



Walk Based Laplacians for Modeling Diffusion on Complex Networks

Francesca Arrigo¹ · Fabio Durastante²

Received: 20 January 2026 / Revised: 16 July 2026 / Accepted: 30 July 2026
© The Author(s) 2026

Abstract

We develop a novel framework for modeling diffusion on complex networks by constructing Laplacian-like operators based on walks around a graph. Our approach introduces a parametric family of walk-based Laplacians that naturally incorporate memory effects by excluding or downweighting backtracking trajectories, where walkers immediately revisit nodes. The framework includes: (i) walk-based Laplacians that count all traversals in the network; (ii) nonbacktracking variants that eliminate immediate reversals; and (iii) backtrack-downweighted variants that provide a continuous interpolation between these two regimes. We establish that these operators extend the definition of the standard Laplacian and also preserve some of its properties. We present efficient algorithms using Krylov subspace methods for computing them, ensuring applicability of our proposed framework to large networks. Extensive numerical experiments on real-world networks validate the modeling flexibility of our approach and demonstrate the computational efficiency of the proposed algorithms, including GPU acceleration.

Keywords Complex network · Laplacian · Diffusion · Nonbacktracking walks · Memory effects

Mathematics Subject Classification 05C82 · 65F60 · 05C50 · 91D30

1 Introduction

Understanding diffusion and information spread in complex networks is fundamental in numerous applications, ranging from epidemiology and social networks to transportation systems and biological processes; see [39] and references therein. The graph Laplacian operator L is a widely used tool for modeling such phenomena, with its classical definition only

✉ Francesca Arrigo
francesca.arrigo@strath.ac.uk
Fabio Durastante
fabio.durastante@unipi.it

¹ Department of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow, G1 1XH Glasgow City, Scotland

² Department of Mathematics, University of Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa, PI, Italy

incorporating nearest-neighbor interactions. Over the past decade, several new extensions have been developed in order to capture longer-range interactions and incorporate memory effects in the exploration of networks [39]. These extensions can be broadly classified into two families; top-down approaches construct non-local diffusion operators through fractional powers of the standard graph Laplacian (e.g., L^γ with $\gamma \in (0, 1)$) [7, 43, 44], enabling super-diffusive and sub-diffusive behaviors; and bottom-up approaches which, on the other hand, explicitly construct transition matrices based on the enumeration of specific path structures [16, 19, 26, 28], offering greater interpretability and flexibility in modeling network exploration. Both these approaches, however, have an extremely high computational cost; indeed, fractional Laplacians require the explicit construction of matrix functions [34] whereas path-based approaches demand costly all-pairs shortest-path computations [30, 46, 52].

The key innovation of this work is the development of a flexible bottom-up framework built on *walks* rather than *paths*. This immediately yields computational advantages while at the same time allowing fine control over diffusion dynamics through the incorporation of memory effects. Our framework further encompasses a parametric family of Laplacian-like operators that naturally exclude or downweight backtracking trajectories, ensuring that walkers are penalised for immediately revisiting nodes they have just left. This enables us to capture realistic exploratory behavior on networks and introducing memory effects in our model.

Our approach unifies and extends several existing concepts. Further, it incorporates the use of nonbacktracking walks, that have so far mainly been studied for centrality measures [3, 4, 31]. Their systematic application to diffusion operators and the incorporation of graded backtrack-downweighting [5] are novel contributions.

The proposed framework retains the same interpretability of previously introduced operators defined in terms of enumerating paths [16, 19, 26, 28]. However, the family of operators we introduce offers several advantages when compared to those; (i) the operators naturally incorporate memory through nonbacktracking and backtrack-downweighted formulations, providing a parametric interpolation between standard diffusion to purely nonbacktracking dynamics; (ii) the matrix functions appearing in our definitions do not need to be built, as their action over a vector can be efficiently computed via Krylov subspace methods without assembling dense matrices. This in turn ensures scalability and applicability of our framework to large networks; and (iii) the diffusion equations induced by the new operators preserve properties of the standard Laplacian, thus ensuring theoretical consistency.

The remainder of the paper is organized as follows. Section 2 presents notation and relevant background on graph theory and the classical diffusion equation. Section 3 develops the main framework, introducing k -walk Laplacians and their backtrack-downweighted variants, establishing their mathematical properties and interpretations through their connection with network centrality concepts. Section 4 addresses computational challenges and proposes efficient algorithms for the computation of matrix–vector products, trace estimation, and linear system solutions tailored to these new operators. Section 5 validates the approach through numerical experiments on real-world networks, comparing diffusion behavior, computational scaling, and the effect of backtrack-downweighting.

2 Background and Notation

One of the most common ways of representing complex networks is by means of graphs. Formally, an *unweighted graph* is defined as an ordered pair of sets $G = (V, E)$, where

$V = \{1, 2, \dots, n\}$ is the set of *nodes* and $E \subseteq V \times V$ is the set of *edges*. Henceforth, the edge $(i, j) \in E$ from node i to node j will be equivalently represented using the notation $i \rightarrow j$. Whenever E is a set (i.e., it does not contain repeated elements) that is symmetric (i.e., $(i, j) \in E \iff (j, i) \in E$) and antireflective (i.e., $(i, i) \notin E$ for all $i \in V$), G will be called *simple*. Throughout this work, we will only consider simple unweighted graphs.

We are interested in linear algebraic representations of graphs. Perhaps the most widely used matrix representation of networks is the adjacency matrix.

Definition 1 Let $G = (V, E)$ be an unweighted graph with n nodes. Its *adjacency matrix* $A \in \mathbb{R}^{n \times n}$ is entrywise defined as:

$$A_{ij} = \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{otherwise} \end{cases} \quad \forall i, j = 1, 2, \dots, n.$$

When the underlying graph G is simple the adjacency matrix is symmetric with zeros on the main diagonal.

The constructions that we will present in the next sections all build on the idea of walks around a graph.

Definition 2 A *walk of length k* from i to j is a sequence of $k + 1$ nodes $i_1, i_2, \dots, i_{k+1}$ such that $i_1 = i$, $i_{k+1} = j$, and $(i_\ell, i_{\ell+1}) \in E$ for all $\ell = 1, 2, \dots, k$. If $i_\ell \neq i_\kappa$ for all $\ell \neq \kappa$, then the walk is called a *path*.

While edges describe immediate connections among nodes, the concept of walk can be used to analyse long range interactions between pairs of nodes. This is showcased by the following classical result.

Proposition 1 ([21, Ex. 34, Chapter 1]) Let A be the adjacency matrix of an unweighted graph $G = (V, E)$. Then, the number of walks of length $k \in \mathbb{N}$ from node i to node j is given by $(A^k)_{ij}$ for all $i, j \in V$.

An immediately related concept is that of connectivity of a simple graph.

Definition 3 A simple graph is deemed *connected* if there exists a walk connecting all its nodes and *disconnected* otherwise. Whenever G is disconnected, its maximal connected subgraphs are referred to as its *connected components*.

In linear algebraic terms, a connected graph G will correspond to an irreducible adjacency matrix A .

The idea of path, paired with Proposition 1, allows one to define a concept of distance between nodes in connected graphs. More specifically, the *geodesic* or *shortest path distance* between nodes i and j in a simple connected graph is defined as the length of the shortest path connecting the two nodes, i.e.,

$$d(i, j) = \min\{k \in \mathbb{N} \mid (A^k)_{ij} > 0\}.$$

Without loss of generality, from this point on we will assume that the graphs under consideration are unweighted, simple, and connected. Whenever a graph is disconnected, our theory can be independently applied to the individual connected components. We will also assume the following notation; $I \in \mathbb{R}^{n \times n}$ is the identity matrix, $O \in \mathbb{R}^{n \times n}$ is the matrix of all zeros, $\mathbf{0}, \mathbf{1} \in \mathbb{R}^n$ are the (column) vector of all zeros and all ones, respectively, $\mathbf{e}_i$ is the i th (column) vector of the standard basis of $\mathbb{R}^n$, and $\|\cdot\|$ denotes the Euclidean norm. Further, $\sigma(M)$ and $\rho(M)$ denote the spectrum and spectral radius of a matrix $M \in \mathbb{R}^{n \times n}$, respectively.

2.1 The graph Laplacian and The Diffusion Equation

The (combinatorial) graph Laplacian is a fundamental matrix in graph theory and spectral graph theory.

Definition 4 Let $G = (V, E)$ be a simple graph with adjacency matrix A . We define the Laplacian matrix as

$$L = D - A,$$

where A is the adjacency matrix of G and $D \in \mathbb{R}^{n \times n}$ is the degree matrix, a diagonal matrix with $D_{ii} = (A\mathbf{1})_i$ for all $i \in V$.

The graph Laplacian is a particular instance of a more general matrix construction that will be useful throughout the paper; see Lemma 1 below. We first recall the following definitions [12].

Definition 5 A matrix $M \in \mathbb{R}^{n \times n}$ is called a *Z-matrix* if all its off-diagonal entries are nonpositive, i.e., $M_{ij} \leq 0$ for all $i \neq j$.

Definition 6 A matrix $M \in \mathbb{R}^{n \times n}$ is called an *M-matrix* if it is a Z-matrix and all its eigenvalues have nonnegative real parts, i.e.,

$$\operatorname{Re}(\lambda) \geq 0 \quad \text{for all } \lambda \in \sigma(M).$$

We note in passing that several equivalent properties can be used to show that a given Z-matrix is an M-matrix; see, e.g., [45, Theorem 2.5.3].

Lemma 1 Let $C \in \mathbb{R}^{n \times n}$ be a symmetric nonnegative matrix. Then, the matrix $M = \operatorname{diag}(C\mathbf{1}) - C$ is a symmetric, singular M-matrix and, moreover, $M\mathbf{1} = \mathbf{0}$.

Proof Symmetry of M holds true by construction, since $C = C^\top$. Moreover, by construction, since $C \geq 0$, the off-diagonal entries of M satisfy $M_{ij} = -C_{ij} \leq 0$ for $i \neq j$. Thus M is a symmetric Z-matrix and $\sigma(M) \subset \mathbb{R}$. We now show that $\sigma(M) \subseteq [0, \infty)$, hence that M is an M-matrix, by proving that M is positive semidefinite. To this end, consider any vector $\mathbf{x} \in \mathbb{R}^n$. Exploiting the definition of M and symmetry of C , it can be shown that $2\mathbf{x}^\top M\mathbf{x} = \sum_{i,j=1}^n C_{ij}(x_i - x_j)^2$, which is nonnegative. Thus M is positive semidefinite. Finally, it is readily verified that $M\mathbf{1} = (\operatorname{diag}(C\mathbf{1}) - C)\mathbf{1} = \operatorname{diag}(C\mathbf{1})\mathbf{1} - C\mathbf{1} = \mathbf{0}$. $\square$

Taking $C = A$ in Lemma 1 immediately gives that the combinatorial graph Laplacian $L = \operatorname{diag}(A\mathbf{1}) - A$ is a singular M-matrix. For the purpose of this work, the relevance of L being an M-matrix lies in its relation to the diffusion equation on a graph. Indeed, non-singular M-matrices are inverse positive. As such, they ensure that the discrete diffusion model inherently respects a maximum principle, physically guaranteeing that the evolution smoothly transports quantities from high concentrations to low concentrations, without producing negative values or spurious oscillation. For singular M-matrices this corresponds to a closed system where mass (or thermal energy) is strictly conserved; the physical process will smooth out spatial gradients into a uniform, non-zero equilibrium state preserving the initial total mass (or thermal energy). Therefore, the evolution under an M-matrix dynamic mimics the behavior of the continuous diffusion operator which satisfies the continuous maximum principle; we refer the interested reader to [15] for a detailed discussion of the connection between the continuous and discrete maximum principles.

Given a row vector $\mathbf{p}_0 \geq 0$ with $\mathbf{p}_0 \mathbf{1} = 1$, the diffusion equation is

$$\begin{cases} \frac{d\mathbf{p}(t)}{dt} = -\mathbf{p}(t)L, & t \in [0, t^*], \\ \mathbf{p}(0) = \mathbf{p}_0, \end{cases} \quad (1)$$

which is solved, for all times t , by the probability distribution

$$\mathbf{p}(t) = \mathbf{p}_0 \exp(-tL).$$

This version of the diffusion equation is often referred to as *Poissonian edge centric* [39], since the transition rate to a given node is determined by its degree; a walker tends to leave high-degree nodes faster than low-degree ones.

Since the combinatorial Laplacian L is a singular M-matrix, its eigenvalues are all non-negative. Hence, the negative sign in (1) ensures that the substance diffusing on the graph flows from regions of higher concentration to lower concentration, mirroring the behavior of heat diffusion in physical systems. The eigenvalues and eigenvectors of L significantly influence the behavior of the diffusion process, particularly in characterizing the long-term dynamics and equilibrium states of (1), which is readily shown to be collinear to $\mathbf{1}$. Indeed, the interesting part of the diffusion phenomena is the transient phase.

3 Going for a Stroll: k -Walk Laplacians and their Backtrack-Downweighted Version

The Laplacian matrix L is defined by only considering immediate connections between nodes in the graph, and therefore it captures primarily local phenomena. Recent literature has shown a growing interest in incorporating longer-range interactions within the definition of variants of the Laplacian. The approaches to doing this can be broadly categorized into two groups; a top-down approach, exemplified by the use of the fractional Laplacian L^γ with $\gamma \in (0, 1)$, which generalizes the discrete Laplacian L through the use of a matrix function [7, 44], and a bottom-up approach, where the paths in the graph are utilized in the definition of new measures [26, 28, 43]. In the following, we will construct a bottom-up approach that considers walks instead of paths. This will allow us to define a parametric family of Laplacians that on the one side capture a walker's short-term memory, and on the other allow to perform long jumps on the graph. These jumps will occur between two nodes i and j with a probability depending not only on the length of the walks connecting i and j , but also on the number of such walks.

We begin by recalling the construction of the k -path transformed Laplacian $\mathcal{L}$, which was introduced in [26, 28] as an alternative to the fractional powers of L . In [28] the authors propose to build a sequence of Laplacians, each encoding information on paths occurring between nodes of a specific length only, if any. If we let $\text{diam}(G) = \max\{d(i, j) \mid i, j \in V\}$ be the *diameter* of the graph, i.e., the length of the longest shortest path in G , then, for each $k = 1, 2, \dots, \text{diam}(G)$ the associated k -path Laplacian $\mathcal{L}_k$ of G is defined entrywise as follows;

$$(\mathcal{L}_k)_{ij} = \begin{cases} \mathcal{N}_k(i), & \text{if } i = j, \\ -1, & \text{if } d(i, j) = k, \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

where $\mathcal{N}_k(i) = |\{j \in V \mid d(i, j) = k\}|$ is the number of nodes at distance k from node i . For $k > \text{diam}(G)$, $\mathcal{L}_k = O$. Clearly, this definition extends Definition 4 in the sense that these

matrices retain the “idea” of how L is built (“degrees” on the diagonal and -1 in selected off-diagonal entries).

The k -path transformed Laplacian, which is predominantly used in practice, is then defined as;

$$\mathcal{L} = \sum_{k=1}^{\infty} t_k \mathcal{L}_k = \sum_{k=1}^{\text{diam}(G)} t_k \mathcal{L}_k, \quad t_k \in \left\{ \frac{1}{k^\beta}, \exp(-\beta k) \right\}_{\beta \in \mathbb{R}_+, k \in \mathbb{N}}. \quad (3)$$

A striking limitation of this approach is that it requires computing and storing the distances between any two nodes in the graph, a computationally expensive and memory-taxing procedure; see Sect. 4 for more details.

Because of this limitation, and because of how much simpler and cheaper it is to count walks rather than paths, we seek to introduce the alternative definition of k -walk Laplacians and of the associated k -walk transformed Laplacian; see Definitions 7 and 8 below. This will be achieved by exploiting Proposition 1. From the modeling viewpoint, for any $k \in \mathbb{N}$ a walker will move from node i to node j on the graph with a weight that is proportional to the number of walks of length k between the two.

Definition 7 For any $k \in \mathbb{N}$, the k -walk Laplacian is defined as

$$\mathbb{L}_k = \text{diag}(A^k \mathbf{1}) - A^k. \quad (4)$$

We point out that, differently from their path-based equivalents $\mathcal{L}_k$, these matrices may be dense even for sparse graphs. The sparsity (or otherwise) of $\mathbb{L}_k$ will depend on the connectivity structure of the graph; for example, whenever A is primitive¹ there exists $k > 0$ such that $A^k > 0$. Still, the computation of $\mathbb{L}_k$ is much simpler and cheaper. In fact, in Sect. 4 we will show that these matrices do not need to be explicitly built, differently from $\mathcal{L}_k$. The next step in our construction is to use the information from all the $\mathbb{L}_k$ to build a matrix that accounts for walks of any length $k = 1, 2, \dots$ in the graph. To ensure the convergence of the resulting matrix and to align with the intuition behind the construction of walk-based centralities [29], we present a definition that includes a monotonically non-increasing sequence of coefficients $\{c_k\}_{k \geq 0}$ to weight the contributions of walks of different lengths.

Definition 8 For any sequence $\{c_k\}_{k \geq 0}$ that is monotonically non-increasing for $k \geq 1$, and such that $\sum_{k=0}^{\infty} c_k A^k$ is convergent, the k -walk transformed Laplacian $\mathbb{L}(\mathbf{c})$ is defined as:

$$\mathbb{L}(\mathbf{c}) = \text{diag} \left(\sum_{k=0}^{\infty} c_k A^k \mathbf{1} \right) - \sum_{k=0}^{\infty} c_k A^k = \sum_{k=1}^{\infty} c_k \mathbb{L}_k.$$

It is easy to verify that, for $\mathbf{c} = (0, 1, 0, 0, \dots)$, we recover $\mathbb{L}(\mathbf{c}) = L$.

We note that, differently from the k -path transformed Laplacian $\mathcal{L}$ in (3), the k -walk transformed Laplacian $\mathbb{L}(\mathbf{c})$ does not reduce to a finite sum. However, whenever the coefficients $\{c_k\}_{k \geq 0}$ in Definition 8 are chosen so that there exists a function f for which $f(x) = \sum_{k \geq 0} c_k x^k$ on an open set $\Omega \subset \mathbb{C}$ containing the spectrum of A [34, Definition 1.1], then

$$\mathbb{L}(f) = \text{diag}(f(A) \mathbf{1}) - f(A). \quad (5)$$

In the following, we will restrict ourselves to sequences of coefficients that allow for the latter representation of the k -walk transfer Laplacian.

¹ An adjacency matrix is said to be primitive if it is irreducible and has only one eigenvalue of maximum modulus; see, e.g., [36, Definition 8.5.0, Theorem 8.5.2].

We postpone the discussion on the computational issues related to the evaluation of $\mathbb{L}(f)$ to Sect. 4. The remainder of this section is devoted to theoretical properties of $\mathbb{L}(f)$ and to one generalization.

3.1 Analogies and Relationship with the Standard Laplacian

We begin by pointing out that we can refine the observation made in Sect. 2.1 about the properties of the diffusion equation (1), expressed through the Laplacian matrix L , and highlight the connections between the classical construction and the framework introduced here.

The following proposition summarizes classical results that we used in Sect. 2 and that we will show extend to $\mathbb{L}(f)$.

Proposition 2 *For an unweighted, simple graph $G = (V, E)$, the Laplacian matrix L satisfies the following.*

- L is symmetric;
- $L\mathbf{1} = \mathbf{0}$; and
- L is an M-matrix.

We have already discussed within Sect. 2.1 the relationship between the diffusion equation and the fact that L is an M-matrix. The following result shows that, under suitable assumptions on f , $\mathbb{L}(f)$ is also an M-matrix, and thus it is possible to define a diffusion equation on the graph extending (1) based on the idea of walks as

$$\begin{cases} \frac{d\mathbf{p}(t)}{dt} = -\mathbf{p}(t)\mathbb{L}(f), & t \in [0, t^*], \\ \mathbf{p}(0) = \mathbf{p}_0. \end{cases} \quad (6)$$

Proposition 3 *Let $G = (V, E)$ be a simple connected graph and let f be a function defined on the spectrum of its adjacency matrix $A \in \mathbb{R}^{n \times n}$; cf. [34, Definition 1.1]. Then the k -walk transformed Laplacian $\mathbb{L}(f)$ in (5) is such that*

- $\mathbb{L}(f)$ is symmetric;
- $\mathbb{L}(f)\mathbf{1} = \mathbf{0}$; and
- If $f(A) \geq 0$, then $\mathbb{L}(f)$ is an M-matrix.

Proof The matrices $f(A)$ and $\mathbb{L}(f)$ are symmetric by construction, since $A^\top = A$. The fact that $\mathbb{L}(f)\mathbf{1} = \mathbf{0}$ follows trivially from (5). Suppose now that $f(A) \geq 0$. Then, by Lemma 1, $\mathbb{L}(f)$ is a singular M-matrix. $\square$

For the purposes of this work, Proposition 3 constrains our choice of f to those functions that preserve nonnegativity. Characterizing these is a well studied problem and several results are available in the literature; see, for example, [13, 23] and references therein. If we restrict our attention to functions f that generate admissible centrality indices, such as those belonging to class $\mathcal{P}$ as introduced in [9],² then the induced $\mathbb{L}(f)$ is an M-matrix. Furthermore, selecting $f \in \mathcal{P}$ allows us to naturally interpret the role of $f(A)$ within (5) in terms of centrality and to provide insight in the diffusion process modeled by $\mathbb{L}(f)$. Indeed, recall

² The class $\mathcal{P}$ contains all analytic functions that can be expressed as sums of power series with strictly positive coefficients on some open neighborhood of 0.

that the vector of total communicability centralities [8] of nodes is defined, for a function f defined on the spectrum of A , as:

$$\mathbf{t}(f) = f(A)\mathbf{1}.$$

When $f(x) = x$, it immediately follows that $\mathbf{t}(f) = (A\mathbf{1})$, and thus that $\mathbb{L}(f) = L$. However, for other choices of $f \in \mathcal{P}$, the introduced k -walk transformed Laplacian is simply generalizing L by replacing degree centrality $A\mathbf{1}$ with the total communicability $\mathbf{t}$ of nodes. With this interpretation, we can deduce that the edge centric diffusion equation (6) induced by $\mathbb{L}(f)$ is one where a walker tends to leave high total communicability nodes faster than low total communicability ones. Clearly, when $f(x) = x$, (6) boils down to (1)

3.2 Backtrack-Downweighted k -Walk Transformed Laplacian

As described in the introduction, we are interested in defining a generalization of the Laplacian operator that accounts for walks on graphs. This we have partially achieved already in the previous section. Our next goal is to further generalize the definition in order to allow the flexibility to downweight (or possibly eliminate) certain walks in our count.

Definition 9 Given a graph $G = (V, E)$ without loops, a walk on G is said to be *backtracking* if it contains at least one instance of the form $i \rightarrow j \rightarrow i$, for $i, j \in V$, *nonbacktracking* (NBT) otherwise.

We now introduce matrices $p_k(A) \in \mathbb{R}^{n \times n}$, for $k \in \mathbb{N}$, which generalize to the NBT setting the idea that powers of the adjacency matrix encode the number of walks of a certain length in the graph; see Proposition 1. These matrices are entrywise defined as

$$(p_k(A))_{ij} = |\{\text{NBT walks of length } k \text{ from } i \text{ to } j\}|, \quad k \in \mathbb{N}, i, j \in V$$

and are symmetric by construction. It can be shown [14], [50, Lemma 1] that the matrices $p_k(A)$ satisfy a three terms recurrence. Setting $p_0(A) = I$, $p_1(A) = A$ and $p_2(A) = A^2 - D$, where D is the degree matrix, for all $k \geq 3$ it holds

$$p_k(A) = A p_{k-1}(A) + (I - D) p_{k-2}(A). \quad (7)$$

This result was generalized in [5], where backtracking was downweighted by a factor $\theta = 1 - \mu$, with $\mu \in [0, 1]$, rather than completely eliminated.

Following [5], we will let $q_k(A)$ denote the *backtrack-downweighted walk* (BTDW) count matrix, where, for brevity, the dependence on μ is not made explicit. More precisely, $(q_k(A))_{ij}$ is the weighted number of distinct walks of length k from node i to node j where the weights are assigned according to the following; for each walk, $i = i_0, i_1, i_2, \dots, i_k = j$, every occurrence of a backtracking step ($i_\ell = i_{\ell+2}$) incurs a downweighting by a factor $1 - \mu$. Then, it can be shown that, setting $q_0(A) = I$, $q_1(A) = A$ and $q_2(A) = A^2 - \mu D$, the BTDW matrices satisfy a three terms recurrence

$$q_{k+1}(A) = A q_k(A) + \mu (\mu I - D) q_{k-1}(A). \quad (8)$$

Equation (8) specifies [5, Theorem 3.1] to the case of undirected graphs.

It is easily verified that for $\mu = 0$, the relation (8) reduces to the counting formula in Proposition 1, while for $\mu = 1$ it recovers (7). We are now in the position to construct the desired extension of the Laplacian matrix.

3.2.1 Building the BTDW Laplacian

The definition of k -walk transformed Laplacian (5) we have put forward considers all traversals on the graph, in contrast to the k -path transformed Laplacian (3), which only allows paths. A middle ground between these two approaches, which specializes the walk count while not becoming too computationally taxing, is to consider downweightings of backtracking walks.

Definition 10 In the above notation, for $\mu \in [0, 1]$, the *backtrack-downweighted (BTDW) k -walk Laplacian* is defined as

$$\mathbb{L}_{k,\mu} = \text{diag}(q_k(A)\mathbf{1}) - q_k(A).$$

There are two substantial differences between the k -path and the BTDW k -walk Laplacians, besides the type of traversals being considered. The first difference is that there exist both a maximal and a minimal k -path Laplacian. Specifically, $\mathcal{L}_k$ is the zero matrix for every $k \geq \text{diam}(G)$, while $\mathbb{L}_{k,\mu}$ is non-zero for every k , as long as $\mu \neq 1$ or G is not a tree. The second difference is that, in the construction of $\mathbb{L}_{k,\mu}$, we consider weighted “adjacency matrices” $q_k(A)$ that account for BTDW walks, whereas the “adjacency matrices” appearing in the construction of the k -path Laplacian are unweighted, as their entries [27, Equation 3.1] equal one if *at least one* path of the given length exists between the corresponding two nodes.

To build BTDW k -walk transfer Laplacian, we need to characterize the convergence of the series $\sum_{k=0}^{\infty} c_k q_k(A)$ and the function it converges to. To this end, we define for any integer $s \geq 0$,

$$f_s(x) = \sum_{k=0}^{\infty} c_{s+k} x^k, \quad (9)$$

and prove a version of [4, Theorem 5.4] for the undirected case.

Proposition 4 Assume that the series (9) for $s = 0$ has finite or infinite radius of convergence $\mathcal{R}$, and let $\mu \in [0, 1]$. Then

$$\phi(A, \mathbf{c}) = \sum_{k=0}^{\infty} c_k q_k(A)$$

converges for $\rho(Z) < \mathcal{R}$, where $\rho(Z)$ is the spectral radius of

$$Z = \begin{bmatrix} O & I \\ \mu(\mu I - D) & A \end{bmatrix}. \quad (10)$$

Furthermore,

$$\phi(A, \mathbf{c}) = \begin{bmatrix} O & I \end{bmatrix} f_2(Z) \begin{bmatrix} q_1(A) \\ q_2(A) \end{bmatrix} + c_1 A + c_0 I = \begin{bmatrix} I & O \end{bmatrix} f_1(Z) \begin{bmatrix} q_1(A) \\ q_2(A) \end{bmatrix} + c_0 I.$$

Proof We begin by expanding the defining series

$$\phi(A, \mathbf{c}) = \sum_{k=0}^{\infty} c_k q_k(A) = c_0 I + c_1 A + \sum_{k=2}^{\infty} c_k q_k(A). \quad (11)$$

It is easy to verify by induction that

$$Z^k \begin{bmatrix} q_1(A) \\ q_2(A) \end{bmatrix} = \begin{bmatrix} q_{k+1}(A) \\ q_{k+2}(A) \end{bmatrix}.$$

This, together with (11), yields:

$$\sum_{k=2}^{\infty} c_k q_k(A) = \sum_{k=0}^{\infty} c_{k+2} q_{k+2}(A) = [O \ I] \left(\sum_{k=0}^{\infty} c_{k+2} Z^k \right) \begin{bmatrix} q_1(A) \\ q_2(A) \end{bmatrix},$$

and thus the first expression for $\phi(A, \mathbf{c})$ is readily recovered using (9) with $s = 2$. A similar derivation allows to obtain the second expression in our statement. Finally, since f_0 has radius of convergence $\mathcal{R}$, analytic functional calculus gives that $f_2(Z)$ converges whenever $\rho(Z) < \mathcal{R}$, which establishes both convergence and the closed-form expressions. $\square$

Corollary 1 For the case $\mu = 1$ the identity in Proposition 4 reduces to:

$$\phi(A, \mathbf{c}) = \sum_{k=0}^{\infty} c_k p_k(A) = [I \ O] f_1(Z) \begin{bmatrix} A \\ A^2 - D \end{bmatrix} + c_0 I.$$

Remark 1 If $\mathcal{R} < \infty$ and $\rho(Z) > \mathcal{R}$ in either Proposition 4 or Corollary 1, one can introduce a parameter $0 < \alpha < 1$ to rescale the series in (9) as $f_s(\alpha x)$. Such α will thus need to be chosen so that $\alpha \rho(Z) < \mathcal{R}$.

We now mirror the construction in Definition 8 and introduce an operator that includes BTDW walks of all possible lengths.

Definition 11 For any sequence $\{c_k\}_{k \geq 0}$ that is monotonically non-increasing for $k \geq 1$ and such that $\sum_{k=0}^{\infty} c_k Z^k$ is convergent, and for any $\mu \in [0, 1]$ we define the BTDW k -walk transformed Laplacian as

$$\mathbb{L}_{\mu}(\mathbf{c}) = \text{diag} \left(\sum_{k \geq 0} c_k q_k(A) \mathbf{1} \right) - \sum_{k \geq 0} c_k q_k(A) = \sum_{k=1}^{\infty} c_k \mathbb{L}_{k, \mu}. \quad (12)$$

The following result, which is the analogue of Proposition 3 for $\mathbb{L}_{\mu}(\mathbf{c})$, describes the characterization we need in order to formulate a diffusion equation employing the newly introduced operator.

Proposition 5 Let $G = (V, E)$ be a simple connected graph and let f_s , $s = 0, 1$, be a function defined on the spectrum of $Z \in \mathbb{R}^{2n \times 2n}$ for Z as in Proposition 4. Then the BTDW k -walk transformed Laplacian $\mathbb{L}_{\mu}(\mathbf{c})$, $\mu \in (0, 1]$, in (12) is such that

- $\mathbb{L}_{\mu}(\mathbf{c})$ is symmetric;
- $\mathbb{L}_{\mu}(\mathbf{c}) \mathbf{1} = \mathbf{0}$; and
- If $\phi(A, \mathbf{c}) \geq 0$, then $\mathbb{L}_{\mu}(\mathbf{c})$ is an M -matrix.

Proof Symmetry holds true by construction. Moreover, $\mathbb{L}_{\mu}(\mathbf{c}) \mathbf{1} = \text{diag}(\phi(A, \mathbf{c}) \mathbf{1}) \mathbf{1} - \phi(A, \mathbf{c}) \mathbf{1} = \mathbf{0}$, which proves the second claim. Finally, suppose that $\phi(A, \mathbf{c}) \geq 0$. By taking $C = \phi(A, \mathbf{c})$ in Lemma 1, it immediately follows that $\mathbb{L}_{\mu}(\mathbf{c}) = \text{diag}(\phi(A, \mathbf{c}) \mathbf{1}) - \phi(A, \mathbf{c})$ is a singular M -matrix. $\square$

We point out that any function $f = f_0 \in \mathcal{P}$ ensures nonnegativity of $\phi(A, \mathbf{c})$ since

$$[\delta_{s,1} I \ \delta_{s,2} I] f_s(Z) \begin{bmatrix} q_1(A) \\ q_2(A) \end{bmatrix} = \sum_{k=0}^{\infty} c_{k+s} q_k(A) \geq 0$$

where $\delta_{i,j}$ is the Kronecker delta. The added value of selecting $f \in \mathcal{P}$ is that any of these functions would allow the interpretation of $\mathbb{L}_{\mu}(\mathbf{c})$ in terms of centralities [3]. Henceforth, for all $\mathbf{c}$ obtained from $f \in \mathcal{P}$ we will write $\mathbb{L}_{\mu}(f)$ to denote $\mathbb{L}_{\mu}(\mathbf{c})$.

By means of Definition 11 we can extend the diffusion equation (6) to

$$\begin{cases} \frac{d\mathbf{p}(t)}{dt} = -\mathbf{p}(t)\mathbb{L}_\mu(f), & t \in [0, t^*], \\ \mathbf{p}(0) = \mathbf{p}_0, \end{cases} \quad \mu \in [0, 1], \quad (13)$$

which reduces exactly to (6) for $\mu = 0$. In fact, it is easy to see that, when $\mu = 0$ and backtracking is not downweighted, $\mathbb{L}_0(\mathbf{c}) = \mathbb{L}(\mathbf{c})$ and moreover Proposition 5 reduces to Proposition 3. For $\mu = 1$ we get instead the NBT version of the k -walk transformed Laplacian. In the following, we provide two explicit examples of NBT Laplacians built using the two matrix functions most widely used in network science applications; the resolvent and the exponential.

The resolvent NBT k -walk transformed Laplacian. We first consider the case of the *resolvent*, when $\mathbf{c} = (1, \alpha, \alpha^2, \dots)$ with $\alpha\rho(Z) < 1$, which is a function in the $\mathcal{P}$ class. Then, by Corollary 1,

$$\phi(A, \mathbf{c})[I - \alpha A - \alpha^2(I - D)] = (1 - \alpha^2)I. \quad (14)$$

Setting $\mathcal{A}(\alpha) = I - \alpha A - \alpha^2(I - D)$, the corresponding NBT k -walk transformed Laplacian is

$$\mathbb{L}_1((1 - \alpha x)^{-1}) = (1 - \alpha^2) \left(\text{diag}(\mathcal{A}(\alpha)^{-1}\mathbf{1}) - \mathcal{A}(\alpha)^{-1} \right), \quad (15)$$

which correspond to the matrix having on the diagonal the NBT Katz centrality vector [3].

Remark 2 The matrix $\mathcal{A}(\alpha)$ is the *deformed graph Laplacian* which has been used to investigate consensus problems leveraging the spectral theory of quadratic eigenvalue problems [40]; furthermore, for undirected graphs, its inverse $\mathcal{A}(\alpha)^{-1}$ is known as the Bethe Hessian [48] and has been employed in spectral clustering tasks.

In general, $\mathcal{A}(\alpha)$ is symmetric but indefinite [40, p. 3050]. However, by choosing α appropriately, positive definiteness can be recovered; in particular, one may select α so that $\mathcal{A}(\alpha)$ becomes an M-matrix. The proposition below characterizes such choices of α , summarizing results from [31].

Proposition 6 *Let G be a connected simple graph with average degree > 2 . Then, for every $0 < \alpha < 1/\rho(Z)$, the deformed graph Laplacian $\mathcal{A}(\alpha)$ is an M-matrix.*

Remark 3 The hypothesis of a connected simple graph having average degree at least two rewrites as $2m > 2n$, where m is the number of undirected edges in the graph. Since normally $m \approx \log(n)$ in real-life networks, this hypothesis is always satisfied in practice.

Remark 4 Eq. 14 and Proposition 6 readily extend to the general case of $\mu \in [0, 1]$, by adapting the results in [5, Theorems 4.1 and 6.1] and employing again Lemma 1 to prove the property of being an M-matrix.

The exponential NBT k -walk transformed Laplacian. Another example of choosing the specific weights for (12) is that of the exponential $f_0(x) = e^x$ in (9), which corresponds to

$$f_1(x) = \sum_{k=1}^{\infty} \frac{1}{(k+1)!} x^k = \frac{e^x - 1}{x} = \varphi_1(x),$$

i.e., to the first φ -function from the theory of exponential integrators [35]

$$\varphi_0(x) = e^x, \quad \varphi_s(x) = \frac{1}{(s-1)!} \int_0^1 e^{(1-\theta)x} \theta^{s-1} d\theta, \quad s = 1, 2, \dots, \quad y \in \mathbb{C}.$$

These functions are entire and as such admit an infinite radius of convergence. The resulting exponential NBT k -walk transformed Laplacian is thus given by

$$\mathbb{L}_1(\varphi_1) = \text{diag} \left([I \ O] \varphi_1(Z) \begin{bmatrix} \mathbf{d} \\ A^2 \mathbf{1} - \mathbf{d} \end{bmatrix} \right) - [I \ O] \varphi_1(Z) \begin{bmatrix} A \\ A^2 - D \end{bmatrix}, \quad (16)$$

where $\mathbf{d} = A\mathbf{1} = D\mathbf{1}$.

4 Computational Considerations

In this section we detail different computational features of the newly introduced Laplacian operators $\mathbb{L}_\mu(\mathbf{c})$, making a direct comparison with their path-based counterpart $\mathcal{L}$ from [26, 28]. We also describe the algorithms implemented to perform the numerical experiments in the next Sect. 5.

The construction of k -path Laplacians $\mathcal{L}_k$ in (2) requires the computation of the geodesic distance $d(\cdot, \cdot)$ for every pair of nodes in the underlying network. This is one of the major disadvantages related to their use. In a directed weighted graph with positive or negative edge weights and no negative cycles this can be solved with the Floyd–Warshall algorithm [30, 46, 52] that has a worst-case performance of $\Theta(n^3)$. If the graph is also sparse then the cost can be brought down to $O(n^2 \log(n) + 2mn)$ with Johnson’s algorithm [37], where m is the number of edges in the graph. Both require storage of n^2 elements. It is also important to observe that, even in situations where only matrix–vector products involving $\mathcal{L}$ are needed, the computational cost is not reduced. Indeed, one still requires the computation of the entire distance matrix. The newly introduced Laplacians $\mathbb{L}_\mu(f)$, for $\mu \in [0, 1]$ and $f \in \mathcal{P}$, on the other hand, allow the implementation and use of significantly lower computational cost algorithms. Furthermore, even though the matrices are dense, they do not need to be stored. Therefore, they share the same advantages as top-down approaches based on the computation of matrix–vector products with the standard Laplacian matrix [7, 10]. Below we briefly address some strategies for tackling these computations efficiently.

4.1 Walk-Based Laplacians

If we compare the expression for $\mathcal{L}$ (3) with the one for $\mathbb{L}_\mu(f)$ (12), we can observe that, from a computational standpoint, the latter allows to compute matrix–vector products without ever assembling the whole matrix $\mathbb{L}_\mu(f)$. Krylov methods [32, 42] provide an efficient approach for approximating objects of the form $f(A)\mathbf{v}$, as we need to do for (12). In the simplest case of a polynomial method, given a matrix $A \in \mathbb{R}^{n \times n}$ and a vector $\mathbf{v} \in \mathbb{R}^n$, we first construct the Krylov subspace of order k , defined as $\mathcal{K}_k(A, \mathbf{v}) = \text{span}\{\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \dots, A^{k-1}\mathbf{v}\}$, by generating an orthonormal basis $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ for it using either the Arnoldi (for $A \neq A^\top$, [47, Algorithm 6.1]) or Lanczos (for $A = A^\top$, [47, Algorithm 6.15]) process. The basis is stored in a rectangular matrix $Q_k = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k]$ which is then used to form the Hessenberg (resp., tridiagonal for $A = A^\top$), matrix $H_k = Q_k^\top A Q_k \in \mathbb{R}^{k \times k}$. The matrix H_k represents the projection of A onto $\mathcal{K}_k(A, \mathbf{v})$. With this, one can then cheaply compute,³ the matrix function $f(H_k)$, since $k \ll n$, and then approximate $f(A)\mathbf{v}$ as $f(A)\mathbf{v} \approx \|\mathbf{v}\| Q_k f(H_k) \mathbf{e}_1$, where we have used the fact that $Q_k^\top \mathbf{v} = \|\mathbf{v}\| \mathbf{e}_1$. Krylov subspace methods provide accurate approximations to $f(A)\mathbf{v}$ without the need to directly compute $f(A)$ for large and sparse

³ The computation of $f(H_k)$ can be performed, e.g., via the Schur–Parlett method [34, §6.4] or, for particular choices of f , by means of specialized methods. We refer the interested reader to [34] for more details.

matrices, an operation that would normally require $O(n^2)$ storage and $O(n^3)$ computation. In some cases it may be convenient to move from polynomial methods to rational ones [24, 32]. In such instances, the cost to augment the dimension of the Krylov space increases, since we must move from the simple computation of matrix–vector products to the solution of a linear system at each space-expansion step. However, this usually has the advantage of significantly reducing the dimension of the space to be generated. Moreover, these methods leverage linear cost algorithms for the efficient solution of sparse linear systems; see Sect. 4.4 below.

4.1.1 Simulating Discrete-Time Diffusion

Using the same procedure, we can also compute the action of the probability transition matrix P of a Markov chain associated with the Laplacians in Definition 11 for $\mu \in [0, 1]$ ⁴ as

$$P = I - \mathbb{D}^{-1} \mathbb{L}_\mu(f), \quad \mathbb{D} = \text{diag}(\mathbb{L}_\mu(f)). \quad (17)$$

Constructing the diagonal matrix $\mathbb{D}$ only requires a single evaluation of the underlying matrix function; afterward, the same routine used to apply the chosen walk-based Laplacian may be reused efficiently.

Remark 5 This is not the case for the fractional Laplacian L^γ , for $\gamma \in (0, 1)$ [7, 10, 44] or for the k -path transformed Laplacian $\mathcal{L}$ (3). For the former, computing the diagonal matrix $D_\gamma = \text{diag}(L^\gamma \mathbf{1})$ needed to define $P = I - D_\gamma^{-1} L^\gamma$ requires n matrix function evaluations; for the latter, one has to actually compute the whole matrix $\mathcal{L}$, i.e., compute the distance matrix associated with the graph.

The matrix P in (17) describes a random walk on the graph in which the transition probability between nodes is inversely weighted with the total communicability measure induced by the selected function $f \in \mathcal{P}$. The evolution of a probability distribution then takes the form

$$\mathbf{p}_{k+1} = \mathbf{p}_k P, \quad \text{given } \mathbf{p}_0 \text{ with } \mathbf{p}_0 \mathbf{1} = 1.$$

We emphasize that, while the distributions $\mathbf{p}(t)$ in (1), (6), and (13) evolve continuously under a Laplacian operator, the discrete-time Markov chain defined by P can be interpreted as the natural discrete analogue of the continuous diffusion equations.

4.1.2 Limitations Due to Floating Point Arithmetic

For both Definitions 8 and 11, a note of caution is required regarding the magnitude of the entries of the matrices considered. When the function f exhibits rapid growth with respect to its argument, e.g., when f is the exponential function, the entries of $\mathbb{L}_\mu(f)$, $\mu \in [0, 1]$ may grow so large to approach the overflow threshold of double-precision arithmetic. This effect is particularly evident on the diagonal elements, as it was previously observed in [2]. In such cases, although Propositions 3 and 5 remain valid, the resulting matrices may be numerically undefined or lead to severe ill-conditioning in floating-point computation. A practical remedy to this issue is to rescale the spectrum by introducing an “inverse temperature” within the function; for example, instead of using $f(x) = \exp(x)$, one may consider $f(x) = \exp(\beta x)$ with $\beta \propto 1/\rho(A)$.

⁴ Recall that, when $\mu = 0$ we recover Definition 8.

4.2 The Resolvent NBT k -Walk Transformed Laplacian: $\mathbb{L}_1((1 - \alpha x)^{-1})$

Computation with the resolvent NBT k -walk transformed Laplacian requires the evaluation of matrix–vector products with $\mathbb{L}_1(f)$ from (15), where $f(x) = (1 - \alpha x)^{-1}$ for suitable values of $\alpha > 0$. Each of these amounts to solving a linear system defined in terms of the deformed Laplacian $\mathcal{A}(\alpha)$. If α is selected within the bounds from Proposition 6, then this solution can be computed using the Conjugate Gradient algorithm. In general, this method requires a suitable preconditioner to ensure fast convergence. Since $\mathcal{A}(\alpha)$ is an M-matrix, we can utilise Algebraic MultiGrid methods (AMG), which are known to be especially effective for such operators. This is because their spectral behavior and smoothing properties align well with the coarse space construction and relaxation strategies used in AMG. For our task, we evaluate the usage of an AMG from the `PSCToolkit` library [22]. Specifically, we employ an aggregation-based multigrid with the decoupled Vaněk, Mandel, and Brezina aggregation [51] in a V-cycle, using one sweep of ℓ_1 -Jacobi as smoother [6] and 30 smoother sweeps as coarse grid solver. We will show in Sect. 5 that this choice is computationally efficient and that it can also be used to perform the computation transparently on one or more GPU accelerators.

4.3 Exponential NBT k -walk Transformed Laplacian: $\mathbb{L}_1(\exp)$

The computation of matrix–vector products with the NBT k -walk framework Laplacian $\mathbb{L}_1(f)$ from (16) requires the evaluation of $\varphi_1(x)$. As previously mentioned, this is a common task when using exponential integrators for the numerical approximation of ODEs [35]. As it generally happens for the computation of matrix functions, the available algorithms are divided between the ones working on small-dimensional matrices [1, 49] and those based on Krylov-type methods [18, 42]. Here, we use the algorithm described and implemented in [42], which exploits a polynomial Krylov method.

4.4 Average Return Probability and its Randomized Estimates

The efficiency of a continuous time random walk can be studied by considering the *return probabilities*, i.e., the probability $[P(t)]_{j,j}$ of having started from node j at time 0 and being back in j at time t following any of the evolution equations

$$\begin{cases} \frac{d\mathbf{p}(t)}{dt} = -\mathbf{p}(t)\mathcal{L}, & t \in [0, t^*], \\ \mathbf{p}(0) = \mathbf{p}_0, \end{cases}$$

for some matrix $\mathcal{L}$. If the initial decay of return probabilities occurs rapidly, the information swiftly traverses from the initial node to the entire network. Conversely, a gradual decay implies that the information can only propagate through the network at a slower pace. The desirability of either scenario depends on the specific phenomenon we aim to describe. The *average return probability* for the evolution equations above is then defined as

$$\hat{p}_0(t) = \frac{1}{n} \sum_{i=1}^n [P(t)]_{ii} = \frac{1}{n} \text{tr}(\exp(-t\mathcal{L})) \quad (18)$$

Algorithm 1 XNYS TRACE- EXP: randomized average return probability estimate

Require: Laplacian matrix $\{L, \mathbb{L}_\mu(f), \mathbb{L}(f)\} \ni \mathfrak{L} \in \mathbb{R}^{n \times n}$, number M of matvecs, times $t \in [0, t^*]$.

Ensure: Estimate $\hat{p}(t) \approx \frac{1}{n} \text{tr}(\exp(-t\mathfrak{L}))$, and error estimate $\hat{e}(t) \approx \left| \hat{p}(t) - \frac{1}{n} \text{tr}(\exp(-t\mathfrak{L})) \right|$ for all $t \in [0, t^*]$.

1: Draw iid isotropic $\omega_1, \dots, \omega_M \in \mathbb{R}^n$

2: $\Omega \leftarrow [\omega_1 \dots \omega_M]$

3: Approximate the spectral radius $\rho(\mathfrak{L})$ of $\mathfrak{L}$, and build the poles $\{\xi_j\}_{j=1}^{k+1}$ of rational approximation of $\exp(-x)$ on the interval $[0, t^* \rho(\mathfrak{L})]$. ▷ Use AAA [41]

4: Build the basis V_{k+1} of $\mathcal{Q}_{k+1}^{\text{block}}(L, \Omega, q_k)$ and the block upper Hessenberg matrix pencil $(\mathcal{H}_k, \mathcal{K}_k)$. ▷ matvecs/solve

5: **for** $t \in [0, t^*]$ **do** ▷ Can be a **parfor**

6: $Y(t) \leftarrow \frac{1}{n} V_k \exp\left(-t \hat{\mathcal{H}}_k \hat{\mathcal{K}}_k^{-1}\right) V_k^\top \Omega$ ▷ Use the *reduced* pencil

7: $\hat{p}(t), \hat{e}(t) \leftarrow \text{XNYS TRACE}(Y(t), \Omega)$

8: **end for**

This quantity allows us to make a global statement on the system's effectiveness in diffusing the information, thus assessing its overall efficiency. In our experiments we will consider $\mathfrak{L} \in \{L, \mathcal{L}, \mathbb{L}_\mu(f)\}$; see Sect. 5 below.

The computation of the average return probabilities in (18) requires the evaluation of the trace of a matrix exponential for different values of t , nominally, until the quantity has changed from the initial value 1 to the final value $1/n$. In general, computing the trace of a matrix function exactly costs $O(n^3)$ computation and $O(n^2)$ storage. We have to consider the entire matrix $\exp(-t\mathfrak{L})$ in order to compute its trace, even if only temporarily. We note, however, that we can be satisfied with an approximation to $\hat{p}_0(t)$, since our goal is purely descriptive and we only need it to be able to compare the different network exploration strategies described in Sect. 3. Two convenient sets of tools for this task are represented by quadrature-based estimators [20] and stochastic trace estimators [38, Section 4].

Here, we focus on the latter, and adapt XNYS TRACE from [25] to our setting, as outlined in Alg. 1. Since we need to compute the matrix exponential for a range of parameters $t \in [0, t^*]$, two distinct regions of approximation arise; one dominated by the larger eigenvalues and another where accuracy on the smaller eigenvalues is essential. For the first region, a polynomial Krylov method would suffice; however, its accuracy diminishes when focusing on the smaller eigenvalues. To address this, we use a Block Rational Krylov method. Recall that this method defines a linear space of block vectors of size $n \times M$, constructed with respect to the matrix $L \in \mathbb{R}^{n \times n}$ and a starting block vector of maximal rank:

$$\Omega = [\omega_1, \omega_2, \dots, \omega_M] \in \mathbb{R}^{n \times M}, \quad \text{rank}(\Omega) = M.$$

To define the block rational Krylov space, we require a sequence of poles $\xi_1, \xi_2, \dots, \xi_k$, which are complex numbers or infinity, all distinct from the eigenvalues of $\mathfrak{L}$. The block rational Krylov space of order $k + 1$, associated with $\mathfrak{L}, \Omega, \xi_j$, is defined as

$$\mathcal{Q}_{k+1}^{\text{block}}(\mathfrak{L}, \Omega, r_k) = r_k(\mathfrak{L})^{-1} \left\{ \sum_{i=0}^k (\mathfrak{L})^i \Omega C_i \right\}, \quad (19)$$

where $\{C_i\}_{i=0}^k$ are matrices of size $M \times M$, and

$$r_k(z) = \prod_{\substack{j=1 \\ \xi_j \neq \infty}}^k (z - \xi_j)$$

is the common denominator of the rational matrix-valued functions associated with the block rational Krylov space. The construction of the basis vectors requires the solution of linear systems with matrix $\mu_k \mathcal{L} - \nu_k I$, and μ_k and $\nu_k \in \mathbb{C}$ obtained from the poles, in the case of Laplacians of the type (5) and (12). However, we only have access to routines that allow us to perform the matrix–vector product (Sect. 4.1). This means that we cannot use a direct solution, but we can rely again on the use of a Krylov-type method for the solution of linear systems fixed to achieve a lower tolerance than that which we require for the entire approximation, e.g., we can use the GMRES method. To generate an orthogonal block basis for this space, we then employ the block rational Arnoldi method [24], which produces

$$V_{k+1} = [\mathbf{v}_1, \dots, \mathbf{v}_{k+1}] \in \mathbb{C}^{n \times (k+1)M}, \quad \mathbf{v}_j \in \mathbb{C}^{n \times M},$$

satisfying the block rational Arnoldi decomposition:

$$\mathcal{L}V_{k+1}\mathcal{K}_k = V_{k+1}\mathcal{H}_k,$$

where $(\mathcal{H}_k, \mathcal{K}_k)$ is an unreduced block upper Hessenberg matrix pencil of size $(k+1)M \times kM$. The block vectors in V_{k+1} span the space $\mathcal{Q}_{k+1}^{\text{block}}(\mathcal{L}, \Omega, r_k)$.

To estimate the average return probabilities, we first sample M independent identically distributed isotropic vectors $\{\omega_i\}_{i=1}^M$, and then construct the Krylov subspaces $\mathcal{Q}_{k+1}^{\text{block}}(\mathcal{L}, \Omega, r_k)$. This involves computing the basis V_{k+1} and the unreduced block upper Hessenberg matrix pencil $(\mathcal{H}_k, \mathcal{K}_k)$. For each $t \in [0, t^*]$, we then compute

$$Y(t) = \frac{1}{n} V_k \exp\left(-t \hat{\mathcal{H}}_k \hat{\mathcal{K}}_k^{-1}\right) V_k^\top \Omega,$$

where the matrix-exponential is now efficiently computed on the reduced pencil—excluding the last block row of size M —($\hat{\mathcal{H}}_k, \hat{\mathcal{K}}_k$). To complete the description, we need to determine the poles $\{\xi_j\}_{j=1}^k$ for the construction of (19). Since we aim for an approximation of the matrix exponential that is accurate for all values of $t \in [0, t^*]$, we select the poles using the AAA algorithm for rational approximation [41]. The support points for the approximation are sampled from the interval $[0, t^* \rho(\mathcal{L})]$.

Finally, we run the trace estimator XNYS TRACE on the matrix $Y(t)$ and the block of vectors Ω for all $t \in [0, t^*]$. This step can be parallelized in shared memory, as the estimates are independent and only require sharing the Krylov subspace basis and the sampling vectors Ω .

5 Numerical Experiments

In this section we analyze the walk-based diffusion proposals introduced in Sects. 3 and 3.2.1 both from the modelling and computational point of view. Specifically, in Sects. 5.1 and 5.2 we investigate the network-exploration properties of the different strategies, while in Sect. 5.3 we analyze the computational challenges posed by the models.

The tests are run on the graphs reported in Table 1 from the SuiteSparse Matrix Collection [17]. The code to reproduce the experiments is available on the GitHub repository [Cirdans-Home/walk-laplacian](https://github.com/Cirdans-Home/walk-laplacian).

The numerical experiments were conducted on the Toeplitz cluster at the *Green Datacenter* of the University of Pisa. We employ the gpu nodes that are each equipped with a total of 256 CPUs, specifically AMD EPYC 7763 64-Core Processors 3.5GHz, with each processor providing 2 threads per core and 64 cores per socket, across 2 sockets. For the modeling and descriptive experiments in Sects. 5.1 and 5.2 we employed MATLAB v. 9.10.0.1602886 (R2021a). The experiments in Sect. 5.3 exploit the PSCToolkit library, and the available

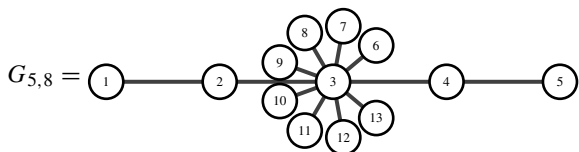
Table 1 Summary of networks in our dataset. For each we report the size ($n = |V|$) and the spectral radius of the matrix Z for $\mu = 1$; see Proposition 4

	Network name	n	$\rho(Z)$
1	Arenas/PGPgiantcompo	10680	41.03
2	Pajek/USpowerGrid	4941	6.23
3	DIMACS10/belgium_osm	1441295	2.58
4	SNAP/ca-CondMat	21363	35.86
5	SNAP/ca-HepPh	11204	243.75
6	Pajek/dictionary28	24831	16.07
7	Belcastro/human_gene1	21853	3335.86
8	Belcastro/human_gene2	14020	2999.83
9	SNAP/roadNet-CA	1957027	3.32
10	SNAP/roadNet-PA	1087562	3.11
11	SNAP/roadNet-TX	1351137	3.56
12	Gleich/usroads-48	126146	2.76

NVIDIA A40 GPUs (memory capacity of 4.6TB and a maximum power usage of 300W). For the latter, the NVIDIA driver v. 535.274.02 and CUDA v. 2.2 are used to manage these GPUs. The system software stack includes GCC v. 12.2.0, OpenBLAS v. 0.3.26, PSBLAS v. 3.9.0, AMG4PSBLAS v. 1.2.0.

5.1 Descriptive Examples

To describe the effect of the different formulations of the Laplacian $\mathbb{L}_\mu(f)$ for $\mu \in [0, 1]$ and $f \in \mathcal{P}$, we consider the toy network $G_{5,8}$. Networks $G_{l,m}$, for $l, m = 1, 2, \dots$ consist of a path with $l = 2k + 1$ vertices and of m additional leaves all connected to the central node in the path. With this notation, the toy network considered is:



The cluster of nodes in the centre acts as a “trap” for a walker. An interesting dynamic would be one for which the steady state probability of having arrived at the end of the path is significant compared with that of having arrived at any other node. We report our results in Fig. 1. We observe how the standard Laplacian L concentrates most of the mass of its steady state at node 3, the center of the trap. The same behavior is observed for all the other variants, except for the NBT k -path transformed Laplacian $\mathbb{L}_1(\text{exp})$. This fits well with the model interpretation. Indeed, if a walker moves from node 3, it is not possible to revisit it as backtracking is eliminated.

As a second example, consider an unlabeled tree generated uniformly at random using Wilf’s algorithm [53]. To build it we use the implementation of the algorithm available in the Python NetworkX (v.3.3) library [33] via a MATLAB interface. We construct a tree with $n = 100$ nodes and consider the probability matrices associated with the standard Laplacian L , the exponential k -walk transformed Laplacian $\mathbb{L}(\text{exp})$, the k -path transformed Laplacian $\mathcal{L}$ with $t_k = \exp(-\beta k)$ in (2), and exponential NBT k -walk transformed Laplacian $\mathbb{L}_1(\text{exp})$.

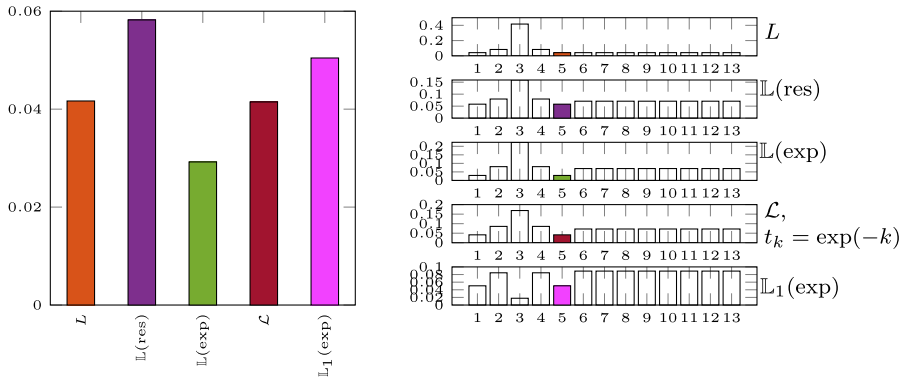


Fig. 1 Stationary distribution for the Markov chain induced by the different Laplacian. On the left panel, we have the probability mass at the target node at the end of the path, on the right one we report the different distributions. $\mathbb{L}(\text{res})$ is used to denote $\mathbb{L}((1 - \alpha x)^{-1})$ for $\alpha = 1/2\rho(A)$

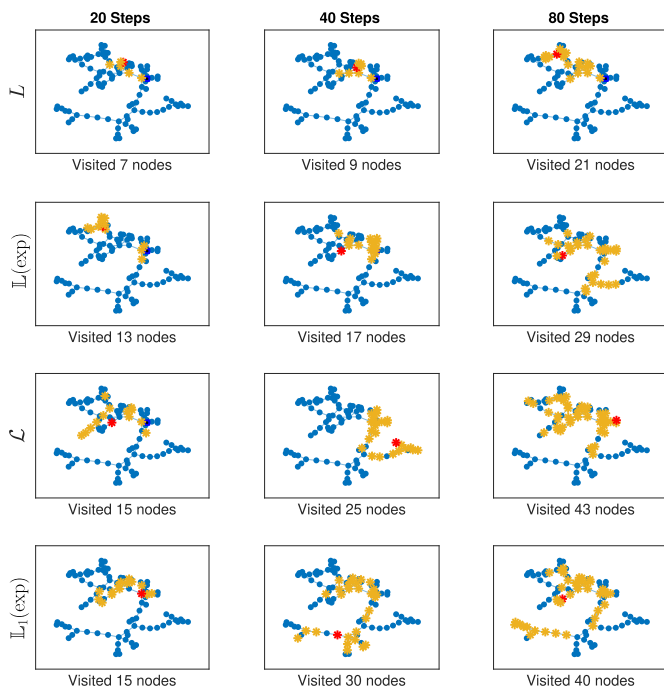
With the associated probability matrices we simulate the connected Markov chain starting from the root node; see Sect. 4.1.1. In Fig. 3a we report the evolution of the chain after 20, 40, and 80 steps showing which nodes of the tree are explored by the different strategies. In all cases other than the classical Laplacian L , the results demonstrate the non-locality of the discrete diffusion process. The k -path transformed Laplacian ($\mathcal{L}$) with exponential weights ($t_k = \exp(-k)$) and the exponential NBT k -walk transformed based Laplacian ($\mathbb{L}_1(\text{exp})$) are the ones that explore more nodes in the tree. To investigate the asymptotic behavior of the different chains we can also look at the gap between the eigenvalue of maximum modulus (1) and the one immediately below it, i.e., the *spectral gap*. From Fig. 3b we observe that the two Laplacians that explore the tree more widely have a larger gap compared to others (Fig. 2).

5.2 Average Return Probabilities

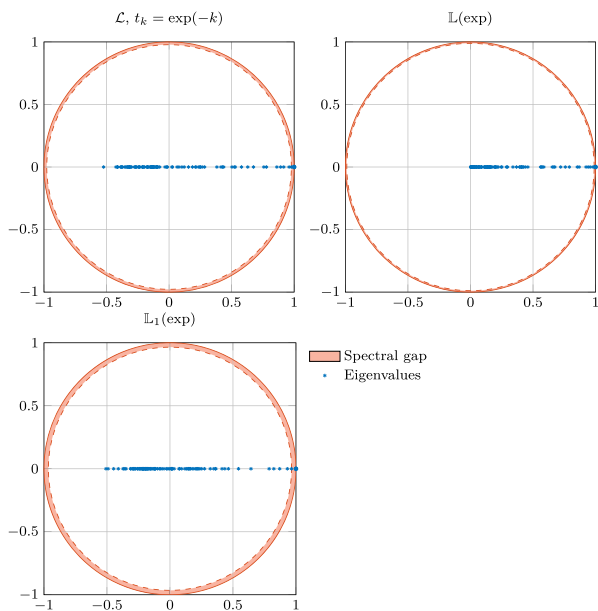
In this section, we perform two types of experiments. For the smaller networks in Table 1, we compute the average return probabilities (exactly) in order to investigate the behavior of different diffusion models. We then evaluate the performance of the randomized procedure in Alg. 1.

In Fig. 3 the graphs are separated based on the network considered. We plot the average return probability (18) versus time on a logarithmic scale, comparing six different Laplacian formulations: the classical Laplacian L , the NBT variant based on Katz centrality $\mathbb{L}_1((1 - \alpha x)^{-1})$, the exponential NBT k -walk transformed Laplacian $\mathbb{L}_1(\text{exp})$, the Katz centrality k -walk Laplacian $\mathbb{L}((1 - \alpha x)^{-1})$, the exponential k -walk transformed Laplacian $\mathbb{L}(\text{exp})$, and two k -path transformed Laplacians $\mathcal{L}$ with different weight functions: $t_k = 1/k$ and $t_k = \exp(-k)$. The average return probabilities are computed exactly, i.e., through the eigenvalue decomposition of the relevant Laplacian matrix.

The behavior observed is consistent across all considered networks. Notably, for exponential k -walk transformed Laplacians, the NBT variant $\mathbb{L}_1(\text{exp})$ exhibits a return probability that decays more slowly than $\mathbb{L}(\text{exp})$, confirming that removing backtracking walks decelerates the exploration process. In all cases except the one shown in Fig. 3b for the network SNAP/ca-CondMat, the fastest convergence is achieved by the $\mathcal{L}$ transformed with weights



(a) Exploration history of the different Markov chains on a tree graph with $n = 100$ nodes. We used orange to color the starting node and yellow to color the nodes visited in the exploration.



(b) Depiction of the spectral gap and eigenvalues for the probability matrices induced by the k -path exponentially transformed Laplacian L , the exponential k -walk transformed Laplacian $L(\exp)$ and its NBT counterpart $L_1(\exp)$.

Fig. 2 Example of discrete-time diffusion on an unlabeled tree graph for Markov chains obtained through different formulations of the Laplacian

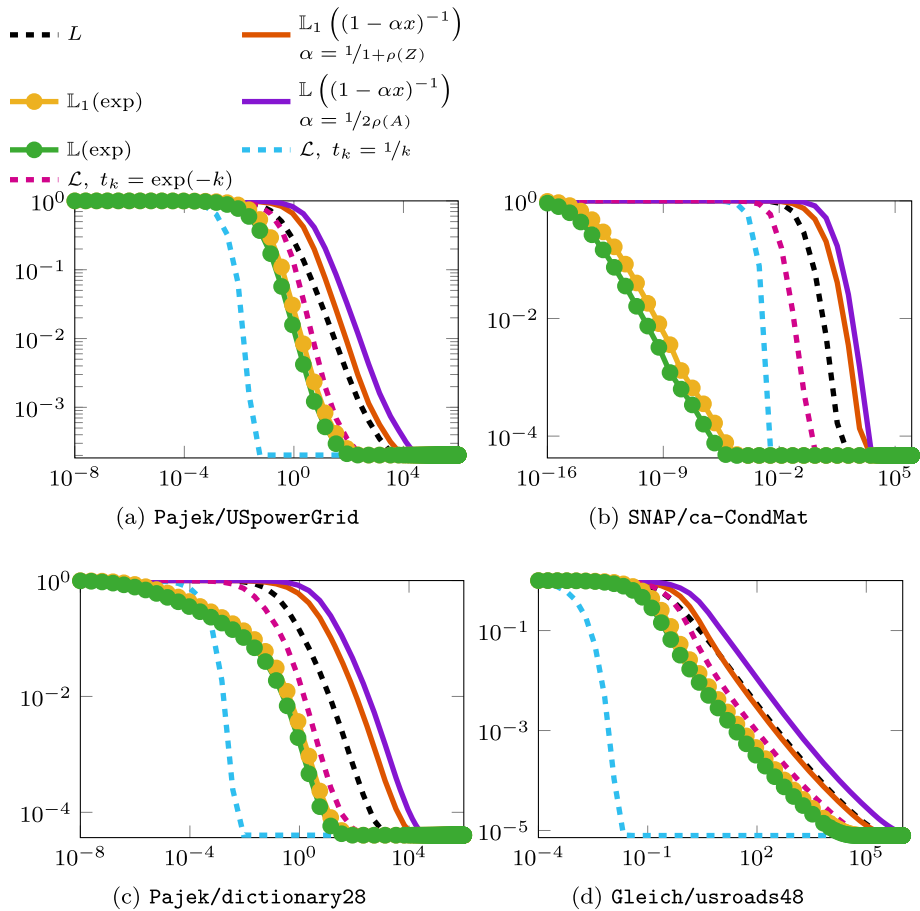


Fig. 3 Average return probabilities (18) computed for matrices from Table 1 for which the computation of the path-based Laplacian was feasible for the dimension. Each subfigure displays results across six different Laplacian formulations on logarithmic time scales

$t_k = 1/k$, corresponding to $\beta = 1$ in (3). In all cases, $\mathcal{L}$ transformed with $t_k = \exp(-k)$ has slower convergence than $\mathbb{L}_\mu(\exp)$, with $\mu = 0, 1$.

Another noteworthy finding is the behavior of the the k -walk transformed Laplacians $\mathbb{L}_\mu((1 - \alpha x)^{-1})$ for $\mu = 0, 1$. Both these result in a slower transition to the steady state compared to the classical Laplacian L .

In Fig. 4 we report the average return probability for the BTDW k -walk Laplacian $\mathbb{L}_\mu(\exp)$ as μ varies in $[0, 1]$. The plot highlights the interpolation effect between the case $\mu = 0$, where the construction coincides with the k -walk transformed Laplacian $\mathbb{L}(\exp)$, and the case $\mu = 1$, which results in the NBT variant $\mathbb{L}_1(\exp)$.

We observe that suppressing backtracking walks slows down the overall exploration of the network, resulting in a higher average return probability compared to the case where all walks are counted, as in $\mathbb{L}(\exp)$. In other words, penalizing or removing immediate reversals limits the walker's ability to escape its local neighborhood, thereby reducing the effective diffusion rate across the graph. This highlights how the parameter μ controls the

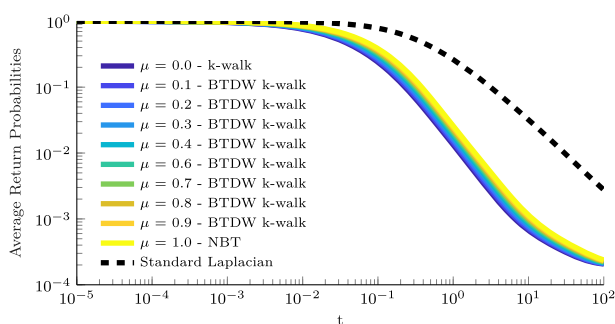


Fig. 4 Average return probabilities (18) computed for the Gleich/usroads48 and the $\mathbb{L}_\mu(\text{exp})$ Laplacian for $0 \leq \mu \leq 1$ showing the effect of down-weighting more and more the NBT walks. The dashed line represents the average return probability of standard Laplacian L

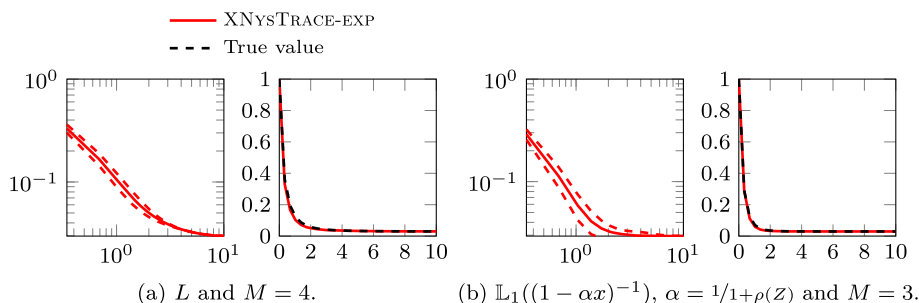


Fig. 5 Randomized approximation of the average return probabilities (18) with the XNYSTRACE-EXP algorithm (Alg. 1) on network Newman/karate and M sample vectors. The dashed lines in the left panels of Fig. 5a and 5b depicts the error estimate obtained from Alg. 1. In the right panels we overlap the estimate obtained from the XNYSTRACE-EXP in Alg. 1 and the exact value computed from the eigenvalue decomposition

balance between local confinement and global exploration; smaller values promote faster spreading by allowing repeated edge traversals, while larger values enforce a more deliberate progression through the network.

5.2.1 Randomized Average Return Probability Estimates

In this section we test the idea of approximating the average return probability (18) via the stochastic estimator XNYSTRACE-EXP, as described in Sect. 4.4. For the construction of the block rational Krylov space in Alg. 1 we use the RKTtoolbox [11].

In Fig. 5a and 5b we report the results obtained for the estimation of the average probability of return using the algorithm on a small-sized network, Newman/karate with $n = 32$ nodes, for which we can compute it exactly and compare the results. Due to the small dimension, we use few sampling vectors in the two cases, $M = 4$ and $M = 3$ respectively. As can be observed, the agreement between the stochastic estimate and the correct value is good enough and in any case adequate to describe the behavior of the two diffusion models.

As a second example, consider two complex networks from Table 1, for which exact computation of the average return probability is infeasible due to memory limitations (SNAP/roadNet-PA and SNAP/roadNet-CA, with $n = 1087562$ and $n = 1957027$ nodes, respectively). We select $M = 30$ sampling vectors and apply Algorithm 1 over the

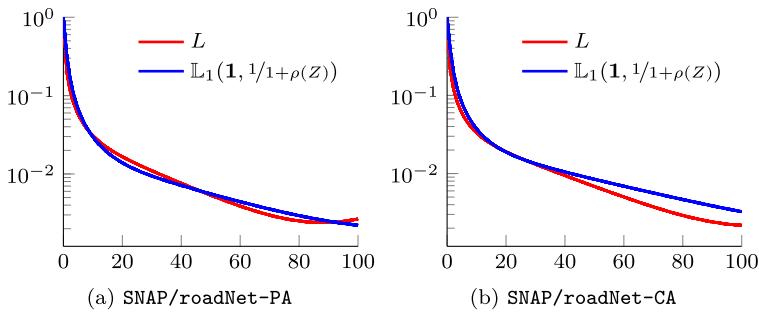


Fig. 6 Randomized approximation of the average return probabilities (18) with the XNYSTRACE-EXP algorithm (Alg. 1). Each experiment is run 10 times, with curves plotted using the same color and stroke thickness per repetition

interval $[0, t^*] = [0, 100]$ with $n_t = 90$ equally spaced nodes. The estimation is repeated ten times, with each repetition plotted in Fig. 6 using consistent stroke and color across curves.

As in the previous case, we focus on the standard Laplacian L as a reference, along with the resolvent NBT k -walk transformed Laplacian $\mathbb{L}_1((1 - \alpha x)^{-1})$. For this, we set $\alpha = 1/(1 + \rho(Z))$. The GMRES algorithm used in the rational Krylov subspace method is configured to achieve a tolerance of 10^{-6} .

The results in Figs. 6a and 6b for networks SNAP/roadNet-PA and SNAP/roadNet-CA can be compared to those in Fig. 3d for the Gleich/usroads48 network. Both networks represent road systems, suggesting a similar reciprocal behavior between the standard and NBT Laplacians, which is observed within the error margin due to randomization. For both networks, the AAA algorithm yields a rational approximation with 13 poles, and the internal GMRES method converges to the required tolerance with iteration counts between 6 and 12, depending on the pole.

5.3 Matrix–Vector Products with Resolvent k -Walk Transformed Laplacian

Here, we conduct tests on the solution of linear systems involving the matrix $\mathcal{A}(\alpha)$ across several examples of real-world networks. For each network, we restrict our analysis to its largest connected component. The parameter α is selected as four equally spaced, logarithmically scaled values between 10^{-3} and $10^{-1/2}$, scaled according to the dominant eigenvalue of the matrix Z in Proposition 4. The Conjugate Gradient algorithm, preconditioned with the AMG method from Sect. 4.2, is employed with a relative tolerance of $\varepsilon = 10^{-9}$. To minimize the construction time of the multigrid hierarchy, OpenMP parallelization is used with 64 threads. Results are displayed in Fig. 7, where we display the time required to apply a matrix–vector product with $\mathbb{L}_\mu((1 - \alpha x)^{-1})$, with α selected as described above.

For each of the twelve networks in Table 1, both CPU and GPU execution times vary only marginally as α changes. This indicates that the matrix structure induced by the resolvent operator does not significantly affect algorithmic stability or solver convergence. In practice, solving the resolvent-regularized system is nearly as efficient as performing a small number of standard sparse matrix–vector multiplications with the original graph Laplacian L . This demonstrates the effectiveness of the AMG preconditioning strategy [22] and the scalability of the approach across a wide range of graph sizes and topologies. The GPU implementation achieves speedups exceeding one order of magnitude for several cases, particularly

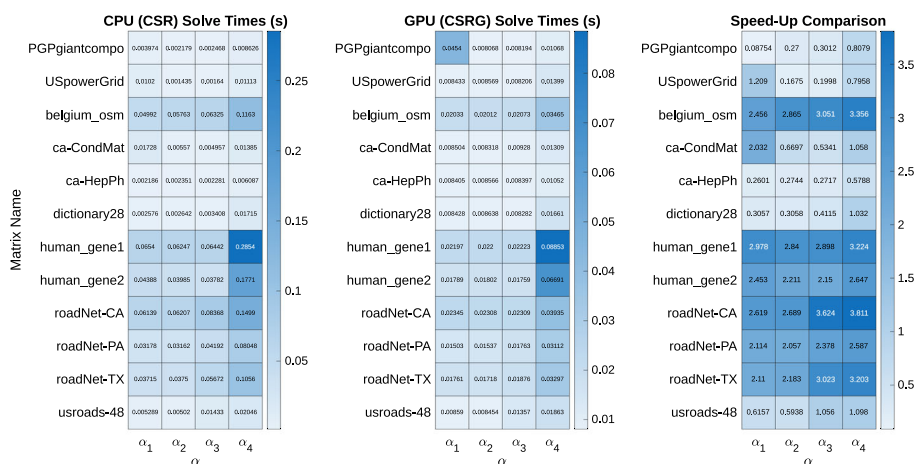


Fig. 7 The two heatmaps show the time required to apply a matrix–vector product with $\mathbb{L}_\mu((1-\alpha)x)^{-1}$. Four equally spaced values of α logarithmically scaled between 10^{-3} and $10^{-1/2}$, based on the dominant eigenvalue of matrix Z (see Proposition 4). Rows correspond to the twelve complex networks detailed in Table 1

on larger and denser networks such as roadNet-CA, roadNet-PA, human_gene2, and usroads-48. Even in instances where acceleration is less dramatic, GPU solve times remain comparable to or are faster than CPU computation. This confirms that reformulating the sparse operator into the CSR format allows efficient utilization of parallel devices. Networks with higher edge counts benefit most, as the available parallelism enables large batches of row operations to proceed concurrently.

Overall, the experiments demonstrate that applying the resolvent-based k -walk transformed Laplacian is computationally practical at scale. The combination of AMG preconditioning and CSR-based GPU kernels yields substantial performance gains without sacrificing numerical robustness, making the method well suited for large network analysis tasks where repeated linear solves are required.

6 Conclusions and Outlook

In this paper, we introduced a novel framework for constructing diffusion operators on complex networks, extending the traditional Laplacian operator through the use of graph-based walks. This approach enables a flexible and nuanced interpretation of network centrality; while the classical Laplacian operator aligns with degree centrality, our generalized construction leverages a generic total communicability centrality, facilitating the use of diverse matrix function-based centrality measures to shape the diffusion dynamics. In our framework we have treated also the inclusion of backtrack-downweighted and nonbacktracking walks, which incorporate memory into the diffusion process thus enhancing its descriptive power.

We proposed and analyzed efficient computational strategies for these operators, along with methods to compute key quantities such as the average return probability of the diffusion process. Demonstrating applicability on real-world networks, we showed how our approach differs from existing path-based diffusion operators, highlighting the distinctive properties and flexibility of this new class of operators. Our framework offers a valuable tool for adapting diffusion dynamics across various network contexts. Future work will focus on new directions

in network analysis and modeling, such as the possibility to consider and generalize the discrete version of differential equations beyond diffusion (e.g., wave equation or Schrödinger equation). From a computational viewpoint, we also plan to integrate the tools contained in the `PSCToolkit` library to be able to generate and treat these problems in a systematic way for large networks on distributed machines equipped with hardware accelerators such as GPUs.

Acknowledgements The authors would like to thank the anonymous reviewer for their constructive feedback, which helped improve this paper.

Funding This work has been partially supported by: Spoke 1 “FutureHPC & BigData” of the Italian Research Center on High- Performance Computing, Big Data and Quantum Computing (ICSC) funded by MUR Missione 4 Componente 2 Investimento 1.4: Potenziamento strutture di ricerca e creazione di “campioni nazionali di R&S (M4C2-19)” - Next Generation EU (NGEU). The visit of F.D. at the University of Strathclyde where some of this research was carried out was funded by “Istituto Nazionale di Alta Matematica (INdAM) - Gruppo Nazionale per il Calcolo Scientifico (GNCS)” project “PASTRAMI - sPline And Solver innovaTions foR Adaptive isogeometric analysis”. Project code CUP E53C24001950001.”

Data Availability All matrices used in this work are drawn from the `SuiteSparse Matrix Collection` [17] (sparse.tamu.edu). The code required to reproduce all experiments and figures is publicly available at github.com/Cirdans-Home/walk-laplacian.

Declarations

Conflict of interest The authors have not disclosed any conflict of interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Al-Mohy, A.H., Higham, N.J.: Computing the action of the matrix exponential, with an application to exponential integrators. *SIAM J. Sci. Comput.* **33**(2), 488–511 (2011). <https://doi.org/10.1137/100788860>
2. Arrigo, F., Durastante, F.: Mittag-Leffler functions and their applications in network science. *SIAM J. Matrix Anal. Appl.* **42**(4), 1581–1601 (2021). <https://doi.org/10.1137/21M1407276>
3. Arrigo, F., Grindrod, P., Higham, D.J., Noferini, V.: Non-backtracking walk centrality for directed networks. *J. Complex Netw.* **6**(1), 54–78 (2018). <https://doi.org/10.1093/comnet/cnx025>
4. Arrigo, F., Grindrod, P., Higham, D.J., Noferini, V.: On the exponential generating function for non-backtracking walks. *Linear Algebra Appl.* **556**, 381–399 (2018). <https://doi.org/10.1016/j.laa.2018.07.010>
5. Arrigo, F., Higham, D.J., Noferini, V.: A theory for backtrack-downweighted walks. *SIAM J. Matrix Anal. Appl.* **42**(3), 1229–1247 (2021). <https://doi.org/10.1137/20M1384725>
6. Baker, A.H., Falgout, R.D., Kolev, T.V., Yang, U.M.: Multigrid smoothers for ultraparallel computing. *SIAM J. Sci. Comput.* **33**(5), 2864–2887 (2011). <https://doi.org/10.1137/100798806>
7. Benzi, M., Bertaccini, D., Durastante, F., Simunec, I.: Non-local network dynamics via fractional graph Laplacians. *J. Complex Netw.* **8**(3), cnaa017, 29 (2020). <https://doi.org/10.1093/comnet/cnaa017>
8. Benzi, M., Klymko, C.: Total communicability as a centrality measure. *Journal of Complex Networks* **1**(2), 124–149 (2013). <https://doi.org/10.1093/comnet/cnt007>
9. Benzi, M., Klymko, C.: On the limiting behavior of parameter-dependent network centrality measures. *SIAM J. Matrix Anal. Appl.* **36**(2), 686–706 (2015). <https://doi.org/10.1137/130950550>

10. Benzi, M., Simunek, I.: Rational Krylov methods for fractional diffusion problems on graphs. *BIT* **62**(2), 357–385 (2022). <https://doi.org/10.1007/s10543-021-00881-0>
11. Berljafa, M., Elsworth, S., Güttel, S.: A rational Krylov toolbox for MATLAB. Tech. rep., Manchester Institute for Mathematical Sciences, University of Manchester (2014). <http://guettel.com/rktoolbox/index.html>
12. Berman, A., Plemmons, R.J.: Nonnegative matrices in the mathematical sciences, *Classics in Applied Mathematics*, vol. 9. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (1994). <https://doi.org/10.1137/1.9781611971262>. Revised reprint of the 1979 original
13. Bharali, G., Holtz, O.: Functions preserving nonnegativity of matrices. *SIAM J. Matrix Anal. Appl.* **30**(1), 84–101 (2008). <https://doi.org/10.1137/050645075>
14. Bowen, R., Lanford III, O.E.: Zeta functions of restrictions of the shift transformation. In: Global Analysis (Proc. Sympos. Pure Math., Vols. XIV, XV, XVI, Berkeley, Calif., 1968), *Proc. Sympos. Pure Math.*, vol. XIV-XVI, pp. 43–49. Amer. Math. Soc., Providence, RI (1970)
15. Ciarlet, P.G.: Discrete maximum principle for finite-difference operators. *Aequationes Math.* **4**, 338–352 (1970). <https://doi.org/10.1007/BF01844166>
16. Cipolla, S., Durastante, F., Tudisco, F.: Nonlocal PageRank. *ESAIM Math. Model. Numer. Anal.* **55**(1), 77–97 (2021). <https://doi.org/10.1051/m2an/2020071>
17. Davis, T.A., Hu, Y.: The University of Florida Sparse Matrix Collection. *ACM Trans. Math. Softw.* **38**(1) (2011). <https://doi.org/10.1145/2049662.2049663>
18. Deka, P.J., Einkemmer, L., Tokman, M.: LeXInt: Package for exponential integrators employing Leja interpolation. *SoftwareX* **21**, 101302 (2023). <https://doi.org/10.1016/j.softx.2022.101302>
19. Diaz-Diaz, F., Estrada, E.: Time and space generalized diffusion equation on graph/networks. *Chaos Solitons Fractals* **156**, Paper No. 111791, 9 (2022). <https://doi.org/10.1016/j.chaos.2022.111791>
20. Díaz Fuentes, R., Donatelli, M., Fenu, C., Mantica, G.: Estimating the trace of matrix functions with application to complex networks. *Numer. Algorithms* **92**(1), 503–522 (2023). <https://doi.org/10.1007/s11075-022-01417-5>
21. Diestel, R.: Graph theory, *Graduate Texts in Mathematics*, vol. 173, third edn. Springer-Verlag, Berlin, Berlin (2005)
22. D’Ambra, P., Durastante, F., Filippone, S.: Parallel Sparse Computation Toolkit. *Software Impacts* **15**, 100463 (2023). <https://doi.org/10.1016/j.simpa.2022.100463>
23. Elhashash, A., Szyld, D.B.: Matrix functions preserving sets of generalized nonnegative matrices. *Electron. J. Linear Algebra* **20**, 673–690 (2010). <https://doi.org/10.13001/1081-3810.1401>
24. Elsworth, S., Güttel, S.: The block rational Arnoldi method. *SIAM J. Matrix Anal. Appl.* **41**(2), 365–388 (2020). <https://doi.org/10.1137/19M1245505>
25. Epperly, E.N., Tropp, J.A., Webber, R.J.: XTrace: Making the Most of Every Sample in Stochastic Trace Estimation. *SIAM J. Matrix Anal. Appl.* **45**(1), 1–23 (2024). <https://doi.org/10.1137/23M1548323>
26. Estrada, E.: Path Laplacian matrices: introduction and application to the analysis of consensus in networks. *Linear Algebra Appl.* **436**(9), 3373–3391 (2012). <https://doi.org/10.1016/j.laa.2011.11.032>
27. Estrada, E., Delvenne, J.C., Hatano, N., Mateos, J.L., Metzler, R., Riascos, A.P., Schaub, M.T.: Random multi-hopper model: super-fast random walks on graphs. *Journal of Complex Networks* **6**(3), 382–403 (2017). <https://doi.org/10.1093/comnet/cnx043>
28. Estrada, E., Hameed, E., Langer, M., Puchalska, A.: Path Laplacian operators and superdiffusive processes on graphs. II. Two-dimensional lattice. *Linear Algebra Appl.* **555**, 373–397 (2018). <https://doi.org/10.1016/j.laa.2018.06.026>
29. Estrada, E., Higham, D.J.: Network properties revealed through matrix functions. *SIAM Rev.* **52**(4), 696–714 (2010). <https://doi.org/10.1137/090761070>
30. Floyd, R.W.: Algorithm 97: Shortest Path. *Commun. ACM* **5**(6), 345 (1962). <https://doi.org/10.1145/367766.368168>
31. Grindrod, P., Higham, D.J., Noferini, V.: The deformed graph Laplacian and its applications to network centrality analysis. *SIAM J. Matrix Anal. Appl.* **39**(1), 310–341 (2018). <https://doi.org/10.1137/17M1112297>
32. Güttel, S.: Rational Krylov approximation of matrix functions: numerical methods and optimal pole selection. *GAMM-Mitt.* **36**(1), 8–31 (2013). <https://doi.org/10.1002/gamm.201310002>
33. Hagberg, A.A., Schult, D.A., Swart, P.J.: Exploring Network Structure, Dynamics, and Function using NetworkX. In: G. Varoquaux, T. Vaught, J. Millman (eds.) *Proceedings of the 7th Python in Science Conference*, pp. 11–15. Pasadena, CA USA (2008). http://conference.scipy.org/proceedings/SciPy2008/paper_2/
34. Higham, N.J.: Functions of matrices. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2008). <https://doi.org/10.1137/1.9780898717778>. Theory and computation

35. Hochbruck, M., Ostermann, A.: Exponential integrators. *Acta Numer* **19**, 209–286 (2010). <https://doi.org/10.1017/S0962492910000048>
36. Horn, R.A., Johnson, C.R.: *Matrix analysis*, 2nd edn. Cambridge University Press, Cambridge (2013)
37. Johnson, D.B.: Efficient algorithms for shortest paths in sparse networks. *J. Assoc. Comput. Mach.* **24**(1), 1–13 (1977). <https://doi.org/10.1145/321992.321993>
38. Martinsson, P.G., Tropp, J.A.: Randomized numerical linear algebra: foundations and algorithms. *Acta Numer* **29**, 403–572 (2020). <https://doi.org/10.1017/s0962492920000021>
39. Masuda, N., Porter, M.A., Lambiotte, R.: Random walks and diffusion on networks. *Phys. Rep.* **716**(717), 1–58 (2017). <https://doi.org/10.1016/j.physrep.2017.07.007>
40. Morbidi, F.: The deformed consensus protocol. *Automatica J. IFAC* **49**(10), 3049–3055 (2013). <https://doi.org/10.1016/j.automatica.2013.07.006>
41. Nakatsukasa, Y., Sète, O., Trefethen, L.N.: The AAA algorithm for rational approximation. *SIAM J. Sci. Comput.* **40**(3), A1494–A1522 (2018). <https://doi.org/10.1137/16M1106122>
42. Niesen, J., Wright, W.M.: Algorithm 919: a Krylov subspace algorithm for evaluating the ϕ -functions appearing in exponential integrators. *ACM Trans. Math. Software* **38**(3), Art. 22, 19 (2012). <https://doi.org/10.1145/2168773.2168781>
43. Riascos, A.P., Mateos, J.L.: Long-range navigation on complex networks using Lévy random walks. *Phys. Rev. E* **86**, 056110 (2012). <https://doi.org/10.1103/PhysRevE.86.056110>
44. Riascos, A.P., Mateos, J.L.: Fractional dynamics on networks: Emergence of anomalous diffusion and Lévy flights. *Phys. Rev. E* **90**, 032809 (2014). <https://doi.org/10.1103/PhysRevE.90.032809>
45. Roger, H., Charles, R.J.: *Topics in matrix analysis*, (1991)
46. Roy, B.: Transitivité et connexité. *C. R. Acad. Sci. Paris* **249**, 216–218 (1959)
47. Saad, Y.: *Iterative methods for sparse linear systems*, second edn. Society for Industrial and Applied Mathematics, Philadelphia, PA (2003). <https://doi.org/10.1137/1.9780898718003>
48. Saade, A., Krzakala, F., Zdeborová, L.: Spectral Clustering of graphs with the Bethe Hessian. In: Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, K. Weinberger (eds.) *Advances in Neural Information Processing Systems*, vol. 27. Curran Associates, Inc., San Diego, CA (2014). <https://shorturl.at/dpwN3>
49. Skaflestad, B., Wright, W.M.: The scaling and modified squaring method for matrix functions related to the exponential. *Appl. Numer. Math.* **59**(3–4), 783–799 (2009). <https://doi.org/10.1016/j.apnum.2008.03.035>
50. Stark, H.M., Terras, A.A.: Zeta functions of finite graphs and coverings. *Adv. Math.* **121**(1), 124–165 (1996). <https://doi.org/10.1006/aima.1996.0050>
51. Vaněk, P., Mandel, J., Brezina, M.: Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing* **56**(3), 179–196 (1996). <https://doi.org/10.1007/BF02238511>. *International GAMM-Workshop on Multi-level Methods* (Meisdorf, 1994)
52. Warshall, S.: A Theorem on Boolean Matrices. *J. ACM* **9**(1), 11–12 (1962). <https://doi.org/10.1145/321105.321107>
53. Wilf, H.S.: The uniform selection of free trees. *J. Algorithms* **2**(2), 204–207 (1981). [https://doi.org/10.1016/0196-6774\(81\)90021-3](https://doi.org/10.1016/0196-6774(81)90021-3)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.