To compute the Helmholtz decomposition, the essential vector operators are needed: divergence, gradient and curl. The divergence represents the volume density of the outward flux of a vector field calculated at an infinitesimal point. On the other hand, the gradient of a scalar field represents the direction in which it increases most quickly, central panel in Figure 3. Finally, the curl describes the infinitesimal circulation of a vector field. It is calculated with the circulation of the vector field along a closed loop divided by the area bounded inside it. Then, the limit to an infinitesimal loop is taken yielding the curl at a given point. It is important to mention that the divergence of a field is a scalar quantity while the curl and the gradient retrieve vector quantities. Mathematically, the curl-free component, $\boldsymbol{\nabla}\Phi$ in equation (1), (irrotational or gradient component) is, by definition, a potential field, which can be calculated by the gradient of a scalar field (scalar potential) Φ. The divergence-free (solenoidal) component is obtained from the curl of a given vector field (vector potential) $\boldsymbol{A}$. Consequently, equation (1) performs the Helmholtz decomposition of a vector field $\mathbf{F}$.

$$\boldsymbol{F} = -\boldsymbol{\nabla}\Phi + \boldsymbol{\nabla} \times \boldsymbol{A} \tag{1}$$

Additionally, the harmonic component can be introduced to the Helmholtz decomposition. Necessarily, this new component must be both irrotational and divergence-free. In other words, if the harmonic field is irrotational, it must come from the gradient of a potential. At the same time, if it is divergence-free, the divergence of the gradient of the scalar field, which is the definition of its Laplacian, must be 0.

$$\dot{\nabla}(\boldsymbol{\nabla}\lambda) = \nabla^2\lambda = 0 \tag{2}$$

Consequently, the resulting harmonic field cannot rotate nor converge or diverge, therefore it is constant in one direction.

Generalizing this vector calculus theorem, the Hodge decomposition aims to apply these concepts to the discrete domain, namely, on graphs.

B. Decomposition of discrete vector fields: The Hodge Decomposition

Mathematically, the weights of a graph can be embedded in many dimensional structures (nodes, edges, triangles and etc.) called *k-dimensional* simplices. Each of these collections forms a subspace of the graph, the dimension of which is $N-1$ where N is the number of nodes of each structure (for more detail on this description see the Appendix and [25]). In this work, the considered weights (mobility flows) will be defined on the *edge-subspace* $\langle\Sigma_1\rangle \equiv E$. In other words, each edge of the network will have its corresponding weight, and the Hodge decomposition will be presented under this description. As mentioned, these weights will represent the net flow ω of individuals through the sidewalks of a pedestrian network. The Hodge decomposition partitions the edge subspace $\langle\Sigma_1\rangle$ into three orthogonal subspaces:

$$\langle\Sigma_1\rangle = G_1 \oplus H_1 \oplus S_1 \tag{3}$$

where G_1, S_1 and H_1 are the gradient, solenoidal and harmonic subspaces and $\oplus$ represents their direct sum. Consequently, the net pedestrian flow on each link of the network, ω, can be expressed as the sum of the gradient component $\omega_g \in G_1$, the harmonic component $\omega_h \in H_1$ and the solenoidal component $\omega_s \in S_1$.

$$\omega = \omega_g + \omega_h + \omega_s \tag{4}$$

Since the Hodge decomposition is the discrete version of the Helmholtz decomposition, the discrete analogues of the continuum vector operators (divergence, gradient and curl) need to be defined. To achieve this, as the underlying graphs are undirected, links need to have a preset arbitrary directionality to determine whether the flow along a link is positive or negative depending on its sense. In other words, the undirected links need to be replaced by directed edges. Additionally, this procedure will avoid double counting, as edge (i,j) is the same as (j,i) in an undirected graph. As the associated edge sense is arbitrary (sidewalks are bidirectional), edges and triangles will be labelled in the ascending order as a reference, $(\overline{i_1,i_2})$ for edges and $(\overline{i_1,i_2,i_3})$ for triangles, where the overline implies that $i_1 < i_2 < i_3$.

Any odd permutation of the indices will be considered a negative orientation of the reference labelling, while the even permutations will be positive. Consequently, in the simplest case of an edge, if ω_{21} is the flow along the edge $(2,1)$, in the reference labelling it will be expressed as $\omega_{21}\,(2,1) = -\omega_{21}\,(\overline{1,2})$. More generally:

$$a_{i_1\dots i_k}(i_1\dots i_k) \equiv \epsilon_{i_1\dots i_k} a_{i_1\dots i_k}(\overline{j_1\dots j_k}) \tag{5}$$

Where, in this work, $k \leq 3$.

With this notation, the discrete vector operators to perform the decomposition can be defined. In the continuum, the divergence measures the net outgoing flux of the vector field at any given point. Since, in the continuum, this operator maps a vector field to a scalar field, in the discrete space the divergence will map edge weights to node weights, and it will be a node property. Consequently, the divergence of a node is defined as the outgoing minus the incoming edge flow. When a node has positive divergence it is considered a source of flow and, when the divergence is negative, that node is a sink. If $\omega_{\overline{i_1 i_2}}$ is the flow associated to the edge $(\overline{i_1 i_2})$, then, the divergence of a node (i_1) is:

$$div(i_1) = \sum_{j>i_1:(\overline{i_1 j})\in\langle\Sigma_1\rangle} \omega_{\overline{i_1 j}} - \sum_{j<i_1:(\overline{j i_1})\in\langle\Sigma_1\rangle} \omega_{\overline{j i_1}} \tag{6}$$

To define the gradient operator, the analogue of a continuous scalar field is needed. In scalar fields, every point in space has its field value associated. In the discrete framework, the points in "space" are the nodes and, in consequence, the discrete analogue of a scalar field is a set of node values, which can represent, for instance, the population density at several areas of a city. Ultimately, the gradient operator applied on a graph performs the difference in node weights between every pair of adjacent nodes. Considering the set of nodes, $(j_1, j_2, ..., j_n) \in \langle\Sigma_0\rangle \equiv V$ where $\langle\Sigma_0\rangle$ is the subspace of nodes, with their corresponding weights $(a_{j_1}, a_{j_2}, ..., a_{j_n})$ expressed as a discrete scalar field η:

$$\eta = \sum_{j\in\langle\Sigma_0\rangle} a_j\,(j)$$

the gradient of η is given by the gradient operator δ_0:

$$\delta_0(\eta) = \sum_{(\overline{ij})\in\langle\Sigma_1\rangle} \left[a_j - a_i\right](\overline{ij}) \tag{7}$$

which yields a set of edges $\{(i,j)\}$ and the gradient between the nodes of each link as their weights $\left[a_j - a_i\right]$ (a vector field in the continuum).

Now that the divergence and gradient operators have been defined in the discrete domain, the gradient component, ω_g, of the total edge flow, ω, can be obtained following the steps of the continuous Helmholtz decomposition.

In the Helmholtz decomposition, the gradient component ($-\boldsymbol{\nabla}\Phi$ in equation (1)) is computed by applying the gradient to an unknown scalar field called the scalar potential, Φ. Consequently, the discrete version of $-\Phi$ is a set of node weights (potentials) η which is a priori unknown. Ultimately, performing the gradient on η, the gradient component of the graph is obtained.

$$\omega_g = \delta_0(\eta) \tag{8}$$

Note here that the gradient operator δ_0 maps the subspace of nodes to the subspace of edges ($\eta \in \langle\Sigma_0\rangle \xrightarrow{\delta_0} \omega_g \in \langle\Sigma_1\rangle$). Conversely, the divergence operator performs the inverse mapping, relating edges to nodes. Consequently, the divergence operator is usually expressed as $-\delta_0^*$ in the literature (a more general description of these operators can be found in the Appendix and [25]) and this nomenclature will be adopted from now on.

It is important to highlight that, unlike in the continuum case, in this formalism, gradient flows start from nodes with low potentials and point towards nodes with higher potentials.

To compute the gradient component, the node potentials that generate it need to be obtained. For that matter, a system of equations can be derived by recalling that the divergence of the original field ω must be equal to the divergence of the gradient component alone ω_g, since the solenoidal and harmonic components are defined as *divergence-free*. Consequently, according to (6) and (8), a system of equations for the set of node potentials η can be obtained:

$$\delta_0^*(\omega) = \delta_0^*(\omega_g) + \cancel{\delta_0^*(\omega_s + \omega_h)} = \delta_0^*(\omega_g) = \delta_0^*(\delta_0(\eta)) \tag{9}$$

Which is a system of equations where $\delta_0^*\delta_0$ is called the graph's *Laplacian* because it is the discrete version of the Laplacian in the continuum (∇^2). Once the node potentials η are found solving (9), computing their gradient yields the gradient component, $\omega_g = \delta_0(\eta)$. In practice, the gradient operator is given by an $m \times n$ matrix ($\mathbf{B}$) where m is the number of edges and n is the number of nodes in the graph:

$$\delta_0 \equiv \boldsymbol{B}_{(\overline{i_0 i_1}),j_0} = \begin{cases} 1 & \text{if } i_1 = j_0 \\ -1 & \text{if } i_0 = j_0 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where $(\overline{i_0 i_1})$ is a given edge of the graph and j_0 is an arbitrary node. Applying this matrix to a column of node weights, one retrieves its gradient given by equation (7).

To compute the solenoidal component, ω_s, the curl operator must be defined in the discrete framework. The circulation over a cycle in a graph can be computed with

the weighted sum of the edges that belong to that cycle. Considering an oriented edge, if the flow follows the arrow's direction, its contribution will be positive, while, if it flows against it, it will contribute negatively. For example, in Figure 4, the value of the circulation around the triangle (012) is $3 + 2 - 3 = 2$. Recalling the definition of curl in the continuum (the circulation of a vector field around an infinitesimal loop) one can extend this concept to the discrete domain by considering the smallest cycle possible in a graph, the triangle. Consequently, the solenoidal component will account for the circulations around the triangles in the graph. This is the canonical way of defining the solenoidal component of the Hodge decomposition [19] [25], which corresponds to the one used in this work. Other approaches consider higher dimensional loops in the solenoidal component, which can be divided into cycles of different sizes [26]. In the canonical description, the curl is given by the δ_1 operator, which maps $\langle\Sigma_1\rangle \rightarrow \langle\Sigma_2\rangle$, that is edges to triangles. When applied to the whole graph, the curl operator computes the circulation of the flow around each triangle of the network and outputs a set of weights (the circulations) associated with each triangle. Ultimately, given a weighted edge $\omega_{\overline{i_1 i_2}}(\overline{i_1 i_2})$ that belongs to one or more triangles, the contribution to the curl of the triangle by the considered edge is computed following (11).

$$\delta_1 \left[\omega_{\overline{i_1 i_2}}(\overline{i_1 i_2})\right] = \sum_{(\overline{i_1 j i_2})\in\Sigma_2} \omega_{\overline{i_1 i_2}}\, \epsilon_{j\overline{i_1 i_2}}\, (\overline{i_1 j i_2}) \tag{11}$$

In practice, the curl operator in (11) can be rewritten with an $f \times m$ matrix ($\mathbf{C}$), where f is the number of triangles in the graph and m is the number of edges.

$$\delta_1 \equiv \boldsymbol{C}_{(\overline{i_0 i_1 i_2}),(\overline{j_0 j_1})} = \begin{cases} 1 & \text{if } i_1 = j_0, i_2 = j_1 \text{or} \\ & i_0 = j_0, i_1 = j_1 \\ -1 & \text{if } i_0 = j_0, i_2 = j_1 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Which corresponds to the transpose of the oriented edge-face incidence matrix or the boundary-2 matrix.

Now that the curl operator is defined, the solenoidal component, ω_s can be obtained by recalling that the curl of the original flow, ω, is strictly the curl of the solenoidal component, as the gradient and harmonic components are *curl-free* by definition.

$$\delta_1(\omega) = \cancelto{0}{\delta_1(\omega_g + \omega_h)} + \delta_1(\omega_s) = \delta_1(\omega_s) \tag{13}$$

Proceeding as for the gradient component, an a priori unknown set of triangle weights is considered, which can be understood as a set of triangle potentials.

$$\gamma = \sum_{(\overline{i_1 i_2 i_3})\in\langle\Sigma_2\rangle} c_{\overline{i_1 i_2 i_3}}(\overline{i_1 i_2 i_3}) \tag{14}$$

In equation (14) $c_{\overline{i_1 i_2 i_3}}$ is the weight associated with the triangle $(\overline{i_1 i_2 i_3})$. The operator δ_1^* performs the inverse mapping than δ_1, that is, triangles to edges. As explained in the appendix, S_1 (the solenoidal subspace) is the image of $\langle\Sigma_2\rangle$ (a subspace of weighted triangles) under δ_1^*. Consequently, if γ contains all the triangles in the graph with their respective potentials, $\omega_s = \delta_1^*(\gamma) \in S_1$ will be the solenoidal component of the Hodge decomposition. Using this definition and equation (13), a system of linear equations is found for the solenoidal component:

$$curl(\omega) \equiv \delta_1(\omega) = \delta_1(\omega_s) = \delta_1(\delta_1^*(\gamma)) \tag{15}$$

where the coefficients of each triangle, $c_{\overline{i_1 i_2 i_3}}$, in γ are unknown and the curl of the graph can be easily computed with (12). When the unknown triangle potentials are found solving (15), the solenoidal component can be computed by $\delta_1^*(\gamma)$, where δ_1^* is just $\mathbf{C}^T$. To simplify and summarize the process, the gradient and solenoidal potentials can be obtained by:

$$\eta = (\mathbf{B}^T\mathbf{B})^{-1}\boldsymbol{B}^T\,\omega \tag{16}$$

$$\gamma = \left(\boldsymbol{C}\boldsymbol{C}^T\right)^{-1}\boldsymbol{C}\,\omega \tag{17}$$

where ω is the column vector of edge weights of the original graph.

Lastly, the harmonic component, ω_h will contain all the larger circulations that cannot be accounted by the curl, that is all loops larger than triangles (squares, pentagons and so forth and bigger simplices). The harmonic component can be obtained by simply subtracting the two previously calculated components from the given graph edge weights. Additionally, since the harmonic component is both curl and divergence-free, it must satisfy (18).

$$(\delta_0\delta_0^* + \delta_1^*\delta_1)\,\omega_h = (\mathbf{B}\mathbf{B}^T + \mathbf{C}^T\mathbf{C})\omega_h = 0 \tag{18}$$

Solving for ω_h is the alternative way to find the harmonic component.

C. A worked example

To illustrate this process, consider the example in Figure 4.

The gradient operator in this case is:

$$\boldsymbol{B} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{19}$$

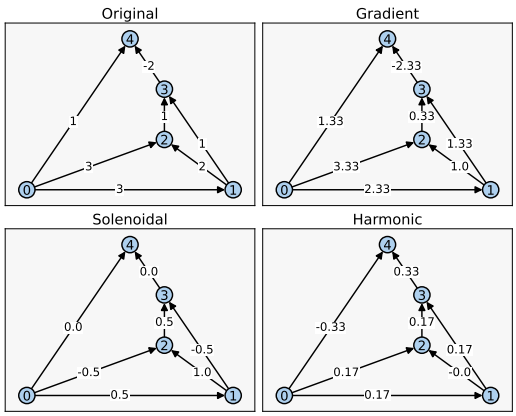


FIG. 4: Hodge decomposition of the original graph.

Using the expression in (9), the system of equations for the potentials of the gradient component with the laplacian matrix reads:

$$\begin{bmatrix} 3 & -1 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 & 0 \\ -1 & 0 & 3 & 0 & 1 \\ 0 & -1 & -1 & 3 & -1 \\ -1 & 0 & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_2 \\ a_4 \end{bmatrix} = \begin{bmatrix} 7 \\ 0 \\ -4 \\ -4 \\ 1 \end{bmatrix} \tag{20}$$

The curl operator in this case is:

$$\boldsymbol{C} = \begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 1 & 0 \end{bmatrix} \tag{21}$$

Next, using expression (15), the system of equations for the potentials of the triangles $\{(012), (123)\}$ for the solenoidal component reads:

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \tag{22}$$

Note that both (20) and (22) are overdetermined systems of equations, which implies that infinite different sets of potentials give rise to the same gradient and solenoidal components of a given graph.

IV. RANDOM WALKS ON SPATIALLY EMBEDDED NETWORKS

To simulate pedestrian mobility on a sidewalk network and analyse the resulting edge flows through the HHD, the agents walking through the graph are chosen to be random walkers. The reason for this choice, as mentioned previously, is that, to explore the structural effects of the network on the agent flow, it is convenient to use an unbiased dynamical model such that there are no preferred destinations. The use of biased mobility models such as gravity models would produce trivial gradient flows, as the field would be fully described by a set of potentials as seen in the related literature [14]. Additionally, real pedestrian mobility is usually bounded by distance or time, making the duration and length of trips an important feature which only agent-based models can capture. Moreover, to achieve sustainable mobility, cities are actively trying to adhere to the *15-minute city* format, [27, 28], where most daily necessities and services, such as work, shopping, education, healthcare, and leisure can be easily reached by a 15-minute walk from any point in the city. Consequently, stationary behaviours may not be the most important since distance and time travelled are relevant in human mobility. Using agent-based models such as random walks, the duration of the simulations can be restricted to account for these characteristics.

When modelling pedestrian behaviour, the simplest and less realistic approach is the one of a Markovian discrete-time random walk (DTRW). In this framework, the agents randomly choose a neighbouring edge at every intersection, jumping from node to node until a maximum amount of transitions is reached. This cutoff may be interpreted as the commuting time of each individual. The main issue with this model is that the geometrical contributions (i.e. edge lengths) to the dynamics are lost, and only the topological effects are captured. To introduce the metric distances of each edge and the geometric structure of the city, the discrete model can be upgraded such that the agents are able to move during a time interval Δt at the typical walking velocity $v = 1.42\, m/s$ (constant velocity random walk). This walking time may be interpreted as a budget which is spent according to the chosen link's length.

Consequently, the time spent to cross a given edge in the DTRW model will be always one iteration while, in the constant velocity model, it will depend on the length of the edge. In order to test the validity of the simulations, a deterministic description of each dynamics is needed. Despite this, the constant speed random walk cannot be modelled analytically (as will be explained later on), which motivates the use of other continuous-time random walk (CTRW) dynamics to get an approximate deterministic result.

In practice, n random walkers have been initialized at each node of a sidewalk network and are able to walk during a given number of steps (in case of discrete jumps) or a finite time Δt. Four different walker dynamics have been used:

1. A discrete-time random walk (DTRW) or Markov Chain.
2. A node-centric continuous-time random walk (CTRW) (model 1 of [11]).
3. A node-centric CTRW on the adjoint network (edge-centric CTRW).
4. A random walk at constant velocity.

The first three dynamics have been simulated and modelled deterministically, while the fourth one has been simulated and used as a benchmark for the two CTRW models presented. The three analytically describable models will be disclosed in depth in the following sections. In order to have an expectation of the net random walk flow along the graph's edges, a method built on [29, 30] has been developed in order to find the edge flows both in discrete and continuous-time Markov chains. Finally, to analyse the validity of the aforementioned dynamical models, the simulation results of each case have been compared to their corresponding deterministic descriptions. Namely, the evolution of the occupation probabilities $p_i(t)$ and the correlation between analytical and simulated node potentials and gradient components has been studied using an RG graph with $r = 0.2$.

A. Deterministic Description of the DTRW Flow

Typically, the transition probabilities of going from node i to node j, $W_{ij} = \frac{A_{ij}}{k_i}$, with k_i being the degree of node i and A_{ij} the adjacency matrix, form the transition matrix $\mathbf{W}$. Unlike other works [29], the random walkers considered have no predefined destination. This condition implies that there are no absorbing nodes, thus, the rows of the transition probability matrix sum up to one. Using this matrix one can easily find the probabilities of being at a given node n at step t:

$$\mathbf{p}(t) = \mathbf{p}(0)W^t \tag{23}$$

where $\mathbf{p}(t)$ is a vector of probabilities containing the probability of being at each node at step t.

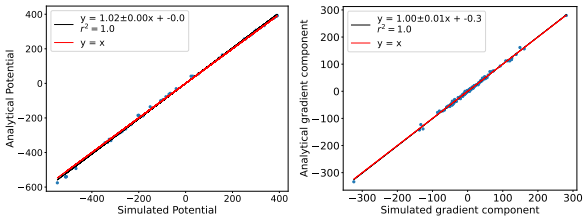
Once the occupation probabilities are obtained in (23), a novel approach is presented where one can compute the expected net flow through edge (i, j) at step t. The expected forward flow from node i to j between step t and step $t + 1$ can be computed using the probability of arriving from any initial node, s, at $t = 0$ to vertex i at step t and then moving to j in the following step. The expected backward flow can be calculated in the same manner but from j to i. Then, the net flow is obtained by the difference between these two expected flows. In mathematical terms:

$$\begin{aligned}\omega'_{ij}(t) &= p_i(t)W_{ij} - p_j(t)W_{ji} \\ &= p_i(0)\sum_s \left[W_{si}\right]^t W_{ij} - p_j(0)\sum_s \left[W_{sj}\right]^t W_{ji}\end{aligned} \tag{24}$$

Now, to find the accumulated edge flow from $t_0 = 0$ to t_1 (total number of forward crossings minus backward crossings from t_0 to t_1), equation (24) must be summed over all time steps:

$$\omega_{ij}(t) = \sum_{t=0}^{t_1} \omega'_{ij}(t) \tag{25}$$

It is important to note that, since the analytical edge flows are the result of a simple difference between the probability of crossing the edge forwards and backwards, they will only be capable to reproduce the gradient components of the simulated flows, as it is an effective difference between potentials. Consequently, the expected deterministic flows have been compared to the gradient component of the simulations (Figure 5).



(a) Potentials (b) Gradient component

FIG. 5: Simulated and predicted potentials and gradient component for an RG graph of 50 nodes and $r = 0.2$ according to the DTRW model. A linear regression is shown in black while the expected correlation is depicted in red.

As observed in Figure 5 the obtained potentials and gradient components agree with the simulations. Additionally, Figure 9 **(a)** shows that the discrete random walk model follows the expected node-probability curves.

Building on the presented description, a novel analytical formulation for the CTRW models has been developed in order to account for the geometrical properties of the network.

B. Analytical description for CTRW

With the aim to model walking agents on a pedestrian network while still considering the length of the streets, a continuous-time Markov chain with exponentially distributed waiting times may be the preferred option. These waiting times can be interpreted as the time a walker spends crossing a given link (street). In a CTRW with discrete space states, the probability of being at a certain node i at time t ($p_i(t)$) is given by the continuous-time master equation:

$$\frac{dp_i(t)}{dt} = \sum_{k\neq i} R_{ki}p_k - R_{ik}p_i \tag{26}$$

where R_{ik} is the probability rate at which a random walker leaves node i and transitions to node k, which is formally known as the transition rate from i to j. These rates can be organized in a transition rate matrix, $\mathbf{R}$, which satisfies the following conditions:

1. $R_{ij} \geq 0$: with $i \neq j$
2. $\sum_j R_{ij} = 0$: $\forall\, i$

3. $R_{ii} = -\sum_{i \neq j} R_{ij}$

which imply that $0 \leq -R_{ii} < \infty$. This description is only valid if the waiting times of the CTRW are exponentially distributed at each node with a rate λ_i, according to the memoryless property of continuous-time Markov chains. These waiting times will be interpreted in the following sections as the average time it takes to traverse the neighbouring links of node i. The transition rates adopt the following form:

$$R_{ij} = \begin{cases} \lambda_i p_{ij} & \text{if } i \neq j \\ -\lambda_i & \text{if } i = j \end{cases} \tag{27}$$

where $p_{ij} = \frac{A_{ij}}{k_i}$ is the transition probability from i to j, which only depends on the degree of node i in the case of unweighted random walks, and A_{ij} is the adjacency matrix of the considered graph.

Solving numerically the system of N coupled differential equations (26), with N being the number of nodes, $\mathbf{p}(t)$, the vector of node-probabilities, is obtained. From this result, the edge flow can be similarly obtained as in DTRW through the probability that a walker reaches node i at time t and then jumps to j in a time dt, which is given by $p_i(t)R_{ij}dt$. Consequently, the net flow of the edge (i, j) between t and $t + dt$ is given by:

$$d\omega_{ij} = \left[p_i(t)R_{ij} - p_j(t)R_{ji}\right]dt. \tag{28}$$

Finally, integrating from $t_0 = 0$ to $t = t_1$, the expected CTRW edge flow over a given time interval for the edge (i, j) is obtained:

$$\omega_{ij} = \int_0^{t_1} \left[p_i(t)R_{ij} - p_j(t)R_{ji}\right] dt \tag{29}$$

In this work, this formalism has been used to retrieve the discrete vector fields from the node-centric and edge-centric CTRW models which will be described in the following section.

C. CTRW models

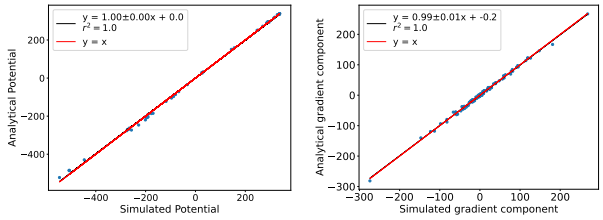
In this section, the two CTRW models will be briefly analysed and validated.

1. *Active Node-centric Random Walks*

The active node-centric random walk, [11], is the simpler approach, in which the walker jumps from each node to the next one with a rate λ_i which depends on the original node. With the aim of taking this model closer to a pedestrian walking at constant speed v, the average waiting time, $1/\lambda_i$, at each node is reinterpreted as the walking time between the original node and the following one. Consequently, this walking time is chosen to be the mean of the walking times of the adjacent edges of node i if the pedestrian moved at v:

$$\lambda_i = \sum_j A_{ij} \frac{vk_i}{d_{ij}} \tag{30}$$

where d_{ij} is the distance of the link (i, j). Following the developed description to compute the expected edge flows (29), Figure 6 shows a perfect correlation between the simulated and analytical potentials and gradient components. Additionally, Figure 9 shows that the evolution of the node probabilities in the simulation follows the expectation values predicted analytically with different stationary values than in the DTRW. These results validate the formulation developed in the previous section.



(a) Potential (b) Gradient component

FIG. 6: Simulated and predicted potentials and gradient components for an RG graph of 50 nodes and $r = 0.2$ according to the active node-centric model. A linear regression is shown in black while the expected correlation is depicted in red.

The node-centric dynamical model is simple and widely used in CTRW literature [11]. Despite this, the reinterpretation of the average waiting time at a node as the walking times of its adjacent edges may be inaccurate if these have heterogeneous lengths. In this case, the mean would not be representative of the actual link distances. To bypass this issue while using a similar formulation, a new approach has been studied, where the random walk is performed directly on the edges.

2. *Edge-centric Continuous-Time Random Walks*

To determine the waiting times for each edge using a similar framework as in the usual node-centric case, the CTRW can be performed on the edges such that the waiting times are specific to each link depending on its length. Consequently, the transition rates are edge dependent and become λ_{ij}. This description reduces to a node-centric random walk on the adjoint graph of the original one (an edge-centric CTRW), where the edges become nodes and links are placed between adjacent edges.

Using the framework provided by equation (26) the continuous time master equation for a random walk on

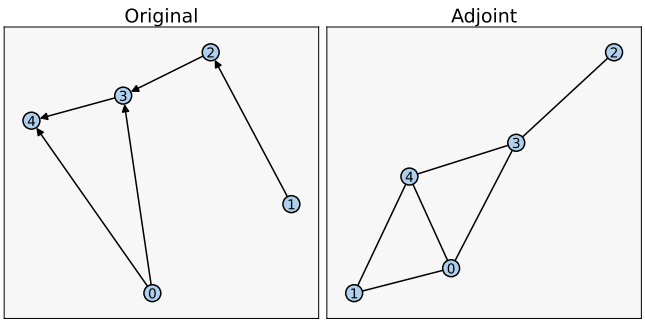


FIG. 7: Example of an adjoint graph.

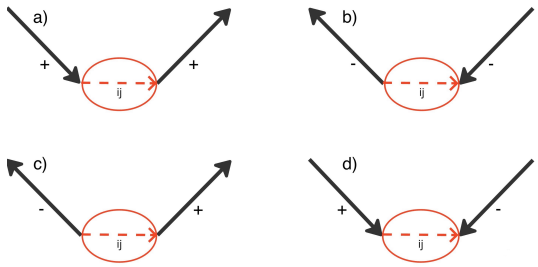


FIG. 8: Possible incoming and outgoing fluxes in the adjoint graph. The dotted and encircled arrow is an edge of the original graph that is represented as a node of the adjoint graph. The black arrows at each side represent the sum of the left and right fluxes respectively.

the edges gives the probabilities of being at edge (i, j) at time t:

$$\frac{dp_{ij}(t)}{dt} = \sum_k \left[p_{jk} \frac{\lambda_{jk}}{k_{jk}} + p_{ki} \frac{\lambda_{ki}}{k_{ki}} \right] - \lambda_{ij} p_{ij} \tag{31}$$

where $k_{ki} = k_k + k_i - 2$ is number of adjacent links of edge (k, i) which corresponds to its degree in the adjoint graph.

Equivalently, equations (28) and (29) can be used to obtain the edge flow in the adjoint graph. Despite this, the edge flows on the original links are needed. Consequently, the vector field in the adjoint graph's edges needs to be converted to flows on the original graph's edges. Unfortunately, this procedure can be ambiguous in some cases. From the flow at the adjoint edges, one can derive the net flow through the adjoint nodes (original edges) by considering the incoming and outgoing flows. Generally, if an original edge is considered, the incoming flow that enters through the base of the arrow is positive, while if it enters from the head (in the opposite sense of the edge) it is considered negative. For the outgoing flow, if it exits the edge through the arrow's head the flow is considered positive while if it exits through the base of the arrow it means that the edge has been traversed in the opposite direction, thus the contribution to the flow is negative. There are four possible combinations of incoming and outgoing flows represented in Figure 8.

As walkers in the edge-centric CTRW start and end at the edges, divergence becomes a problem when counting the flow through them. In a divergence-free case, the incoming, $\sum_k e_{ki \to ij}$, and outgoing, $\sum_l e_{ij \to jl}$, flows are equal, and the net flow through a given edge (i, j) is just (32) as all the incoming flow passes through the edge:

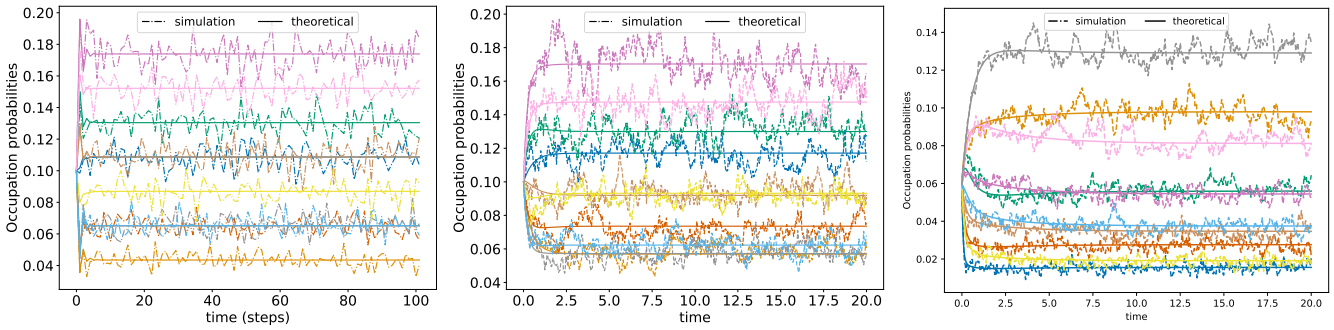
$$\omega_{ij} = \frac{1}{2} \left(\sum_k e_{ki \to ij} + \sum_l e_{ij \to jl} \right) \tag{32}$$

where $e_{ki \to ij}$ is the flow through the adjoint edge that connects the real edges (k, i) and (i, j). When divergence is not null, some walkers may start or end at each edge. This is problematic because in usual random walks, the edge simply links two nodes and it is not a state where walkers can be born or die. Consequently, when a walker performs its last jump to edge (i, j) and dies there, a decision needs to be made on whether it ends on node j, crossing (i, j) and contributing to its flow, or on i without crossing the last link. The same problem is faced with the first jump, as the starting node in the initial edge needs to be decided.

In the simulation, this problem is approached by adding a certain crossing probability for the walkers that start or finish at a given edge. When a walker performs the initial jump, it crosses the starting edge with a probability of 0.5. On the other hand, when a walker performs the final jump, one can rudimentarily assume that walkers will cross the final edge half the times they end there, thus with a probability of 0.5. Despite that, a more refined approach is to consider that the crossing probability in the final step is given by the cumulative exponential distribution with λ_{ij} depending on the arrival edge, $P(\tau \leq t_1 - t)$. This probability depends on the remaining time-budget of that walker $(t_1 - t)$, and it is the probability that a further jump takes place before the remaining time has passed. Consequently, these undetermined countings are flattened and distributed among the two nodes of the considered edge.

While in the simulations both methods can be used, in the analytical description, only the first procedure can be computed, as there is no notion of the individual walkers and their remaining time. This ultimately implies that the net analytical edge flow is the mean between the incoming and outgoing flow, which is conveniently given by equation (32).

In the simulation, the forwards and backward crossings of edge (i, j) are computed considering the nth jump of the random walk from the edge (i, j) and the previous jump $n - 1$. Essentially, if the walker enters the edge through the base of the arrow in the step $n - 1$ and leaves at step n through its head, a positive counting will be added. If the edge is traversed in the opposite direction the contribution to the flow will be negative. If a walker comes and leaves through the same extreme (base or head) of the arrow it will not contribute to its flow (see Table I).



(a) DTRW. (b) Node centric CTRW. (c) Edge-centric CTRW

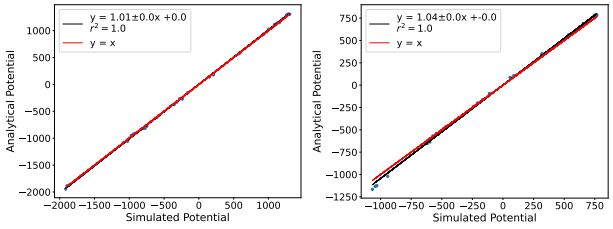
FIG. 9: Simulated and predicted node ((a) and (b)) and edge (c) occupation probabilities for an RG graph of 10 nodes and $r = 0.47$ with 20 walkers starting at each node. The nodes of the graph exhibit 7 different degrees ($k_i = 2, 3, 4, 5, 6, 7, 8$).

For the edge-centric random walk, the simulated occupation probabilities at the edges fluctuate around the analytical evolution (Figure 9, **(c)**).

$\Delta\omega_{ij}$	Jump	
	n-1	n
1	$ki(ik) \rightarrow ij$	$ij \rightarrow jl\ (lj)$
-1	$kj(jk) \rightarrow ij$	$ij \rightarrow il\ (li)$
0	$ki(ik) \rightarrow ij$	$ij \rightarrow ki\ (ik)$
0	$kj(jk) \rightarrow ij$	$ij \rightarrow kj\ (jk)$

TABLE I: Flow contributions to edge (i, j) according to the previous jump. The arrows represent transitions between adjacent edges and edge (i, j).

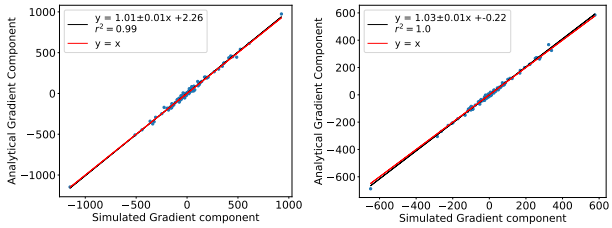
Considering the two presented approaches to account for the final jumps in the simulation, the correlation of the obtained potentials and gradient components is shown in figures 10 and 11 respectively.



(a) Acceptance probability of 0.5. (b) Exponential acceptance probability of $P(\tau \leq t_1 - t)$.

FIG. 10: Analytical vs. simulated potentials in the edge-centric random walk model using the two different crossing probabilities in the simulation. A linear regression is shown, as well as the expected correction curve in red. The analysis has been performed on the RG graph with 50 nodes and $r = 0.2$.

Although both descriptions show similar results, applying an exponentially distributed probability yields slightly better results with different RG graphs.



(a) Acceptance probability of 0.5. (b) Exponential acceptance probability of $P(\tau \leq t_1 - t)$.

FIG. 11: Analytical vs. Simulated gradient components in the edge-centric random walk model using the two different crossing probabilities in the simulation. A linear regression is shown, as well as the expected correction curve in red. The analysis has been performed on the RG graph with 50 nodes and $r = 0.2$.

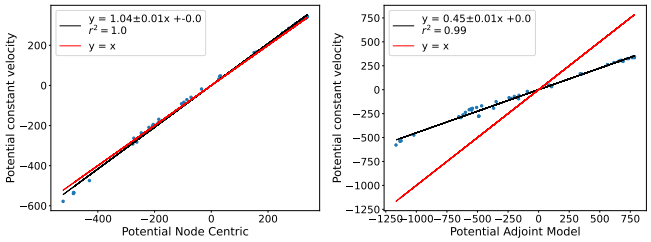
V. RESULTS

The previous random walk dynamics have been applied to the toy graph models mentioned above, which have clear structural differences in order to obtain a flavour of the impact of the geometry on the obtained flow decomposition.

A. Validity of the Deterministic Models Against a Constant Velocity Random Walk

Ultimately, the goal of the aforementioned random walk models is to approximate, both deterministically and in the simulations, pedestrians randomly moving through the sidewalks of a city at a constant speed, v, from $t_0 = 0$ to $t = t_1$. These dynamics cannot be deterministically modelled directly because the jump intervals do not follow any probability distribution. The time spent crossing a link (i, j) is always $\tau_{ij} = \frac{\Delta x_{ij}}{v}$, and the total elapsed time depends on the path taken by the random walker. Overall, this type of dynamics can only be

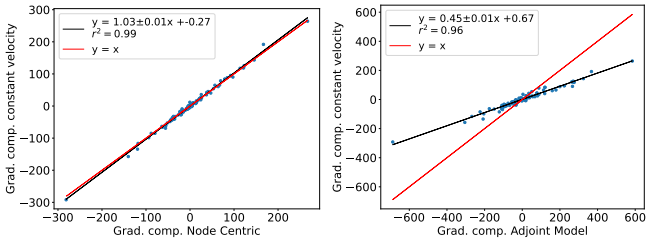
simulated, and the previously discussed models, which do have an analytical characterization, attempt to be an approximation to the desired pedestrian dynamics. In order to measure their validity against the constant speed random walk, the potentials and gradient components obtained with the deterministic approach have been compared to the simulation of the constant speed pedestrian model.



(a) Node-centric vs. constant speed random walk.

(b) Edge-centric random walk vs. constant speed random walk.

FIG. 12: Comparison of the potentials of the node-centric and edge-centric random walks (exponential acceptance probability) using their analytical description against the constant velocity model. The graph considered is an RG of 50 nodes and $r = 0.2$. A linear regression is shown as well as the expected correlation in red.



(a) Node-centric vs. constant speed random walk.

(b) Edge-centric vs. constant speed random walk.

FIG. 13: Comparison of the gradient components of the node-centric and edge-centric random walks (exponential acceptance probability) using the analytical models against the constant velocity model. The graph considered is an RG of 50 nodes and $r = 0.2$. A linear regression is shown as well as the expected correlation in red.

As observed both in Figures 12 and 13, the node-centric model with averaged transition rates shows correlations with the constant speed pedestrian model, with $r^2 = 0.99$, and is a good approximation when the average edge length is close to the real lengths of the edges incident to the node. The Edge-centric random walk shows a good correlation with $r^2 = 0.96$, but the fitted slope is not 1. The reason for these results could reside in the way the edge crossings are counted. Ultimately, this model needs a transformation from a random walk on the edges to a random walk on the nodes to retrieve an edge flow. This process is not unique and some assumptions are needed to treat cases such as if the last or the first jumps are considered to cross the first and last edges respectively, leaving room for uncertainty and error.

B. Geometrical Role of the Hodge Potential in the Discrete- and Continuous-Time Random Walks

To understand how the given graph geometry affects the flows and potentials retrieved from the Hodge decomposition, the vector field components induced by the different random walks on the PBC-modified lattice have been obtained (Figure 2). This graph presents two clearly distinguishable regions with different node densities and degrees. By analysing the node potentials, the aim is to observe how the differences in node densities influence the random walk behaviour.

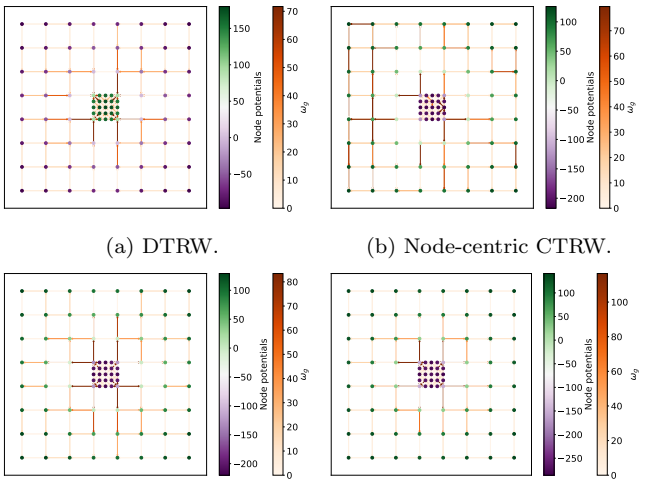
It is well known that a discrete random walk has a stationary occupation probability which is proportional to the degree.

$$p_i^{st} = \frac{k_i}{\sum_{j=1}^{N} k_j} = \frac{k_i}{2E} \tag{33}$$

being E the number of edges. This property implies that the nodes with a higher degree, the dense region of the modified PBC lattice, will have larger occupation probabilities, and, thus, behave like attractors. Consequently, their potentials will be larger (more positive) than the nodes with lower degrees, and gradient flows will point towards regions with larger edge density as seen in Figure 14. In this setup, the results obtained are a direct consequence of the degree of a node in the discrete random walk dynamics.

Accordingly, in CTRWs intuition may lead to expect that communities of close and interconnected nodes will also behave like attractors (thus, again, larger potentials), as walkers that enter these regions would struggle to escape from them as if they were in a labyrinth. Despite this, the potentials of the dense region of the lattice retrieved from all the CTRW models show to be smaller than the potentials of the outer (and less connected) nodes (Figure 14 **(b), (c)** and **(d)**). This implies that in general denser regions will usually show a repulsive character, contrary to the results obtained with DTRWs. To understand these counter-intuitive results one has to consider that the computation of the "attractiveness" or potential of a node involves its divergence. Generally, nodes with negative divergence (sinks) will exhibit high potentials and vice-versa. While it is true that both types of walkers, DTRW or CTRW, will perform more jumps in the denser region of the lattice, only the initial and last steps contribute to the divergence since the vector field is constructed upon net flow. Consequently, if a walker performs many jumps in the denser region but ends outside of it, this region of the graph will not be considered attractive in terms of potential but neutral.

The cause of the attractiveness/repulsiveness of high-density areas can be understood by interpreting the available walking time, t_1, as a budget. In DTRWs, all edges have the same cost regardless of their length, while, in CTRWs, shorter edges will have a lower traversing cost. This ultimately implies that, with the same available budget, CTRWs will be able to perform more jumps inside the dense region in the lattice. Consequently, the probability of escaping from that region will be higher in CTRWs, as the walker will have more chances to leave the cluster. When this happens, the "budget" left is quickly spent on the longer outer edges and the walker has a lower chance to reach back in the denser region.



(a) DTRW. (b) Node-centric CTRW.

(c) Constant v CTRW. (d) Edge-centric CTRW.

FIG. 14: Gradient components (ω_g) and node potentials obtained through all the studied dynamics on the modified lattice with PBC. Edges have been reoriented to follow the direction of the flow.

Ultimately, Figure 14 seems to indicate that the node's potential, hence the attractiveness of a node, will be a result of two competing network properties, the node's degree, which generates high potentials when it is large, and the lengths of its neighbouring edges, which, in turn, define its transition rate which determine how much of its budget is spent in the jumps to the adjacent nodes. It is seen that short average neighbouring link distances tend to decrease the node's potential, which is the case in dense and connected regions.

While the two outcomes of the competition between node degree and edge length are evident in Figure 14, to further analyze the effect of incorporating link distances into the random walk dynamics, the evolution of the gradient component and its potentials has been studied while transitioning from the DTRW to CTRW model. This has been done by continuously varying the length of the edges, since, when all the edge lengths of the network are equal, the DTRW and CTRW models are equivalent.

Specifically, the edges in the denser region are enlarged from their original length to the distance between nodes in the regular lattice. To see if the denser region is repulsive or attractive, the difference between the average of the potentials of the inner nodes and the average of the potentials of the outer nodes (ΔV_{io}) is shown against $\theta = \frac{d_{mod} - d_{real}}{d_{lattice} - d_{real}}$, where $d_{lattice} = 1$ is the length of the regular lattice spacings, d_{real} is the original length of each inner edge and d_{mod} is the modified length of the inner edges which ranges from $d_{lattice}$ to d_{real}.

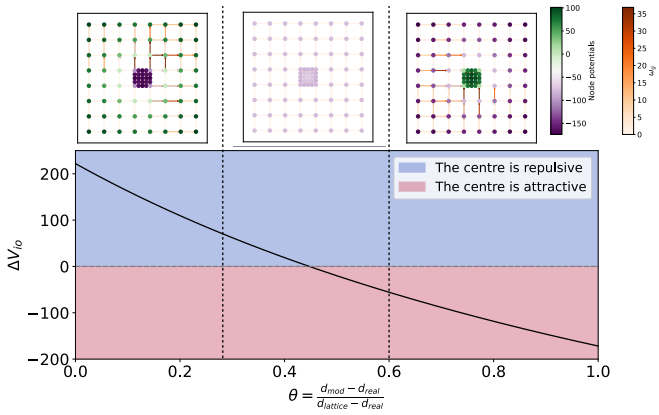


FIG. 15: Difference of average potential between the inner and outer nodes of the cluster. The analytical node-centric CTRW model has been computed for different link distances. On top, the node potentials and gradient components are shown for different values of θ, which is 0 for the real distances and 1 when the lengths of the inner edges are equal to the regular lattice link length.

Results in Figure 15 show the dominance shift between the node degree and the edge lengths in the average potential difference between the inner and outer regions of the cluster. Initially, for lengths close to the real ones, the inner region is repulsive as seen for CTRWs. In this situation, the edge lengths are crucial in how the random walk budget is spent, and their effect outweighs the one of the degrees. As inner link distances grow towards the lattice ones, a transition is observed, where the effects of the degree and edge distances are compensated and the potential difference between the two regions is zero, producing no net flow. Finally, for larger inner-link distances, where all the edges of the graph have a similar length, the potentials resemble the ones of DTRWs and the inner region is attractive.

C. Origin of the Cyclic Components

So far, the emphasis of the study has been put on the gradient component of the decomposition and its originating potentials. The reason for this is that the deterministic description and computation of the discrete and continuous random walk flows, Eqs. 24-25 and 28-29, for each edge flow are obtained as the difference between two node-dependent values that can be directly reinterpreted as node potentials. Despite the transition probabilities

(W_{ij}) and rates (R_{ij}) being defined for each edge, both only depend on the starting node i, as $W_{ij} = 1/k_i$ and $R_{ij} = \lambda_i/k_i$. Consequently, the deterministic approach cannot generate cyclic (solenoidal and harmonic) flows, and the resulting field can be fully described by the gradient component, as the results in Figure 6 showed.

Conversely, in the simulations of all the random walk models (DTRW, both CTRWs and the constant-velocity case) the cyclic components are never zero if any loops are present in the graph. Two hypotheses are presented to understand why the developed framework does not predict cyclic flows. The first option is that the harmonic and solenoidal components simply cannot be computed with this approach and another formalism needs to be used. The second explanation is that, if the deterministic description does not predict the cyclic components, then they are a result of stochasticity and their expected values are really 0 as obtained in the deterministic framework. This last interpretation would be compelling because it would allow the isolation of the noise contributions of the random walk flow into separate components.

Starting from the first hypothesis, another approach to computing the solenoidal component is first to calculate the expected circulations around each triangle of the graph and then, from those values apply the δ_1^* operator to retrieve the solenoidal component. To calculate the expected circulation of flow around a triangle $\left\langle \phi_{\overline{ijk}} \right\rangle$ in the DTRW during a total amount of steps Δt, equation (34) can be implemented.

$$\left\langle \phi_{\overline{ijk}} \right\rangle = \sum_{t=0}^{Dt} \Big\{ p_i(t)p_j(t)p_k(t) \Big[3T_{ij}T_{jk}T_{ki} - 3T_{ik}T_{kj}T_{ji} + T_{ij}T_{ji}(T_{ki} - T_{kj}) + T_{jk}T_{kj}(T_{ij} - T_{ik}) + T_{ik}T_{ki}(T_{jk} - T_{ji}) \Big] \Big\} \quad (34)$$

Equation (34) computes the expected value of the circulation around a triangle ($\overline{ijk}$) when having one walker at each of its nodes. Essentially it is the probability of simultaneously having a walker on each of the nodes of the triangle a time t and each making a forward or backward step accounting for all the possible contributions to the circulation. The expected value is calculated for all the simultaneous combinations of forwards/backward steps. Ultimately, as the transition probabilities only depend on the departure node at each step, all the elements $(T_{ij} - T_{ik})$ will be zero as $T_{ij} = T_{ik} \quad \forall j, k$. Consequently, in general, the predicted circulations will also be 0 if the same development is performed considering one, two, or any number of simultaneous walkers in the triangle.

If a similar framework is applied to the continuous case, the expected circulation around any triangle will also be zero because the transition rates are also node-dependent $R_{ij} = R_{ik} \quad \forall\, j, k \neq i$. The only model which might not have this issue is the edge-centric CTRW, as the edge walking time is edge dependent.

Since the deterministic approach is not able to reproduce the observed curl and harmonic components in the simulations, the stochastic cyclic components have been analysed in order to find properties that may reveal their origin.

To see if there are any stationary values of the cyclic component, the mean squared flows (MSFs) of the cyclic, gradient and total components have been computed for different values of the simulation times t_1.

$$\left\langle \omega_g^2 \right\rangle = \frac{1}{E} \sum_{i,j} (\omega_{ij}^g)^2 \quad ; \quad \left\langle \omega_{cycl}^2 \right\rangle = \frac{1}{E} \sum_{i,j} (\omega_{ij}^{cycl})^2 \quad (35)$$

where, in the last expression E is the number of edges of the network and ω_{ij} is the edge flow of (i, j) of each component.

Results in Figure 16 show the MSFs of each component. Firstly, the gradient MSF reaches a stationary value in all the graphs. This behaviour is expected by the deterministic prediction, where the evolution of the node probabilities in Figure 9 reaches a steady state. Looking at the cyclic component, the cyclic MSF seems to grow linearly with the duration of the simulation, with exponents 1.03 ± 0.03 for the ER, 0.98 ± 0.02 for the PBC modified lattice and 0.96 ± 0.03 for the RG cases. This behaviour shows that the components that generate circulatory flows do not reach a stationary value. Additionally, the amplitudes of the fluctuations along the fitted curves grow in time.

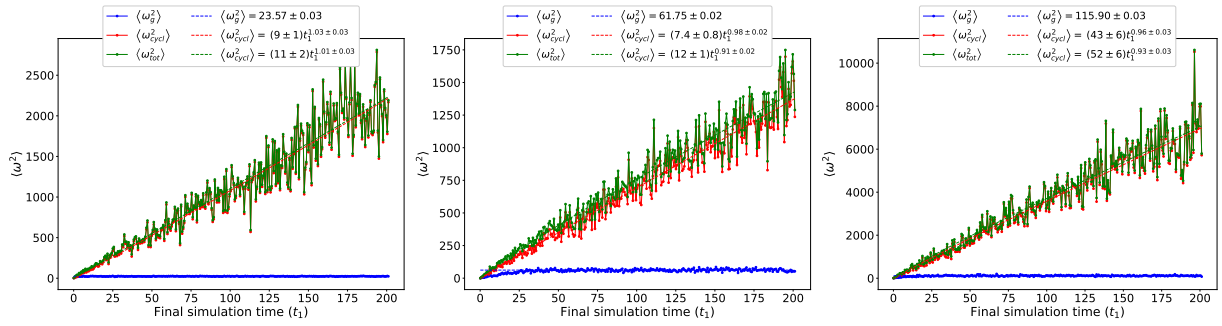
This linear behaviour of the MSF is reminiscing of the typical linear dependence of the Mean-Squared Displacement (MSD) in geometric random walks where $\left\langle (x - x_0)^2 \right\rangle = 2Dt$ with D being the diffusion coefficient.

The effect of this behaviour can also be seen in the evolution of the importance or strength ratio of each component. As the HHD divides the edge domain into three orthogonal subspaces, the $l^2 - norm$ is conserved. This property allows to quantify the importance or weight of each component [17, 18]. Namely, the norm of the original graph is the sum of the norm of each component.

$$1 = \frac{\|\omega_g\|^2}{\|\omega\|^2} + \frac{\|\omega_s\|^2}{\|\omega\|^2} + \frac{\|\omega_h\|^2}{\|\omega\|^2} = \eta_g + \eta_s + \eta_h \quad (36)$$

where η_g, η_s and η_h are the relative strengths of the gradient, solenoidal and harmonic components respectively.

The strength ratios in Figure 17 reveal a power law behaviour with a negative exponent very close to -1 in the presented cases. Additionally, an initial transient behaviour is observed. Understandably, as the MSF of the cyclic component increases and the gradient MSF remains constant, the gradient strength ratio decreases while the cyclical components become more dominant. Lastly, the initial values of the gradient and cyclic

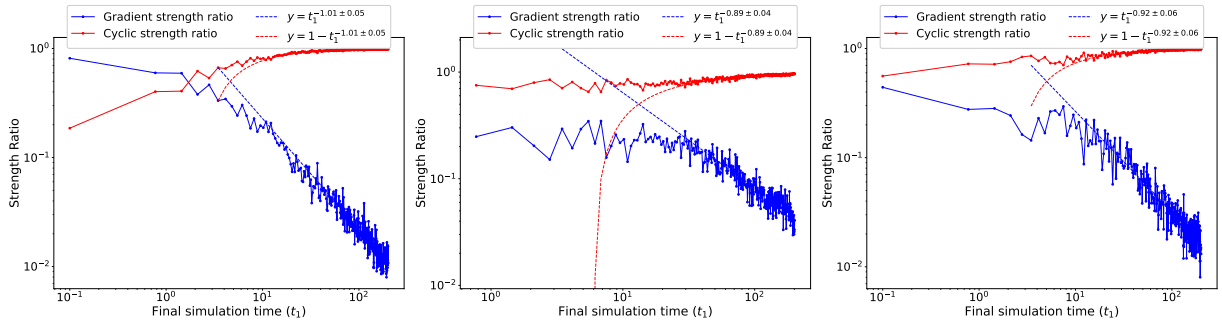


(a) Erdös-Rényi with 50 nodes and $p = 0.1$.

(b) PBC modified lattice.

(c) Random Geometric with 50 nodes and $r = 0.2$.

FIG. 16: Gradient, cyclic and total MSFs for different total simulation times. A power-law is fitted to the cyclic and total components, and the stationary value of the gradient MSF is also shown. Simulations use 20 walkers starting at each node.



(a) Erdös-Rényi.

(b) PBC modified lattice.

(c) Random geometric.

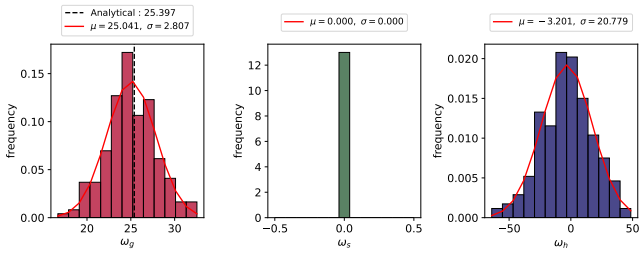
FIG. 17: Gradient, cyclic strength ratios for different total simulation times using the three model graphs mentioned above. A power law is fitted to the gradient ratio. The axes ar in log-log scale.

strength ratios are very different for the ER and the lattice cases. In the modified lattice, the cyclic strength ratio is initially very high (larger than the gradient ratio), while in the ER case, the gradient strength ratio is initially higher. Interestingly, since the modified lattice is completely regular except for the modified cluster, the degree and link distance distributions are very narrow. This leads to a more uniform distribution of the node probabilities which, in turn, generates a weaker gradient component. Additionally, there is a large number of loops, the circulation of which is captured by the cyclic component. Hence, due to the regularity of the graph and the larger presence of loops, the cyclic ratio is expected to be more important than in the ER case, where the edge length distribution is broader.

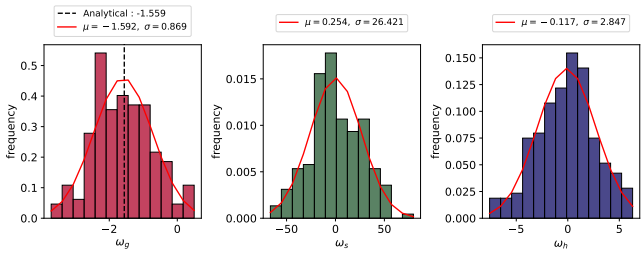
Finally, by focusing only on one of the edges of the graph and performing the node-centric CTRW simulation 200 times, the distribution of the flow and its components on that edge can be obtained.

In Figure 18, the histograms for the components of two different edges of the PBC-modified lattice are shown. Subfigure **(a)** corresponds to an edge of the regular, grid-structured region of the lattice. As expected, the cyclical component is fully captured in the harmonic component as the edge does not belong to any triangle. Subfigure **(b)** is representative of an edge in the dense region. As the considered edge is part of a triangle, the solenoidal component is present and contains the majority of the cyclic flow. On another note, the mean of the gradient component distribution on both subfigures **(a)** and **(b)** is close to the expected analytical value calculated using the formalism in section IV. Conversely, the dominant cyclic components have a standard deviation one order of magnitude higher than the one of the gradient component, and the mean of the distribution is close to zero. Again, the cyclic flows do not seem to point towards any expected value different than zero, and, instead, their fluctuations are much larger than the ones of the gradient component.

These results point towards the conclusion that the cyclic components are a consequence of the own stochasticity of the process, but a deeper and more complete analysis is needed to back this hypothesis and see if there are any structural correlations embedded in the cyclic flow.



(a) Edge of the regular region of the lattice.



(b) Edge of the modified region of the lattice.

FIG. 18: Histograms of each component of the flow along two edges in the PBC modified lattice. A Gaussian distribution has been fitted, and the values of the mean and the variance are shown in each subplot. For the gradient component, the deterministic flow is also shown.

VI. DISCUSSION AND CONCLUSIONS

This work has been focused on the impact of the topology and geometry of urban infrastructure and sidewalk networks on pedestrian mobility and its dynamics. Consequently, the Hodge decomposition has been used to partition random walk edge flows on spatially embedded graphs, described as discrete vector fields, into components that give information about divergent (gradient component) and cyclical (solenoidal and harmonic components) contributions. This description may be useful to visualize inherent structural city centres through the potentials and gradient flows, while the cyclic components may contain information about how more complex or more regular urban configurations affect the pedestrian flow. To isolate structural effects and remove artefacts or biases induced by the underlying dynamics, four different random walk models have been used and three of them have been modelled deterministically. To accommodate for the limited travel range of pedestrian mobility, the random walkers have been attributed a temporal budget interpreted as the maximum walking time. A new methodology inspired in [29] has been developed to obtain the deterministic time evolution of the expected net edge flow in both continuous-time (CTRW) and discrete-time (DTRW) random walks. This description has been validated through the correlation between the expected value of the edge flows and the simulated gradient components as well as with the expected and simulated node potentials.

After the used dynamics and deterministic approach were validated, the analysis of the resulting potentials for the DTRW and CTRW revealed the competing effect between the node degree and edge length in the attractiveness of densely connected regions of the graph. Ultimately, for CTRWs the clustered regions seem to exhibit a repelling behaviour while, for DTRWs, these regions appear to be attractive. These diametrically opposed outcomes have been explained through the differences in how the temporal budget is spent in DTRWs and CTRWs. This transition in the node potentials and the attractive nature of close and connected regions has been captured as a function of the edge length. This effect has been reinterpreted as a continuous transition between CTRWs and DTRWs.

Finally, since cyclical components (harmonic and solenoidal) are not captured in the deterministic framework developed in this work, their origin has been discussed. The analysis of the mean-squared flow (MSF) for gradient and cyclic flows reveals that the former tends to a stationary value as expected deterministically, while the cyclic MSF grows linearly with the simulation time t_1. Additionally, the fluctuations of the cyclic MSF also grow for increasing values of the time budget. These results are also confirmed in the analysis of the gradient and cyclic strength ratios, which reveal a transient phase followed by a power-law behaviour.

Lastly, the histograms for each component of the flow of two distinct edges for a fixed simulation time show that the mean of the distribution of gradient components coincides with the deterministic expected flow. Additionally, the cyclic components exhibit large standard deviations and are centred around the zero value.

All the results obtained from the cyclical components point towards the idea that the stochastic nature of the random process can be captured in the cyclic components of the simulated random walk flow. Despite this, a more in-depth analysis is needed to back this idea and see if the cyclic components present any correlations with the graph's structure and geometry.

This work has provided an understanding of the random walk dynamics on simple graphs through the Hodge decomposition, which will serve as the foundation for future works where this approach will be applied to analyse the effect of the geometry of real urban networks and other transport layers on the human flow.

Finally, other factors should be studied as diverse random walk initial distributions or dynamical behaviours. This would help to simulate pedestrian behaviour more accurately, as regions with more nodes (intersections) do not usually imply more density of population and real pedestrian dynamics are not random.

APPENDIX

Graphs as simplicial complexes

Formally, an undirected graph is a pair $G = (V, E)$ where V and E represent a vertex (also referred to as a node) and an edge set respectively. Usually, vertices are denoted as $v_1, v_2, ..., v_n$ and an edge between v_i and v_j as e_{ij}.

A graph can also be described by higher dimensional structures, needing a more general notation. A k-dimesional *simplex* is a fully connected (every vertex is connected to all the others) subgraph containing $k+1$ vertices. Following this definition, a 0-simplex is a vertex, a 1-simplex is an edge, a 2-simplex is a triangle and so forth.

In the previous definition of graph, only the lowest dimensional simplices were used. Alternatively, one can define a graph as a collection of simplices of different dimensions $G = (\Sigma_0, \Sigma_1, ..., \Sigma_k)$, where Σ_i contains fully connected vertex subsets of size $i+1$.

Applying this formalism, one can also define a set of weights for each simplex. Take as an example a neighbourhood where the nodes are intersections and the edges are a given set of bidirectional streets. One could assign the net flow of cars through a street as a weight in a given edge. Similarly, one can assign weights to any simplicial structure, from a vertex to large simplices.

In undirected graphs, one can define an orientation for the edges and bigger structures such that they are not considered more than once. Looking at the example in figure 19, the triangle (123) is the same as (231) or other index permutations. Taking the ascending order of its indices as a reference, a given simplex could be labelled with any odd or even permutation of the indices. The notation $\overline{v_0 v_1 ... v_m}$ describes $0 \leq v_1 < v_2 < ... < v_m \leq m$ the ascending order of any given simplex. To avoid the labelling confusion, a k-simplex can be oriented such that the positive orientation is associated with $\overline{v_0 v_1 ... v_m}$ and any even permutation, while the negative orientation is associated with the odd permutations. This can be described by the Levi-Civita tensor acting on a form. In this case, the form is a given simplex with its weight associated. Note that, formally, a form maps a vector to a real number, in this case, the form maps a simplex $(\overline{v_0 v_1 .. v_k})$ to its associated weight $(w_{\overline{01...k}})$. Ultimately, any permutation of a simplex can be written as:

$$a_{i_1 ... i_k}(i_1 ... i_k) \equiv \epsilon_{i_1 ... i_k} a_{i_1 ... i_k}(\overline{j_1 ... j_k}) \tag{37}$$

where $a_{i_1 ... i_k}$ is a real number associated to the simplex, $\overline{j_1 ... j_k}$ is the ascending permutation of $i_1 ... i_k$ and the Levi-Civita tensor is either 1 if the number of permutations with respect to the ascending ordering is even or -1 if it is odd.

Each set of simplices defines a vector space. In Figure 19 the vector space $< \Sigma_2 >$ has 2 dimensions (contains 2 triangles). Mappings can be defined from $\langle \Sigma_k \rangle$ to $\langle \Sigma_{k+1} \rangle$ and vice versa. Given a simplicial complex (a graph) $G = (\Sigma_0, \Sigma_1, ..., \Sigma_K)$, for $0 \leq k < K$ the operator $\delta_k : \langle \Sigma_k \rangle \rightarrow \langle \Sigma_{k+1} \rangle$ can be defined:

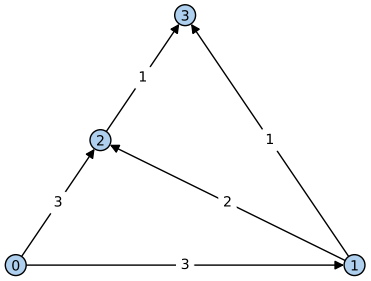


FIG. 19: Simple graph with four nodes and four edges.

$$\delta_k(\overline{i_1 ... i_{k+1}}) = \sum_{(\overline{i_1 ... j ... i_{k+1}}) \in \Sigma_{k+1}} \epsilon_{j \overline{i_1 ... i_{k+1}}} (\overline{i_1 ... j ... i_{k+1}}) \tag{38}$$

Where an extra vertex j, fully connected to the previous $k-simplex$, has been added to the correct place (according to the ascending ordering) in the $k-simplex$ $(\overline{i_1 ... i_{k+1}})$. To give an example, considering the edge (12), $\omega = 2(12)$, in Figure 19, the δ_1 operator maps this edge to the triangles in which it is contained.

$$\delta_1(\omega) = 2\left[(312) + (012)\right] = 2\left[(123) + (012)\right]$$

This operator maps $\langle \Sigma_k \rangle$ and $\langle \Sigma_{k+1} \rangle$. Using the inner product (see [25] for the proof), one can define the mapping in the other direction $\delta_k^* : \langle \Sigma_{k+1} \rangle \rightarrow \langle \Sigma_k \rangle$ as:

$$\delta_k^*(\overline{i_1 ... i_{k+2}}) = \sum_{q=1}^{k+2} (-1)^{q-1} (\overline{i_1 ... \hat{i_q} ... i_{k+2}}) \tag{39}$$

where $(\overline{i_1 ... \hat{i_q} ... i_{k+2}})$ is $(\overline{i_1 ... i_{k+2}})$ with the index i_q removed if $(\overline{i_1 ... \hat{i_q} ... i_{k+2}})$ exists in $\langle \Sigma_k \rangle$. Taking as an example triangle (123) in Figure 19:

$$\delta_k^*(123) = -(23) + (13) - (12)$$

A. Hodge Decomposition Under the Simplicial Complex Framework

Considering, as previously, $G = (\Sigma_0, ..., \Sigma_K)$, the Hodge decomposition partitions a subspace $\langle \Sigma_k \rangle$, with $0 \leq k < K$, into a direct sum of three orthogonal subspaces:

$$\langle \Sigma_k \rangle = G_k \oplus H_k \oplus S_k \tag{40}$$

where G_k is the image of $\langle\Sigma_{k-1}\rangle$ under δ_{k-1}, S_k is the image of $\langle\Sigma_{k+1}\rangle$ under δ_k^* and H_k contains elements of $\langle\Sigma_k\rangle$ orthogonal both to G_k and S_k. The Hodge decomposition in the simplicial formalism allows the decomposition of any field embedded in any dimension $k < K$. In this work, the subspace considered has been the edge subspace, of dimension 1.

B. Code Availability

All the code and data used to reproduce the plots and results of this work are available at GitHub https://github.com/COSIN3-UOC/Hodge_decomposition.git.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisors, Javier Borge-Holthoefer and Albert Solé-Ribalta, for their invaluable support, mentorship, and time devoted to this research. Their guidance and expertise have played a pivotal role in shaping this work and expanding my understanding of urban complex networks. I also extend my gratitude to Josep Perelló Palou for his availability and for serving as a crucial link between the two parties involved in this collaboration.

[1] L. Bettencourt and G. West, "A unified theory of urban living," *Nature*, vol. 467, pp. 912–3, 10 2010. [Online]. Available: https://doi.org/10.1038/467912a

[2] M. Batty, K. Axhausen, F. Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis, and Y. Portugali, "Smart cities of the future," *The European Physical Journal Special Topics*, vol. 214, pp. 481–518, 11 2012. [Online]. Available: https://doi.org/10.1140/epjst/e2012-01703-3

[3] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gomez, and A. Arenas, "Centrality in interconnected multilayer networks," *Physica D: Nonlinear Phenomena*, vol. 323, 11 2013.

[4] J. A. Ur, *Patterns of Settlement in Sumer and Akkad*. Oxford and New York: Routledge, 2013, pp. 131–155.

[5] D. Helbing, "A fluid dynamic model for the movement of pedestrians," 1998.

[6] F. Simini, G. Barlacchi, M. Luca, and L. Pappalardo, "A deep gravity model for mobility flows generation," *Nature Communications*, vol. 12, no. 1, p. 6576, Nov 2021. [Online]. Available: https://doi.org/10.1038/s41467-021-26752-4

[7] F. Simini, M. C. González, A. Maritan, and A.-L. Barabási, "A universal model for mobility and migration patterns," *Nature*, vol. 484, no. 7392, pp. 96–100, Apr 2012. [Online]. Available: https://doi.org/10.1038/nature10856

[8] D. Volchenkov and P. Blanchard, "Markov chain methods for analyzing urban networks," *Journal of Statistical Physics*, vol. 132, no. 6, pp. 1051–1069, Sep 2008. [Online]. Available: https://doi.org/10.1007/s10955-008-9591-2

[9] Y. Cheng, T. Zhang, and J. Wang, *Multi-agent System Model for Urban Traffic Simulation and Optimizing Based on Random Walk*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 703–711. [Online]. Available: https://doi.org/10.1007/978-3-642-12990-2_82

[10] H. Xiong, L. Yao, H. Tan, and W. Wang, "Pedestrian walking behavior revealed through a random walk model," *Discrete Dynamics in Nature and Society*, vol. 2012, p. 405907, Dec 2012. [Online]. Available: https://doi.org/10.1155/2012/405907

[11] J. Petit, R. Lambiotte, and T. Carletti, "Classes of random walks on temporal networks with competing timescales," *Applied Network Science*, vol. 4, no. 1, p. 72, Sep 2019. [Online]. Available: https://doi.org/10.1007/s41109-019-0204-6

[12] J. D. Noh and H. Rieger, "Random walks on complex networks," *Phys. Rev. Lett.*, vol. 92, p. 118701, Mar 2004. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.92.118701

[13] F. Wang, W. Chen, Y. Zhao, T. Gu, S. Gao, and H. Bao, "Adaptively exploring population mobility patterns in flow visualization," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 8, pp. 2250–2259, 2017.

[14] M. Mazzoli, A. Molas, A. Bassolas, M. Lenormand, P. Colet, and J. J. Ramasco, "Field theory for recurrent mobility," *Nature Communications*, vol. 10, no. 1, p. 3895, Aug 2019. [Online]. Available: https://doi.org/10.1038/s41467-019-11841-2

[15] T. Aoki, S. Fujishima, and N. Fujiwara, "Urban spatial structures from human flow by hodge–kodaira decomposition," *Scientific Reports*, vol. 12, no. 1, p. 11258, Jul 2022. [Online]. Available: https://doi.org/10.1038/s41598-022-15512-z

[16] K. Polthier and E. Preuß, "Identifying vector field singularities using a discrete hodge decomposition," in *Visualization and Mathematics III*, H.-C. Hege and K. Polthier, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 113–134.

[17] D. V. Anand, S. Das, and M. K. Chung, "Hodge-decomposition of brain networks," 2022. [Online]. Available: https://arxiv.org/abs/2211.10542

[18] T. Haruna and Y. Fujiki, "Hodge decomposition of information flow on small-world networks," *Frontiers in Neural Circuits*, vol. 10, 2016. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fncir.2016.00077

[19] U. Kan and E. López, "Layered hodge decomposition for urban transit networks," in *Complex Networks & Their Applications X*, R. M. Benito, C. Cherifi, H. Cherifi, E. Moro, L. M. Rocha, and M. Sales-Pardo, Eds. Cham: Springer International Publishing, 2022, pp. 804–815.

[20] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*. New York: Elsevier, 1976.

[21] P. Erdös and A. Rényi, "On random graphs i," *Publicationes Mathematicae Debrecen*, vol. 6, pp. 290–297, 1959.

[22] T. M. J. Fruchterman and E. M. Reingold, "Graph drawing by force-directed placement," *Software: Practice and Experience*, vol. 21, no. 11, pp. 1129–1164, 1991. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.4380211102

[23] A. A. Hagberg, D. A. Schult, and P. J. Swart, "Exploring network structure, dynamics, and function using networkx," in *Proceedings of the 7th Python in Science Conference*, G. Varoquaux, T. Vaught, and J. Millman, Eds., Pasadena, CA USA, 2008, pp. 11 – 15.

[24] H. Helmholtz, "Über integrale der hydrodynamischen gleichungen, welche den wirbelbewegungen entsprechen." vol. 1858, no. 55, pp. 25–55, 1858. [Online]. Available: https://doi.org/10.1515/crll.1858.55.25

[25] J. Johnson and T. Goldring, "Discrete hodge theory on graphs: A tutorial," *Computing in Science & Engineering*, vol. 15, pp. 42–55, 09 2013.

[26] A. Strang, "Applications of the helmholtz-hodge decomposition to networks and random processes," Ph.D. dissertation, 2020, copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works. [Online]. Available: https://www.proquest.com/dissertations-theses/applications-helmholtz-hodge-decomposition/docview/2461030154/se-2

[27] C. Moreno, Z. Allam, D. Chabaud, C. Gall, and F. Pratlong, "Introducing the "15-Minute City": Sustainability, Resilience and Place Identity in Future Post-Pandemic Cities," HAL, Post-Print hal-03549665, Mar. 2021. [Online]. Available: https://ideas.repec.org/p/hal/journl/hal-03549665.html

[28] D. Rhoads, A. Solé-Ribalta, and J. Borge-Holthoefer, "The inclusive 15-minute city: Walkability analysis with sidewalk networks," *Computers, Environment and Urban Systems*, vol. 100, p. 101936, 03 2023.

[29] M. E. J. Newman, "A measure of betweenness centrality based on random walks," 2003.

[30] R. Banisch, N. Djurdjevac Conrad, and C. Schütte, "Reactive flows and unproductive cycles for random walks on complex networks," *The European Physical Journal Special Topics*, vol. 224, no. 12, pp. 2369–2387, Sep 2015. [Online]. Available: https://doi.org/10.1140/epjst/e2015-02417-8