SPECTRAL STRUCTURE OF PATH-LENGTH DENSITY MATRICES OF UNROOTED BINARY TREES

Daniele Catanzaro, Raffaele Pesenti, Roberto Ronco

LIDAM Discussion Paper CORE
2026 / 13

UCLouvain

LIDAM
CORE | Louvain Institute of Data Analysis and Modeling in economics and statistics

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email: immaq-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/core-discussion-papers

SPECTRAL STRUCTURE OF PATH-LENGTH DENSITY MATRICES OF UNROOTED BINARY TREES*

DANIELE CATANZARO†, RAFFAELE PESENTI‡, AND ROBERTO RONCO§

Abstract. The *Path-Length Density Matrix* (PLDM) of an *Unrooted Binary Tree* (UBT) is the real symmetric matrix with zero diagonal and off-diagonal entries $2^{-\tau_{ij}}$, where τ_{ij} denotes the number of edges on the unique path between leaves i and j. Here, UBTs are trees with $n \geq 3$ leaves and all internal vertices of degree 3, and their PLDMs arise naturally in computational phylogenetics, network design, and information theory. In this article, we study PLDMs of UBTs and investigate how their spectra reflect the structure of the underlying trees. We prove that the smallest eigenvalue of any PLDM of a UBT is $-1/4$, with multiplicity equal to the number of leaf pairs sharing a common parent. We also show that specific eigenvalues characterize the presence of particular subtree configurations. For UBTs of minimum diameter, we derive closed-form expressions for the eigenvalues; for UBTs of maximum diameter, we relate the spectrum to an auxiliary Kac–Murdock–Szegő-type matrix and obtain asymptotic results for the empirical spectral distribution and the edge spectral densities. These results show that the spectrum of a PLDM encodes explicit structural information about the underlying tree and suggest that spectral methods may be useful for optimization over this matrix class.

Key words. spectrum, path-length density matrices, unrooted binary trees, structured symmetric matrices, computational phylogenetics.

MSC codes. 15A18, 05C05, 05C50, 90C27.

1. Introduction.

Terminal encodings are matrix representations of trees that depend only on information associated with a distinguished subset of vertices, typically called *terminals*. Such encodings arise naturally in inverse optimization problems in which one seeks to reconstruct or identify a tree structure from incomplete or aggregated data available only at the terminals. Classical examples include the *Huffman Coding Problem* [17] and the *Balanced Minimum Evolution Problem* (BMEP) [4].

A prominent example of terminal encoding is the *Path-Length Matrix* (PLM) [4]. Given a tree T with $n \geq 3$ leaves, the PLM of T is the symmetric integer matrix $\boldsymbol{\tau}$ of order n with zero diagonal and off-diagonal entries τ_{ij}, where τ_{ij} is the number of edges on the unique path in T connecting leaves i and j. Figure 1 shows an *Unrooted Binary Tree* (UBT), i.e., a tree with $n \geq 3$ leaves and all internal vertices of degree 3, together with its associated PLM. The central object of this article is the closely related *Path-Length Density Matrix* (PLDM). For a UBT T, the PLDM is the real symmetric $n \times n$ matrix with zero diagonal and off-diagonal entries $2^{-\tau_{ij}}$. Figure 1 also shows the PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ associated with the displayed UBT.

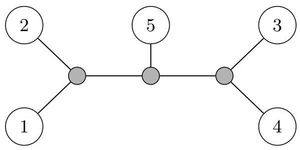
The extent to which matrix spectra encode structural information about trees is a classical question in spectral graph theory; see, e.g., [2, 5, 8, 10]. In general, spectra yield only partial information: unweighted trees are not uniquely determined by the spectrum of either their adjacency matrix [18] or their PLM [15], and similar non-uniqueness phenomena occur for rooted binary trees [14]. Nonetheless, spectral

*Submitted to the SIAM Journal on Matrix Analysis and Applications (SIMAX) on April 23, 2026.

†Center for Operations Research and Econometrics, Université Catholique de Louvain, Voie du Roman Pays 34, 1348, Louvain-la-Neuve, Belgium (daniele.catanzaro@uclouvain.be).

‡Venice School of Management, University Ca' Foscari, San Giobbe, Cannaregio 837, I-30121 Venice, Italy (pesenti@unive.it).

§Institute of Marine Engineering, Italian National Research Council, Via De Marini 6, I-16149, Genoa, Italy (roberto.ronco@cnr.it).



$$\boldsymbol{\tau} = \begin{pmatrix} 0 & 2 & 4 & 4 & 3 \\ 2 & 0 & 4 & 4 & 3 \\ 4 & 4 & 0 & 2 & 3 \\ 4 & 4 & 2 & 0 & 3 \\ 3 & 3 & 3 & 3 & 0 \end{pmatrix} \qquad \mathbf{2}^{-\boldsymbol{\tau}} = \begin{pmatrix} 0 & 2^{-2} & 2^{-4} & 2^{-4} & 2^{-3} \\ 2^{-2} & 0 & 2^{-4} & 2^{-4} & 2^{-3} \\ 2^{-4} & 2^{-4} & 0 & 2^{-2} & 2^{-3} \\ 2^{-4} & 2^{-4} & 2^{-2} & 0 & 2^{-3} \\ 2^{-3} & 2^{-3} & 2^{-3} & 2^{-3} & 0 \end{pmatrix}$$

FIG. 1. *On the left, an example of an Unrooted Binary Tree (UBT), i.e., a tree graph with $n \geq 3$ leaves, $n-2$ internal vertices each of degree 3, and $2n-3$ edges. The UBT shown here has 5 leaves, 3 internal vertices, and 7 edges. At the center and on the right, the Path-Length Matrix (PLM) $\boldsymbol{\tau}$ and the Path-Length Density Matrix (PLDM) $\mathbf{2}^{-\boldsymbol{\tau}}$ of this UBT, respectively.*

data often reveal meaningful structural features. For example, Schwenk [18] showed that cospectral trees must contain certain exchangeable rooted subtrees, indicating that spectral signatures can reflect underlying combinatorial components. In a related direction, Araújo Lima and de Aguiar [1] studied Laplacian spectra of rooted binary trees in extremal diameter regimes and observed clustering phenomena for highly balanced trees.

Our perspective is different. Motivated by optimization questions related to the BMEP and by the combinatorial structure of PLDMs of UBTs, we study the extent to which the spectrum of a PLDM reflects the topology of the underlying tree. This question is also relevant for optimization: in the BMEP, one seeks a UBT whose PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ minimizes $\mathrm{tr}(\mathbf{D}\,\mathbf{2}^{-\boldsymbol{\tau}})$ for a given distance matrix $\mathbf{D}$. Understanding which structural features of a UBT are encoded by the spectrum of its PLDM may therefore help clarify the structure of feasible solutions and motivate future algorithmic developments.

Main results and structure of the article. After introducing notation and preliminary results in Section 2, we study the eigenspaces of PLDMs of UBTs in Section 3. We prove that the smallest eigenvalue of any PLDM is $-1/4$, with multiplicity equal to the number of *cherries*, i.e., pairs of leaves at distance 2, in the underlying UBT. We also show that specific eigenvalues characterize the presence of particular subtree configurations. For UBTs attaining the maximum possible diameter or having a maximally balanced structure, we derive explicit spectral descriptions and show that both families can be analyzed through a common auxiliary Kac–Murdock–Szegő-type matrix. In Section 4, we establish sharper eigenvalue bounds and study asymptotic spectral behavior. Finally, in Section 5, we present conjectures supported by computational evidence and discuss possible extensions of the present work.

2. Notation and preliminaries.

Throughout the article, we shall use lowercase letters for scalars (e.g., $c \in \mathbb{R}$), bold capital letters for matrices (e.g., $\mathbf{A} \in \mathbb{R}^{n\times n}$), bold lowercase letters for vectors (e.g., $\mathbf{v} \in \mathbb{R}^n$), and Greek letters to denote the eigenvalues of a given matrix (e.g., the eigenvalue λ of a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$).

Intervals, vectors, and matrices. Given two integers a, b, with $a \leq b$, we denote by $[a, b]$ the set of integers c such that $a \leq c \leq b$. If $a = 1$, we simply write $[b]$.

Unless stated otherwise, any vector norm is understood to be Euclidean, and any matrix $\mathbf{A}$ is assumed to be real, symmetric, and of order n. For an integer $n \geq 1$, we denote by $\mathbf{0}_n$ and $\mathbf{1}_n$ the column vectors in $\mathbb{R}^n$ whose entries are all equal to 0 and 1, respectively. We denote by $\mathbf{e}_i$ the i-th canonical unit vector of $\mathbb{R}^n$, and by $\mathbf{e}_i^{(k)}$ the i-th canonical unit vector of the subspace $\mathbb{R}^k$, for $k \in [n-1]$. For example, if $n = 4$, then $\mathbf{e}_1^{(2)} = (1, 0)^\top$ and $\mathbf{e}_1 = (1, 0, 0, 0)^\top$. Given a vector $\mathbf{v} \in \mathbb{R}^n$, we write $x_{[k]}$ to denote $k \geq 2$ consecutive entries of $\mathbf{v}$ equal to x. For example, if $\mathbf{v} = (2, 2, 1, 1, 2)^\top$, then we write $\mathbf{v} = (2_{[2]}, 1_{[2]}, 2)^\top$.

We denote the determinant, trace, rank, kernel, and image of $\mathbf{A}$ by $\det(\mathbf{A})$, $\mathrm{tr}(\mathbf{A})$, $\mathrm{rank}(\mathbf{A})$, $\ker(\mathbf{A})$, and $\mathrm{im}(\mathbf{A})$, respectively. We refer to the i-th row and column vectors of $\mathbf{A}$ as $\mathbf{A}_{i\cdot}$ and $\mathbf{A}_{\cdot i}$, respectively, and we denote by $\mathrm{diag}(a_{11}, a_{22}, \ldots, a_{nn})$ the diagonal matrix whose diagonal entries are a_{11}, a_{22}, …, a_{nn}. Finally, we denote by $\mathbf{I}_n$ the identity matrix of order n and, for an integer $k \in [n]$, we denote by $\hat{\mathbf{I}}_{n-k} = \mathrm{diag}(1_{[k]}, 0_{[n-2k]})$ the diagonal matrix whose first k diagonal entries are equal to 1, and whose remaining $n-2k$ entries (if any) are equal to 0.

For a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$, we denote by $m_g(\lambda)$ the geometric multiplicity of an eigenvalue λ and simply refer to it as the *multiplicity* of λ, since geometric and algebraic multiplicities coincide for real symmetric matrices. We always assume that the eigenvalues of $\mathbf{A}$ are ordered non-decreasingly, so that $\lambda_i \leq \lambda_j$ whenever $i < j$.

Finally, we will use the *Gershgorin circle theorem*, according to which, for some $k \in [n]$, every eigenvalue λ of an $n \times n$ matrix $\mathbf{A} = \{a_{ij}\}$ satisfies

$$|\lambda - a_{kk}| \leq \sum_{\substack{j=1 \\ j\neq k}}^{n} |a_{kj}|.$$

Trees, UBTs, and their terminal encodings. Here we introduce notation, definitions, and standard results for general trees, UBTs, and their terminal encodings. We begin by observing that each UBT induces a family of PLMs (and, hence, of PLDMs), one for each permutation of its leaf labels. Any two such matrices are related by a simultaneous permutation of rows and columns and therefore share the same eigenvalues; their eigenvectors are related by the corresponding permutation of entries. Throughout the article, we exploit this equivalence freely: whenever a spectral property of the PLDM of a given UBT is under study, we select whichever leaf labeling is most convenient for the argument at hand. In addition, we denote by

- $\mathcal{U}_n$ the set of the $(2n-5)!! = 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n-5)$ distinct UBTs with $n \geq 3$ leaves (see OEIS A000672 in [16]);
- Θ_n the set of the PLDMs associated with the UBTs in $\mathcal{U}_n$.

The following theorem characterizes Θ_n:

THEOREM 2.1 (Catanzaro et al. [4]). *An integer matrix $\mathbf{2}^{-\tau}$ of order $n \geq 3$ belongs to Θ_n if and only if it satisfies:*

$$
\begin{array}{lll}
\text{(integrality conditions)} & \tau_{ij} \in [2, n-1] & \forall \textit{ distinct } i,j \in [n], \\
\text{(symmetry conditions)} & \tau_{ji} = \tau_{ij} & \forall \textit{ distinct } i,j \in [n], \\
\text{(strong triangle inequalities)} & & \\
 & \tau_{ij} + \tau_{jk} - \tau_{ik} \geq 2 & \forall \textit{ distinct } i,j,k \in [n], \\
\text{(Kraft's equalities)} & \displaystyle\sum_{\substack{j\in[n] \\ j\neq i}} 2^{-\tau_{ij}} = \frac{1}{2} & \forall\ i \in [n],
\end{array}
$$

and, for all distinct $i, j, p, q \in [n]$, exactly one of the following conditions:

$$
\begin{array}{ll}
 & \tau_{ij} + \tau_{pq} + 2 \leq \tau_{ip} + \tau_{jq} = \tau_{iq} + \tau_{jp}, \\
\text{(strong four-point conditions)} & \tau_{ip} + \tau_{jq} + 2 \leq \tau_{ij} + \tau_{pq} = \tau_{iq} + \tau_{jp}, \\
 & \tau_{iq} + \tau_{jp} + 2 \leq \tau_{ij} + \tau_{pq} = \tau_{ip} + \tau_{jq}.
\end{array}
$$

A *cherry* $\langle i, j\rangle$ of a UBT $T \in \mathcal{U}_n$ is the subgraph induced by the leaves i and j of T together with their unique adjacent internal vertex, hereafter referred to as the

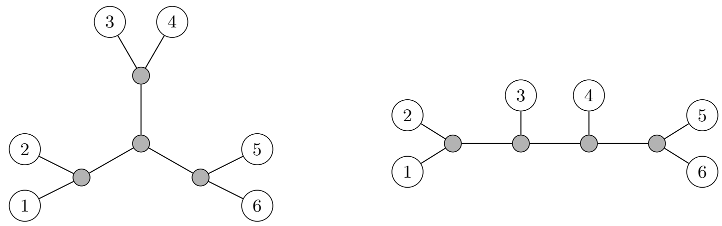


FIG. 2. *The fundamental most-balanced UBT (left) and the fundamental caterpillar (right) for* $n = 6$.

cherry parent and denoted by $\mathrm{par}\langle i, j\rangle$. Since UBTs are planar graphs, we refer to the labeling obtained by ordering the leaves clockwise starting from a leaf labeled 1 and belonging to a cherry as the *fundamental labeling*; the associated UBT, PLM, and PLDM are called the *fundamental UBT*, the *fundamental PLM*, and the *fundamental PLDM*, respectively. We refer to the *structure* of a UBT T as the tree obtained from T by removing all the vertex labels. Figure 2 shows the fundamental labeling of the only two distinct UBT structures in $\mathcal{U}_6$.

The *diameter* of a tree T is the number of edges on the longest path between any pair of its leaves. A *caterpillar* $T \in \mathcal{U}_n$ is a UBT having the maximum diameter, namely $n-1$. A caterpillar has only two cherries, because the endpoints of a longest leaf-to-leaf path must each belong to a distinct cherry as each internal vertex has degree 3. In addition, every internal vertex belonging to such path, except for those belonging to these two cherries, is adjacent to exactly one of the remaining $n-4$ leaves. The two cherries of a *fundamental caterpillar* $T^{\mathrm{c}} \in \mathcal{U}_n$ are $\langle 1, 2\rangle$ and $\langle n-1, n\rangle$. The UBT on the right in Figure 2 is the fundamental caterpillar with 6 leaves.

Two leaves $i, j \in [n]$ of a UBT $T \in \mathcal{U}_n$ with associated PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ belong to the cherry $\langle i, j\rangle$ if and only if the corresponding (i, j)-entry of matrix $\mathbf{2}^{-\boldsymbol{\tau}}$ is equal to $1/4$. In this case, rows i and j of $\mathbf{2}^{-\boldsymbol{\tau}}$ are called *twins* [3], in the sense that

$$\begin{aligned} 2^{-\tau_{ij}} &= \frac{1}{4}, \\ 2^{-\tau_{ir}} &= 2^{-\tau_{jr}} \qquad \forall\, r \in [n] \setminus \{i, j\}. \end{aligned}$$

Throughout the article, we shall assume without loss of generality that twin rows are consecutive, i.e., $j = i + 1$. Since every UBT $T \in \mathcal{U}_n$ has at least two cherries, every PLDM in Θ_n contains at least two pairs of twin rows.

The notion of cherry admits the following natural generalization. We refer to a cherry of T as a *0th-order cherry*. For $d \geq 1$, if T contains two vertex-disjoint $(d-1)$th-order cherries C and C' whose root vertices are both adjacent to a common vertex v, then the subgraph induced by C, C', and v is called a *dth-order cherry* of T, and v is called the *parent* of C and C'. Figure 3 illustrates four 0th-order cherries, two 1st-order cherries, and one 2nd-order cherry.

To streamline notation, for a dth-order cherry C with 2^{d+1} leaves, we write $\langle \ell\rangle$ and $\mathrm{par}\langle \ell\rangle$, where ℓ denotes the ordered set of leaves of C. The ordering is chosen so that the first and last 2^d leaves correspond to the two $(d-1)$th-order cherries contained as subtrees in C. For example, $\langle i, i+1, i+2, i+3\rangle$ denotes the 1st-order cherry formed by the cherries $\langle i, i+1\rangle$ and $\langle i+2, i+3\rangle$. Hereafter, we write *cherry* in place of *0th-order cherry*.

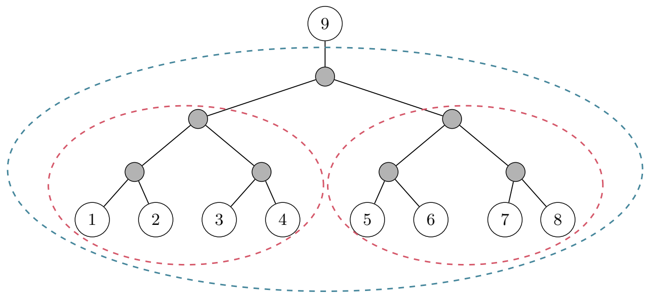


FIG. 3. *A UBT displaying a 2nd-order cherry. The subtree delimited by the dashed blue line is the 2nd-order cherry* $\langle 1,2,3,4,5,6,7,8\rangle$. *It is composed of the two 1st-order cherries* $\langle 1,2,3,4\rangle$ *and* $\langle 5,6,7,8\rangle$, *highlighted by dashed red lines, which are in turn composed of the 0th-order cherries* $\langle 1,2\rangle$ *and* $\langle 3,4\rangle$, *and* $\langle 5,6\rangle$ *and* $\langle 7,8\rangle$, *respectively.*

Contraction of a cherry in a UBT. Let $T \in \mathcal{U}_n$ denote a UBT with c cherries, and assume without loss of generality that $\langle i, i+1\rangle$ is a cherry of T. The *contraction* of T with respect to $\langle i, i+1\rangle$ is defined as the operation that deletes leaves i and $i+1$ and their incident edges, and relabels their parent $\mathrm{par}\langle i, i+1\rangle$ as i. The resulting tree T' belongs to $\mathcal{U}_{n-1}$, as the operation removes a leaf while preserving connectivity and the degrees of all remaining vertices. Moreover, T' has c cherries if there exists a leaf k such that $\tau_{ik} = \tau_{i+1,k} = 3$; otherwise, it has $c-1$ cherries.

Figures 4 and 5 illustrate both cases. In particular, Figure 4 shows a UBT T having a 1st-order cherry on the left, and a 0th-order cherry and a leaf not belonging to any cherry on the right. Figure 5 shows the result of two distinct contractions: on the left, the UBT with two cherries obtained by contracting T with respect to $\langle 1,2\rangle$; on the right, the UBT with three cherries obtained by contracting T with respect to $\langle 6,7\rangle$.

The PLDM $\mathbf{2}^{-\boldsymbol{\tau}'}$ of the contracted tree $T' \in \mathcal{U}_{n-1}$ is obtained from the PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ of $T \in \mathcal{U}_n$ via the following two operations:

(i) multiply by 2 each entry in row i and column i, since for each leaf $k \in [n] \setminus \{i, i+1\}$ the distance from k to the vertex relabeled i in T' is exactly one less than its distance to either i or $i+1$ in T;

(ii) remove row $i+1$ and column $i+1$.

By analogy with T', we shall also refer to $\mathbf{2}^{-\boldsymbol{\tau}'}$ as a *contraction* of $\mathbf{2}^{-\boldsymbol{\tau}}$. We emphasize that, after a contraction, the ith row and column of $\mathbf{2}^{-\boldsymbol{\tau}'}$ correspond to the vertex labeled i, which is a leaf of T'. Figure 6 displays the UBT obtained by contracting

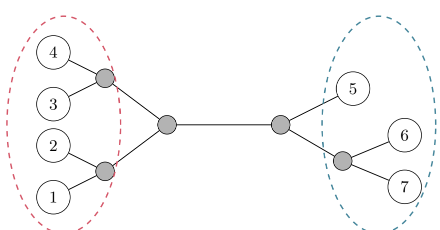


FIG. 4. *An example of a UBT* T *on 7 leaves with three cherries:* $\langle 1,2\rangle$, $\langle 3,4\rangle$, *and* $\langle 6,7\rangle$. *The first two cherries are enclosed by a dashed red line, while the second cherry is enclosed by a dashed blue line together with leaf 6. Contracting cherries* $\langle 1,2\rangle$ *and* $\langle 6,7\rangle$ *leaves a single cherry inside both dashed lines.*

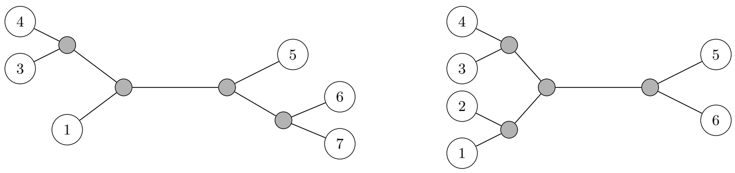


FIG. 5. *UBTs obtained from the UBT in Figure 4 by cherry contraction. Left: contraction of* $\langle 1, 2\rangle$. *Right: contraction of* $\langle 6, 7\rangle$.

the tree in Figure 1, together with the associated PLM and PLDM.

Let $T \in \mathcal{U}_n$ denote a UBT with at least c cherries, and let $T' \in \mathcal{U}_{n-2c}$ denote the UBT obtained from T by simultaneously contracting any c vertex-disjoint cherries. We say that T' is obtained from T through a *c-contraction*, and we refer to the PLDM $\mathbf{2}^{-\boldsymbol{\tau}'}$ of T' as the *c-contraction* of the PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ of T. If T' is obtained by contracting all cherries of T, we say that T' is obtained from T through a *full contraction*.

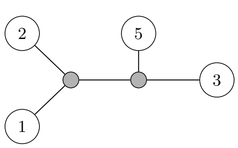
Let $k \in \mathbb{N}$ and $n = 3 \cdot 2^k$. We say that a UBT $T \in \mathcal{U}_n$ is *most-balanced* if it has $3 \cdot 2^{k-1}$ cherries and, for every $h \in [k]$, the UBT obtained after h successive full contractions has $3 \cdot 2^{k-h}$ cherries. Equivalently, T is formed by three kth-order cherries sharing a common parent. In particular, a most-balanced UBT has minimum diameter, namely, $2k + 2$. The *fundamental most-balanced UBT* $T^{\mathrm{b}} \in \mathcal{U}_n$ is the most-balanced UBT whose cherries are exactly $\langle 2i - 1, 2i\rangle$ for all $i \in [n/2]$. The UBT displayed on the left of Figure 2 is the fundamental most-balanced UBT with $n = 3 \cdot 2^1 = 6$ leaves.

3. On the eigenspaces of PLDMs of UBTs. In this section, we study the spectrum of the PLDMs in Θ_n and its relation to the structural properties of the associated UBTs. Specifically, we establish a connection between certain substructures of UBTs, such as cherries and dth-order cherries, and some particular dyadic eigenvalues of their PLDMs. In addition, Section 3.1 provides a closed-form expression for the eigenvalues of PLDMs of most-balanced UBTs, and Section 3.2 characterizes the spectra of the PLDMs associated with caterpillars. We begin with the following proposition.

PROPOSITION 3.1. Let $\mathbf{2}^{-\boldsymbol{\tau}}$ denote a PLDM in Θ_n. Then,

(i) the greatest eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}}$ is equal to $\frac{1}{2}$, with associated eigenvector $\mathbf{1}_n$;
(ii) $\mathbf{2}^{-\boldsymbol{\tau}}$ has an eigenvalue equal to $-\frac{1}{4}$ for each cherry $\langle i, j\rangle$ of T, with associated eigenvector $\mathbf{e}_i - \mathbf{e}_j$.

Proof. (i). By (Kraft's equalities), each row of the matrix $\mathbf{2}^{-\boldsymbol{\tau}}$ sums to $\frac{1}{2}$. The characteristic equation $\det(\mathbf{2}^{-\boldsymbol{\tau}} - \lambda \mathbf{I}_n) = 0$ is then satisfied when $\lambda = \frac{1}{2}$ and this value is maximum due to the Gershgorin circle theorem.



$$\boldsymbol{\tau}' = \begin{pmatrix} 0 & 2 & 3 & 3 \\ 2 & 0 & 3 & 3 \\ 3 & 3 & 0 & 2 \\ 3 & 3 & 2 & 0 \end{pmatrix} \qquad \mathbf{2}^{-\boldsymbol{\tau}'} = \begin{pmatrix} 0 & 2^{-2} & 2^{-3} & 2^{-3} \\ 2^{-2} & 0 & 2^{-3} & 2^{-3} \\ 2^{-3} & 2^{-3} & 0 & 2^{-2} \\ 2^{-3} & 2^{-3} & 2^{-2} & 0 \end{pmatrix}$$

FIG. 6. *On the left, a UBT T' with 4 leaves obtained as a contraction of the UBT T in Figure 1 with respect to the cherry $\langle 3, 4\rangle$. At the center and on the right, the Path-Length Matrix (PLM) $\boldsymbol{\tau}$ and the Path-Length Density Matrix (PLDM) $\mathbf{2}^{-\boldsymbol{\tau}}$ of T', respectively.*

(ii). Without loss of generality, assume that $\langle 1, 2\rangle$ is one of the cherries of T. Then, the first two rows of $\mathbf{2}^{-\boldsymbol{\tau}}$ are twins, hence $2^{-\tau_{1,2}} = 2^{-\tau_{2,1}} = \frac{1}{4}$ and, $2^{-\tau_{k,1}} = 2^{-\tau_{k,2}}$ for all $k \in [3, n]$. Now, consider the vector $\mathbf{v} = \mathbf{e}_1 - \mathbf{e}_2 = (1, -1, 0, \ldots, 0)$ and observe that the following chain of equalities holds:

$$\begin{aligned}\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{v} &= \left[-\frac{1}{4}, \frac{1}{4}, \left(2^{-\tau_{3,1}} - 2^{-\tau_{3,2}}\right), \ldots, \left(2^{-\tau_{n,1}} - 2^{-\tau_{n,2}}\right)\right]^\top \\ &= \left[\frac{1}{4}, -\frac{1}{4}, 0, \ldots, 0\right]^\top = -\frac{1}{4}\mathbf{v}.\end{aligned}$$

Thus, $\mathbf{v}$ is an eigenvector of $\mathbf{2}^{-\boldsymbol{\tau}}$ with associated eigenvalue $-\frac{1}{4}$. ∎

In what follows, we will see that the PLDMs of a most balanced UBT and of a caterpillar, respectively, are characterized by distinctive spectra exhibiting opposite extremal structures, yet governed by a single auxiliary matrix.

Let $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ denote a PLDM whose associated UBT $T \in \mathcal{U}_n$ has at least c cherries. Without loss of generality, we assume that the first $2c$ rows of $\mathbf{2}^{-\boldsymbol{\tau}}$ correspond to the first c cherries of T, so that, for each $i \in [c]$, the i-th cherry is $\langle 2i - 1, 2i\rangle$. By Proposition 3.1, the matrix $\mathbf{2}^{-\boldsymbol{\tau}}$ has $-1/4$ as an eigenvalue with multiplicity at least c. To study the remaining $n - c$ eigenvalues, we introduce the auxiliary matrices $\mathbf{U}, \hat{\mathbf{U}} \in \mathbb{R}^{n\times(n-c)}$, defined as:

$$\mathbf{U}_{\cdot j} = \begin{cases}\mathbf{e}_{2j-1} + \mathbf{e}_{2j} & \text{if } j \in [c], \\ \mathbf{e}_{j+c} & \text{if } j \in [c+1,\, n-c],\end{cases} \qquad \forall\, j \in [n-c], \tag{3.1}$$

and

$$\hat{\mathbf{U}}_{\cdot j} = \begin{cases}\mathbf{e}_{2j} & \text{if } j \in [c], \\ \mathbf{e}_{j+c} & \text{if } j \in [c+1,\, n-c],\end{cases} \qquad \forall\, j \in [n-c]. \tag{3.2}$$

In light of the definitions of $\mathbf{U}$ and $\hat{\mathbf{U}} \in \mathbb{R}^{n\times(n-c)}$, we define the matrices $\mathbf{V}, \mathbf{W} \in \mathbb{R}^{(n-c)\times(n-c)}$ as follows:

$$\mathbf{V} = \mathbf{U}^\top\mathbf{U} = \operatorname{diag}(2_{[c]}, 1_{[n-2c]}), \tag{3.3}$$

$$\mathbf{W} = \hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{U}. \tag{3.4}$$

Finally, we observe that $\mathbf{U}$ and $\hat{\mathbf{U}}$ satisfy the following simple properties:

$$(\mathbf{U}\hat{\mathbf{U}}^\top)_{\cdot j} = \begin{cases}\mathbf{0} & \text{if } j \in [2c] \text{ and odd} \\ \mathbf{e}_{j-1} + \mathbf{e}_j & \text{if } j \in [2c] \text{ and even} \\ \mathbf{e}_j & \text{otherwise}\end{cases} \qquad \forall\, j \in [n], \tag{3.5}$$

and

$$\hat{\mathbf{U}}^\top\mathbf{U} = \mathbf{I}_{n-c}. \tag{3.6}$$

Example 3.2. Consider the PLDM in Figure 1 and, in particular, the $c = 2$ cherries of the associated UBT. Then, we have that

$$\mathbf{U} = \begin{pmatrix}1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix}, \qquad \hat{\mathbf{U}} = \begin{pmatrix}0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{pmatrix},$$

$$\mathbf{W} = \begin{pmatrix} 2^{-2} & 2^{-3} & 2^{-3} \\ 2^{-3} & 2^{-2} & 2^{-3} \\ 2^{-2} & 2^{-2} & 0 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathrm{diag}(2_{[2]}, 1_{[1]}) \tag{3.7}$$

and

$$\mathbf{U}\hat{\mathbf{U}}^\top = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \hat{\mathbf{U}}^\top \mathbf{U} = \mathbf{I}_3.$$

∎

The following proposition holds:

PROPOSITION 3.3. *Let $\mathbf{2}^{-\boldsymbol{\tau}}$ denote the PLDM of a UBT $T \in \mathcal{U}_n$ with c cherries. Then, the matrices $\mathbf{W}$ and $\mathbf{2}^{-\boldsymbol{\tau}}$ have the same spectrum, except for the eigenvalues equal to $-\frac{1}{4}$ associated with the c cherries of T.*

Proof. We assume that the first $2c$ rows of $\mathbf{2}^{-\boldsymbol{\tau}}$ correspond to the first c cherries of T, and that matrices $\mathbf{U}$, $\hat{\mathbf{U}}$, $\mathbf{V}$, and $\mathbf{W}$ are defined accordingly. By Proposition 3.1(ii), the eigenvectors associated with the first c cherries, i.e., $\mathbf{e}_{2i-1} - \mathbf{e}_{2i}$ for $i \in [c]$, span the c-dimensional eigenspace

$$S = \{\mathbf{v} \in \mathbb{R}^n : \; v_{2i-1} = -v_{2i} \; \forall i \in [c], \text{ and } v_k = 0 \; \forall k > 2c\},$$

Now, consider the $(n-c)$-dimensional subspace

$$S^\perp = \{\mathbf{v} \in \mathbb{R}^n : v_1 = v_2, \ldots, v_{2c-1} = v_{2c}\}$$

Since S and $S^\perp$ are orthogonal complements, their direct sum is $\mathbb{R}^n$. Hence, $S^\perp$ is the subspace generated by the eigenvectors of $\mathbf{2}^{-\boldsymbol{\tau}}$ not associated with the first c cherries of T. Let $\mathbf{v} \in S^\perp$ and note that, by (3.5),

$$\mathbf{v} = \mathbf{U}\hat{\mathbf{U}}^\top \mathbf{v}.$$

Now let λ be an eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}}$ with associated eigenvector $\mathbf{v} \in S^\perp$, and set $\mathbf{u} = \hat{\mathbf{U}}^\top \mathbf{v}$. Then

$$\begin{aligned} \mathbf{W}\mathbf{u} &= \hat{\mathbf{U}}^\top \mathbf{2}^{-\boldsymbol{\tau}} \mathbf{U}\mathbf{u} = \hat{\mathbf{U}}^\top \mathbf{2}^{-\boldsymbol{\tau}} \mathbf{U}\hat{\mathbf{U}}^\top \mathbf{v} \\ &= \hat{\mathbf{U}}^\top \mathbf{2}^{-\boldsymbol{\tau}} \mathbf{v} = \lambda \hat{\mathbf{U}}^\top \mathbf{v} = \lambda \mathbf{u}, \end{aligned}$$

which proves the claim. □

The matrix $\mathbf{W}$ enjoys several additional properties that are instrumental in characterizing the $n-c$ eigenvalues of the PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ not associated with the first c cherries of T. In particular, $\mathbf{W}$ admits the block decomposition

$$\mathbf{W} = \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix} \tag{3.8}$$

where $\mathbf{W}_{11} \in \mathbb{R}^{2c \times 2c}$ and $\mathbf{W}_{22} \in \mathbb{R}^{(n-2c)\times(n-2c)}$ are symmetric, $\mathbf{W}_{21} = 2\mathbf{W}_{12}^\top$, $\mathbf{W}_{12} \in \mathbb{R}^{2c \times (n-2c)}$, and $\mathbf{W}_{21} \in \mathbb{R}^{(n-2c)\times 2c}$. The diagonal entries of $\mathbf{W}_{11}$ are all equal to $\frac{1}{4}$, and each off-diagonal entry (i,j) is given by $2^{-\tau_{2i-1,\,2j-1}+1}$. The diagonal entries of $\mathbf{W}_{22}$ are all equal to 0, and each off-diagonal entry (i,j) is given by $2^{-\tau_{2c+i,\,2c+j}}$. On the other hand, each entry (i,j) of $\mathbf{W}_{12}$ is given by $2^{-\tau_{2i-1,\,2c+j}}$. In addition, the largest eigenvalue of $\mathbf{W}$ is $\frac{1}{2}$ because $\sum_{j=1}^{n-c} w_{ij} = \frac{1}{2}$ for each $i \in [n-c]$.

Example 3.4. Consider the matrix $\mathbf{W}$ in (3.7). Its block decomposition (3.8) reads as follows:

$$\mathbf{W}_{11} = \begin{pmatrix} 2^{-2} & 2^{-3} \\ 2^{-3} & 2^{-2} \end{pmatrix}, \quad \mathbf{W}_{12} = \begin{pmatrix} 2^{-3} \\ 2^{-3} \end{pmatrix}, \quad \mathbf{W}_{21} = \begin{pmatrix} 2^{-2} & 2^{-2} \end{pmatrix}, \quad \mathbf{W}_{22} = \begin{pmatrix} 0 \end{pmatrix}. \qquad \blacksquare$$

Now, recall that $\mathbf{2}^{-\boldsymbol{\tau}}$ is a PLDM of a UBT $T \in \mathcal{U}$ with at least c cherries. Then, the following proposition shows that matrices $\mathbf{V}$ and $\mathbf{W}$ naturally induce the c-contraction of $\mathbf{2}^{-\boldsymbol{\tau}}$.

PROPOSITION 3.5. *Let $T \in \mathcal{U}_n$ denote a UBT with at least c cherries and $T' \in \mathcal{U}_{n-2c}$ denote the UBT obtained from T by contracting its first c cherries. The matrices $\mathbf{V}$ and $\mathbf{W}$ associated with a PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ of UBT T induce the c-contraction $\mathbf{2}^{-\boldsymbol{\tau}'}$ of $\mathbf{2}^{-\boldsymbol{\tau}}$ through the following equation:*

$$\mathbf{V}\mathbf{W} - \frac{1}{2}\hat{\mathbf{I}}_{n-c} = \mathbf{2}^{-\boldsymbol{\tau}'}. \tag{3.9}$$

Proof. We assume that the first $2c$ rows of the PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ correspond to the first c cherries of T, and that matrices $\mathbf{U}$, $\hat{\mathbf{U}}$, $\mathbf{V}$, and $\mathbf{W}$ are defined accordingly.

Analogously to (3.8), we partition $\mathbf{2}^{-\boldsymbol{\tau}'}$ into blocks conformal with those of $\mathbf{W}$ as follows,

$$\mathbf{2}^{-\boldsymbol{\tau}'} = \begin{pmatrix} \mathbf{2}_{11}^{-\boldsymbol{\tau}'} & \mathbf{2}_{12}^{-\boldsymbol{\tau}'} \\ \mathbf{2}_{21}^{-\boldsymbol{\tau}'} & \mathbf{2}_{22}^{-\boldsymbol{\tau}'} \end{pmatrix}, \tag{3.10}$$

where $\mathbf{2}_{11}^{-\boldsymbol{\tau}'} \in \mathbb{R}^{2c\times 2c}$ and $\mathbf{2}_{22}^{-\boldsymbol{\tau}'} \in \mathbb{R}^{(n-2c)\times(n-2c)}$ are symmetric, $\mathbf{2}_{12}^{-\boldsymbol{\tau}'} \in \mathbb{R}^{2c\times(n-2c)}$, and $\mathbf{2}_{21}^{-\boldsymbol{\tau}'} = (\mathbf{2}_{12}^{-\boldsymbol{\tau}'})^\top$. We now express the entries of each block of (3.10) in terms of the entries of $\mathbf{2}^{-\boldsymbol{\tau}}$. The diagonal entries of both $\mathbf{2}_{11}^{-\boldsymbol{\tau}'}$ and $\mathbf{2}_{22}^{-\boldsymbol{\tau}'}$ are zero. Each off-diagonal entry (i,j) of $\mathbf{2}_{11}^{-\boldsymbol{\tau}'}$ is $2^{-\tau_{2i-1,\,2j-1}+2}$, because the leaves i and j of T' correspond to the parents of the cherries $\langle 2i-1, 2i\rangle$ and $\langle 2j-1, 2j\rangle$ in T, so the corresponding path length decreases by two. Likewise, each off-diagonal entry (i,j) of $\mathbf{2}_{22}^{-\boldsymbol{\tau}'}$ is equal to $2^{-\tau_{2c+i,\,2c+j}}$, since the leaves $c+i$ and $c+j$ of T' correspond to the leaves $2c+i$ and $2c+j$ of T, and hence the path length is preserved. Finally, each entry (i,j) of $\mathbf{2}_{12}^{-\boldsymbol{\tau}'}$ is equal to $2^{-\tau_{2i-1,\,2c+j}+1}$, because the leaf i of T' corresponds to the parent of the cherry $\langle 2i-1, 2i\rangle$ in T, whereas the leaf $c+j$ corresponds to the leaf $2c+j$, so the path length decreases by one.

By definition of $\mathbf{V}$ in (3.3) and the block decomposition (3.8) of the matrix $\mathbf{W}$, the following chain of equalities holds:

$$\mathbf{V}\mathbf{W} - \frac{1}{2}\hat{\mathbf{I}}_{n-c} = \begin{pmatrix} 2\mathbf{I}_{2c} & 0 \\ 0 & \mathbf{I}_{n-c} \end{pmatrix} \begin{pmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix} - \frac{1}{2}\begin{pmatrix} \mathbf{I}_c & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2\mathbf{W}_{11} - \frac{1}{2}\mathbf{I}_c & 2\mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{pmatrix}.$$

It is immediate to verify that the diagonal entries of $2\mathbf{W}_{11} - \frac{1}{2}\mathbf{I}_c$ and $\mathbf{W}_{22}$ are zero. Furthermore, by direct comparison between the entries of the blocks in (3.8) with those in (3.10), we deduce that

$$2\mathbf{W}_{11} - \frac{1}{2}\mathbf{I}_c = \mathbf{2}_{11}^{-\boldsymbol{\tau}'}, \qquad 2\mathbf{W}_{12} = \mathbf{2}_{12}^{-\boldsymbol{\tau}'},$$

$$2\mathbf{W}_{21} = \mathbf{2}_{21}^{-\boldsymbol{\tau}'}, \qquad \mathbf{W}_{22} = \mathbf{2}_{22}^{-\boldsymbol{\tau}'}.$$

Thus, the statement follows. ∎

Proposition 3.5 allows us to extend Proposition 3.1 to higher-order cherries.

THEOREM 3.6. *Let $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ denote the PLDM of a UBT $T \in \mathbb{U}_n$. Assume that T contains a dth-order cherry $C = \langle i, i+1, \ldots, i+2^{d+1}-1\rangle$, for some $i \in \big[n - (2^{d+1}-1)\big]$. Then, $\mathbf{2}^{-\boldsymbol{\tau}}$ has an eigenvalue*

$$\frac{2^{d+1}-3}{2^{d+2}}, \tag{3.11}$$

with associated eigenvector

$$\mathbf{x} = \sum_{k=0}^{2^d-1} \mathbf{e}_{i+k} - \sum_{k=2^d}^{2^{d+1}-1} \mathbf{e}_{i+k}. \tag{3.12}$$

Proof. If $d = 0$, the claim follows directly from Proposition 3.1. Assume now that $d \geq 1$, and suppose by induction that the statement holds for every UBT containing a $(d-1)$th-order cherry. Without loss of generality, also assume that $i = 1$, so that $C = \langle 1, 2, \ldots, 2^{d+1}\rangle$. Then, T has the $c = 2^d$ cherries $\langle 1,2\rangle$,$\langle 3,4\rangle$,…,$\langle 2c-1, 2c\rangle$, which are contained as subtrees in C. In addition, the first $2c$ rows of $\mathbf{2}^{-\boldsymbol{\tau}}$ correspond to such cherries, and the matrices $\mathbf{U}$, $\hat{\mathbf{U}}$, $\mathbf{V}$, and $\mathbf{W}$ are defined accordingly.

Let T' denote the UBT obtained from T through a c-contraction with respect to $\langle 1,2\rangle$,$\langle 3,4\rangle$,…,$\langle 2c-1, 2c\rangle$, and let $\mathbf{2}^{-\boldsymbol{\tau}'}$ denote the PLDM associated with T'. By construction, T' contains a $(d-1)$th-order cherry induced by its first c leaves; each of such leaves is a parent of exactly one of the c cherries above in T. Hence, by the inductive hypothesis, $\mathbf{2}^{-\boldsymbol{\tau}'}$ has eigenvalue $\frac{2^d-3}{2^{d+1}}$ with associated eigenvector

$$\mathbf{y} = \sum_{i=1}^{c/2} \mathbf{e}_i^{(n-c)} - \sum_{i=c/2+1}^{c} \mathbf{e}_i^{(n-c)} = (1_{[c/2]}, -1_{[c/2]}, 0_{[n-2c]})$$

As $\mathbf{2}^{-\boldsymbol{\tau}'}\mathbf{y} = \frac{2^d-3}{2^{d+1}}\,\mathbf{y}$ and $\mathbf{V}^{-1} = \mathrm{diag}(\frac{1}{2}_{[c]}, 1_{[n-c]})$ we have:

$$\begin{aligned}\mathbf{V}^{-1}\mathbf{2}^{-\boldsymbol{\tau}'}\mathbf{y} &= \tfrac{2^d-3}{2^{d+1}}\mathbf{V}^{-1}\left(\sum_{i=1}^{c/2} \mathbf{e}_i^{(n-c)} - \sum_{i=c/2+1}^{c} \mathbf{e}_i^{(n-c)}\right) = \\ &= \tfrac{2^d-3}{2^{d+1}}\left(\frac{1}{2}\sum_{i=1}^{c/2} \mathbf{e}_i^{(n-c)} - \frac{1}{2}\sum_{i=c/2+1}^{c} \mathbf{e}_i^{(n-c)}\right) = \tfrac{2^d-3}{2^{d+2}}\mathbf{y}\end{aligned}$$

Then, by Proposition 3.5 and, in particular, equation (3.9), the following chain of equalities also holds:

$$\mathbf{V}^{-1}\mathbf{2}^{-\boldsymbol{\tau}'}\mathbf{y} = \left(\hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{U} - \tfrac{1}{4}\,\hat{\mathbf{I}}_{n-c}\right)\mathbf{y} = \hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{U}\mathbf{y} - \tfrac{1}{4}\,\hat{\mathbf{I}}_{n-c}\mathbf{y} = \tfrac{2^d-3}{2^{d+2}}\mathbf{y}.$$

which in turn implies

$$\hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{U}\mathbf{y} - \tfrac{1}{4}\,\hat{\mathbf{I}}_{n-2}\mathbf{y} = \tfrac{2^d-3}{2^{d+2}}\mathbf{y} \quad \Rightarrow \quad \hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{U}\mathbf{y} = \tfrac{2^{d+1}-3}{2^{d+2}}\mathbf{y}$$

By observing that, in (3.12), $\mathbf{x}$ is equal to $\mathbf{U}\mathbf{y}$, we obtain

$$\begin{aligned}\hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{U}\mathbf{y} = \tfrac{2^{d+1}-3}{2^{d+2}}\mathbf{y} \quad &\Rightarrow \quad \mathbf{U}\hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{U}\mathbf{y} = \tfrac{2^{d+1}-3}{2^{d+2}}\mathbf{U}\mathbf{y} \\ &\Rightarrow \quad \mathbf{U}\hat{\mathbf{U}}^\top\mathbf{2}^{-\boldsymbol{\tau}}\mathbf{x} = \tfrac{2^{d+1}-3}{2^{d+2}}\mathbf{x}.\end{aligned}$$

Now, set $\mathbf{z} = \mathbf{2}^{-\boldsymbol{\tau}}\mathbf{x}$. As $x_1 = x_2, x_3 = x_4, \ldots, x_{2c-1} = x_{2c}$ and the pairs of rows $(1,2)$, $(3,4), \ldots,$ $(2c-1, 2c)$ are twins rows of $\mathbf{2}^{-\boldsymbol{\tau}}$, it follows that $z_1 = z_2, z_3 = z_4, \ldots, z_{2c-1} = z_{2c}$. Therefore, by (3.5), the condition $\mathbf{z} = \mathbf{U}\hat{\mathbf{U}}^\top \mathbf{z}$ holds. Hence, $\mathbf{U}\hat{\mathbf{U}}^\top \mathbf{2}^{-\boldsymbol{\tau}}\mathbf{x} = \mathbf{2}^{-\boldsymbol{\tau}}\mathbf{x}$, and the statement follows. ∎

Proposition 3.5 shows that $\mathbf{W}$ can be written as a linear combination of the PLDM $\mathbf{2}^{-\boldsymbol{\tau}'}$ and the matrix $\hat{\mathbf{I}}_{n-c}$. Combining this representation with the structural properties of $\mathbf{W}$ yields the following proposition.

Proposition 3.7. *Let $\mathbf{2}^{-\boldsymbol{\tau}}$ denote the PLDM of a UBT $T \in \mathcal{U}_n$ having $c = \frac{n}{2}$ cherries, with $n \geq 6$ and even. Then, the matrix $\mathbf{W}$ associated with $\mathbf{2}^{-\boldsymbol{\tau}}$ and the full contraction $\mathbf{2}^{-\boldsymbol{\tau}'}$ of $\mathbf{2}^{-\boldsymbol{\tau}}$ satisfy*

$$\mathbf{W} = \tfrac{1}{2}\mathbf{2}^{-\boldsymbol{\tau}'} + \tfrac{1}{4}\hat{\mathbf{I}}_{n-\frac{n}{2}}. \tag{3.13}$$

In addition, denoted by $\lambda_1, \lambda_2, \cdots, \lambda_c$ and $\mu_1, \mu_2, \cdots, \mu_c$ the eigenvalues of $\mathbf{W}$ and $\mathbf{2}^{-\boldsymbol{\tau}'}$, respectively, we have that

$$\lambda_i = \tfrac{1}{2}\mu_i + \tfrac{1}{4}, \qquad \forall\, i \in [c]. \tag{3.14}$$

Proof. We first observe that (3.9) can be rewritten as

$$\mathbf{W} = \mathbf{V}^{-1}\left(\mathbf{2}^{-\boldsymbol{\tau}'} + \frac{1}{2}\hat{\mathbf{I}}_{n-c}\right). \tag{3.15}$$

By the definition of $\mathbf{V}$ and the assumption $n - c = n/2$, equation (3.15) is equivalent to (3.13). Finally, (3.14) follows from (3.13) by applying the standard result on the eigenvalues of a sum with a scalar multiple of the identity matrix. ∎

3.1. On the spectrum of the PLDM of the fundamental most-balanced UBT. In this section, we determine a closed-form expression for the spectrum of the PLDM of a most-balanced UBT. To this end, we first note that a most-balanced UBT $T \in \mathcal{U}_{3 \cdot 2^k}$ can be seen as the union of a kth-order cherry and a $(k-1)$th-order cherry having disjoint vertex sets and adjacent parents. For example, the fundamental most-balanced UBT displayed on the left of Figure 2 may be viewed as consisting of the 0th-order cherry $\langle 1, 2\rangle$ and the 1st-order cherry $\langle 3,4,5,6\rangle$.

In light of the previous propositions and observations, the following result holds:

Theorem 3.8. *For a fixed $k \in \mathbb{N}$, the PLDM $\mathbf{2}^{-\boldsymbol{\tau}}$ of a most-balanced $T \in \mathcal{U}_{3 \cdot 2^k}$ has eigenvalues*

- $\frac{2^d - 3}{2^{d+1}}$ *with multiplicity* $3 \cdot 2^{k-d}$*, for each* $d \in [k]$*;*
- $\frac{2^{k+1} - 3}{2^{k+2}}$ *with multiplicity* 2*;*
- $\frac{1}{2}$ *with multiplicity* 1*.*

Proof. By definition of a dth-order cherry, for each $d \in [k]$, T contains $3 \cdot 2^{k-d}$ distinct $(d-1)$th-order cherries. Hence, by Theorem 3.6, for each $d \in [k]$, $\mathbf{2}^{-\boldsymbol{\tau}}$ has eigenvalue $\frac{2^d-3}{2^{d+1}}$ with multiplicity $3 \cdot 2^{k-d}$. In addition, T has a unique kth-order cherry. Hence, $\mathbf{2}^{-\boldsymbol{\tau}}$ has eigenvalue $\frac{2^{k+1}-3}{2^{k+2}}$. Moreover, as $\mathbf{2}^{-\boldsymbol{\tau}}$ is a PLDM, by Proposition 3.1, $\frac{1}{2}$ is also an eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}}$. Finally, by using the fact that $\mathrm{tr}(\mathbf{2}^{-\boldsymbol{\tau}}) = 0$ and that

$$\sum_{d=1}^{k} 3 \cdot 2^{k-d}\frac{2^d - 3}{2^{d+1}} + \frac{2^{k+1} - 3}{2^{k+2}} + \frac{1}{2} = -\frac{2^{k+1} - 3}{2^{k+2}},$$

we can conclude that the eigenvalue $\frac{1}{2}$ has multiplicity 1, whereas the eigenvalue $\frac{2^{k+1}-3}{2^{k+2}}$ has multiplicity 2. Thus, the statement follows. □

3.2. On the spectrum of the PLDM of the fundamental caterpillar. In this section, we establish the eigenvalues of the PLDM of the fundamental caterpillar in $\mathcal{U}_n$, hereafter denoted by $\mathbf{C}(n)$. We begin by redefining the matrices $\mathbf{U}$, $\hat{\mathbf{U}}$, and $\mathbf{V}$ as follows to match the positions of the two cherries $\langle 1, 2\rangle$ and $\langle n-1, n\rangle$ in the fundamental caterpillar $T^{\mathrm{c}} \in \mathcal{U}_n$:

$$\mathbf{U}_{\cdot j} = \begin{cases} \mathbf{e}_1 + \mathbf{e}_2 & \text{if } j = 1, \\ \mathbf{e}_{j+1} & \text{if } j \in [2,\, n-3], \\ \mathbf{e}_{n-1} + \mathbf{e}_n & \text{if } j = n-2, \end{cases} \quad \forall\, j \in [1, n-2],$$

$$\hat{\mathbf{U}}_{\cdot j} = \begin{cases} \mathbf{e}_2 & \text{if } j = 1, \\ \mathbf{e}_{j+1} & \text{if } j \in [2,\, n-3], \\ \mathbf{e}_n & \text{if } j = n-2, \end{cases} \quad \forall\, j \in [1, n-2],$$

$$\mathbf{V} = \mathbf{U}^\top \mathbf{U} = \mathrm{diag}(2, 1_{[n-4]}, 2).$$

Observe that this redefinition still satisfies (3.4) and (3.6).

We also introduce the dyadic matrix $\mathbf{S} \in \mathbb{R}^{(n-2)\times(n-2)}$ having $s_{ij} = 2^{-|i-j|}$ as its generic entry, which satisfies the following property:

$$\mathbf{C}(n) = \frac{1}{4}(\mathbf{U}\mathbf{S}\mathbf{U}^\top - \mathbf{I}_{n-2}). \tag{3.16}$$

Example 3.9. Consider the fundamental caterpillar T^{c} in Figure 2. Then, the PLDM $\mathbf{C}(6)$ associated with T^{c} reads as follows

$$\mathbf{C}(6) = \begin{pmatrix} 0 & 2^{-2} & 2^{-3} & 2^{-4} & 2^{-5} & 2^{-5} \\ 2^{-2} & 0 & 2^{-3} & 2^{-4} & 2^{-5} & 2^{-5} \\ 2^{-3} & 2^{-3} & 0 & 2^{-3} & 2^{-4} & 2^{-4} \\ 2^{-4} & 2^{-4} & 2^{-3} & 0 & 2^{-3} & 2^{-3} \\ 2^{-5} & 2^{-5} & 2^{-4} & 2^{-3} & 0 & 2^{-2} \\ 2^{-5} & 2^{-5} & 2^{-4} & 2^{-3} & 2^{-2} & 0 \end{pmatrix} \tag{3.17}$$

and, accordingly the matrices $\mathbf{U}$, $\hat{\mathbf{U}}$ and $\mathbf{V}$ associated with $\mathbf{C}(6)$ are

$$\mathbf{U} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \hat{\mathbf{U}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{V} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

In addition,

$$\mathbf{S} = \begin{pmatrix} 1 & 2^{-1} & 2^{-2} & 2^{-3} \\ 2^{-1} & 1 & 2^{-1} & 2^{-2} \\ 2^{-2} & 2^{-1} & 1 & 2^{-1} \\ 2^{-3} & 2^{-2} & 2^{-1} & 1 \end{pmatrix}$$

Then, it is easy to verify that the following conditions hold:

$$\mathbf{C}(6) = \frac{1}{4}(\mathbf{U}\mathbf{S}\mathbf{U}^\top - \mathbf{I}_6) \quad \text{and} \quad \mathbf{W} = \hat{\mathbf{U}}^\top \mathbf{C}(6)\mathbf{U} = \begin{pmatrix} 2^{-2} & 2^{-3} & 2^{-4} & 2^{-4} \\ 2^{-2} & 0 & 2^{-3} & 2^{-3} \\ 2^{-3} & 2^{-3} & 0 & 2^{-2} \\ 2^{-4} & 2^{-4} & 2^{-3} & 2^{-2} \end{pmatrix}.$$

∎

In light of these new definitions, the following proposition holds:

PROPOSITION 3.10. *The matrix* $\mathbf{W}$ *associated with* $\mathbf{C}(n)$ *satisfies*

$$\mathbf{W} = \frac{1}{4}(\mathbf{S}\mathbf{V} - \mathbf{I}_{n-2}), \tag{3.18}$$

where $\mathbf{S} \in \mathbb{R}^{(n-2)\times(n-2)}$ *is a dyadic matrix having* $s_{ij} = 2^{-|i-j|}$ *as its generic entry.*

Proof. By (3.4) and (3.16), we have that

$$\mathbf{W} = \hat{\mathbf{U}}^\top \mathbf{C}(n)\mathbf{U} = \frac{1}{4}(\hat{\mathbf{U}}^\top \mathbf{U}\mathbf{S}\mathbf{U}^\top \mathbf{U} - \hat{\mathbf{U}}^\top \mathbf{I}_{n-2}\mathbf{U}).$$

Then, since $\hat{\mathbf{U}}^\top \mathbf{U} = \mathbf{I}_{n-2}$ and $\mathbf{U}^\top \mathbf{U} = \mathbf{V}$:

$$\mathbf{W} = \frac{1}{4}(\mathbf{I}_{n-2}\mathbf{S}\mathbf{V} - \hat{\mathbf{U}}^\top \mathbf{I}_{n-2}\mathbf{U}) = \frac{1}{4}(\mathbf{S}\mathbf{V} - \mathbf{I}_{n-2}),$$

thus, the statement follows. □

The spectral properties of matrix $\mathbf{S}$ have been studied by Fikioris [11], who proved the following result that is instrumental to characterize the spectrum of the fundamental caterpillar:

PROPOSITION 3.11 (Fikioris [11], Theorem 2.1). *Let*

$$\mathbf{S} = \left[2^{-|i-j|}\right]_{i,j\in[n]} \in \mathbb{R}^{n\times n}.$$

Then, $\mathbf{S}$ *is invertible and its inverse is the tridiagonal matrix*

$$\mathbf{S}^{-1} = \frac{1}{3}\begin{pmatrix} 4 & -2 & 0 & 0 & 0 & \dots & 0 \\ -2 & 5 & -2 & 0 & 0 & \dots & 0 \\ 0 & -2 & 5 & -2 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \ddots & & \vdots \\ & \vdots & \ddots & -2 & 5 & -2 & 0 \\ & & & & -2 & 5 & -2 \\ 0 & \dots & \dots & & 0 & -2 & 4 \end{pmatrix}. \tag{3.19}$$

Furthermore, the eigenvalues $\gamma_1 \le \gamma_2 \le \ldots \le \gamma_n$ *of* $\mathbf{S}$ *are*

$$\gamma_i = \frac{3}{5 - 4\cos(\chi_i)}, \quad \forall\ i \in [n], \tag{3.20}$$

where χ_i *is given by the unique root* $\frac{i\pi}{n} \le \chi_i \le \frac{(i+1)\pi}{n+1}$ *of*

$$\begin{cases} 2\cos\frac{(n+1)\chi}{2} = \cos\frac{(n-1)\chi}{2} & \text{if } i \text{ even} \\ 2\sin\frac{(n+1)\chi}{2} = \sin\frac{(n-1)\chi}{2} & \text{if } i \text{ odd} \end{cases}$$

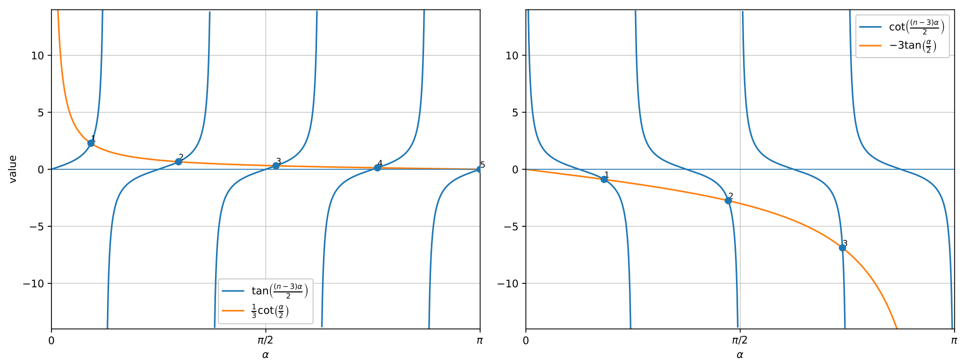


FIG. 7. *A graphical representation of the solutions to* (3.22) *for* $n = 11$.

S is usually referred to as a *Kac-Murdock-Szegö matrix* [13] in the literature. In light of Fikioris' result and the properties of both tridiagonal matrices [9, 19] and the Kac-Murdock-Szegö matrix [11, 13], the following result (whose proof is reported in Appendix A for the sake of conciseness) holds:

THEOREM 3.12. *Let* $\mathbf{C}(n)$ *denote the PLDM of the fundamental caterpillar* T *with* $n \geq 4$ *leaves. Then, the eigenvalues* $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ *of* $\mathbf{C}(n)$ *are*

$$\lambda_i = \begin{cases} -\frac{1}{4}, & \text{if } i \in [2], \\ \frac{1}{4}\left(\frac{3}{5-4\cos\alpha_i} - 1\right), & \text{if } i \in [3, n-1], \\ \frac{1}{2}, & \text{if } i = n, \end{cases} \tag{3.21}$$

where the scalars α_i*,* $i \in [3, n-1]$*, are the* $n-3$ *distinct solutions in the interval* $0 \leq \alpha \leq \pi$ *of*

$$\tan\left(\frac{(n-3)\alpha}{2}\right) = \frac{1}{3}\cot\left(\frac{\alpha}{2}\right), \tag{3.22a}$$

$$\cot\left(\frac{(n-3)\alpha}{2}\right) = -3\tan\left(\frac{\alpha}{2}\right). \tag{3.22b}$$

Figure 7 shows a numerical example related to the application of Theorem 3.12 for $n = 11$. In particular, the figure shows the solutions to (3.22a) and (3.22b) on the left and right, respectively.

The approach used to establish the spectra of the PLDMs of the fundamental most-balanced UBT and the fundamental caterpillar can, in principle, be extended to any UBT in $\mathcal{U}_n$. The main difficulty lies in redefining, on a case-by-case basis, the matrices $\mathbf{U}$, $\hat{\mathbf{U}}$, and $\mathbf{V}$, as well as the relationships linking $\mathbf{V}$ and $\mathbf{W}$. A general formulation of this approach remains an open problem.

4. On the extremal properties of PLDMs of UBTs. In the previous sections, we have shown that $-1/4$ is the smallest eigenvalue of the PLDMs of the fundamental most-balanced UBT and the fundamental caterpillar, respectively. Moreover, the multiplicity of this eigenvalue equals the number of cherries in the corresponding UBTs. By using the notion of a contraction, in this section, we extend this result to arbitrary PLDMs in Θ_n. To this end, consider a UBT $T \in \mathcal{U}_n$ and its contraction T' on the cherries $\langle i, i+1 \rangle$, for some $i \in [n]$. Denote by $\mathbf{2}^{-\boldsymbol{\tau}}$ and $\mathbf{2}^{-\boldsymbol{\tau}'}$ the PLDMs of

T and T', respectively, and by $\mathbf{E}$ a symmetric matrix in $\mathbb{R}^{n\times n}$, whose i-th row and column coincide with the i-th row and column of $\mathbf{2}^{-\boldsymbol{\tau}}$, respectively, i.e,

$$\mathbf{E} = \mathbf{e}_i(\mathbf{2}_{i\cdot}^{-\boldsymbol{\tau}})^\top + \mathbf{2}_{i\cdot}^{-\boldsymbol{\tau}}\mathbf{e}_i^\top.$$

Finally, let $\mathbf{R}$ denote a matrix in $\mathbb{R}^{n\times(n-1)}$ defined as

$$\mathbf{R}_{\cdot j} = \begin{cases} \mathbf{e}_j, & \text{if } j \in [1,\, i], \\ \mathbf{e}_{j+1}, & \text{if } j \in [i+1,\, n-1], \end{cases} \qquad \forall\, j \in [1,\, n-1]$$

and hereafter referred to as the *removal matrix*. Then, it is easy to realize that $\mathbf{2}^{-\boldsymbol{\tau}}$ and $\mathbf{2}^{-\boldsymbol{\tau}'}$ are related by the following equation:

$$\mathbf{2}^{-\boldsymbol{\tau}'} = \mathbf{R}^\top \left(\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}\right)\mathbf{R}. \tag{4.1}$$

Example 4.1. Consider again the fundamental caterpillar T^{c} in Figure 2 and its PLDM $\mathbf{C}(6)$ provided in (3.17). Suppose we perform a contraction of T^{c} with respect to the cherry $\langle 1, 2\rangle$. We then show that this operation can be carried out by applying (4.1) with respect to the first row of $\mathbf{C}(6)$, i.e.,

$$\mathbf{2}_{1\cdot}^{-\boldsymbol{\tau}} = \begin{pmatrix} 0 & 2^{-2} & 2^{-3} & 2^{-4} & 2^{-5} & 2^{-5}\end{pmatrix}^\top.$$

To this end, first observe that

$$\mathbf{E} = \mathbf{e}_1(\mathbf{2}_{1\cdot}^{-\boldsymbol{\tau}})^\top + \mathbf{2}_{1\cdot}^{-\boldsymbol{\tau}}\mathbf{e}_1^\top = \begin{pmatrix} 0 & 2^{-2} & 2^{-3} & 2^{-4} & 2^{-5} & 2^{-5} \\ 2^{-2} & 0 & 0 & 0 & 0 & 0 \\ 2^{-3} & 0 & 0 & 0 & 0 & 0 \\ 2^{-4} & 0 & 0 & 0 & 0 & 0 \\ 2^{-5} & 0 & 0 & 0 & 0 & 0 \\ 2^{-5} & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The associated removal matrix R and $\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}$ are given by:

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad \mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E} = \begin{pmatrix} 0 & 2^{-1} & 2^{-2} & 2^{-3} & 2^{-4} & 2^{-4} \\ 2^{-1} & 0 & 2^{-3} & 2^{-4} & 2^{-5} & 2^{-5} \\ 2^{-2} & 2^{-3} & 0 & 2^{-3} & 2^{-4} & 2^{-4} \\ 2^{-3} & 2^{-4} & 2^{-3} & 0 & 2^{-3} & 2^{-3} \\ 2^{-4} & 2^{-5} & 2^{-4} & 2^{-3} & 0 & 2^{-2} \\ 2^{-4} & 2^{-5} & 2^{-4} & 2^{-3} & 2^{-2} & 0 \end{pmatrix}.$$

Now, premultiplying $\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}$ by $\mathbf{R}^\top$ and, then, postmultiplying by $\mathbf{R}$ has the effect of removing the second row and column from $\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}$, leading to

$$\mathbf{2}^{-\boldsymbol{\tau}'} = \mathbf{R}^\top\left(\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}\right)\mathbf{R} = \begin{pmatrix} 0 & 2^{-2} & 2^{-3} & 2^{-4} & 2^{-4} \\ 2^{-2} & 0 & 2^{-3} & 2^{-4} & 2^{-4} \\ 2^{-3} & 2^{-3} & 0 & 2^{-3} & 2^{-3} \\ 2^{-4} & 2^{-4} & 2^{-3} & 0 & 2^{-2} \\ 2^{-4} & 2^{-4} & 2^{-3} & 2^{-2} & 0 \end{pmatrix}$$

which encodes the PLDM of $\mathbf{C}(5)$ and, incidentally, is equivalent to the PLDM shown in Figure 1 through symmetric permutations of rows and columns. ∎

Equation (4.1) and the particular structure of the matrices $\mathbf{E}$ and $\mathbf{R}$ can be exploited to study how the multiplicity of the eigenvalue $-1/4$ of a PLDM $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ changes after a contraction on the associated UBT. Specifically, consider the vector $\mathbf{v} = \mathbf{2}_{i\cdot}^{-\boldsymbol{\tau}}$ and observe that $\operatorname{im}(\mathbf{E})$ is contained in $\operatorname{span}\{\mathbf{e}_i, \mathbf{v}\}$. As $\mathbf{e}_i^\top \mathbf{v} = 0$, the vectors $\mathbf{e}_i$ and $\mathbf{v}$ are linearly independent, hence $\operatorname{rank}(\mathbf{E}) = 2$. In addition, as $\mathbf{E}\mathbf{e}_i = \mathbf{v}$ and $\mathbf{E}\mathbf{v} = \|\mathbf{v}\|^2 \mathbf{e}_i$, the restriction of $\mathbf{E}$ to $\operatorname{span}\{\mathbf{e}_i, \mathbf{v}\}$, with respect to the basis $\{\mathbf{e}_i, \mathbf{v}\}$, is represented by the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & \|\mathbf{v}\|^2 \\ 1 & 0 \end{pmatrix}.$$

Because the eigenvalues are invariant under a change of basis, the nonzero eigenvalues of $\mathbf{E}$ are the eigenvalues of $\mathbf{A}$, i.e., they are equal to $\pm\|\mathbf{v}\|$. Thus, $\mathbf{E}$ has one positive eigenvalue $\|\mathbf{v}\|$, one negative eigenvalue $-\|\mathbf{v}\|$, and one eigenvalue 0 with multiplicity $n-2$. To further advance our analysis, we recall the following two results from the literature, which will be instrumental in the proof of Theorem 4.4.

Proposition 4.2 (Poincaré separation theorem, [12], Cor. 4.3.37). *Let $m, n \in \mathbb{N}$ satisfy $m < n$. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ denote a symmetric matrix, and let $\mathbf{A}' = \mathbf{P}^\top \mathbf{A}\mathbf{P} \in \mathbb{R}^{m\times m}$, where $\mathbf{P} \in \mathbb{R}^{n\times m}$ is a matrix with orthonormal columns. If $\mu_1 \le \cdots \le \mu_n$ and $\nu_1 \le \cdots \le \nu_m$ are the eigenvalues of $\mathbf{A}$ and $\mathbf{A}'$, respectively, then*

$$\mu_i \le \nu_i \le \mu_{i+n-m}, \qquad \forall\, i \in [m].$$

Proposition 4.3 ([12], Cor. 4.3.3). *Let $\mathbf{A}, \mathbf{E} \in \mathbb{R}^{n\times n}$ denote two symmetric matrices, and suppose that $\mathbf{E}$ has exactly one positive and one negative eigenvalue. If $\lambda_1 \le \cdots \le \lambda_n$ and $\mu_1 \le \cdots \le \mu_n$ are the eigenvalues of $\mathbf{A}$ and $\mathbf{A}+\mathbf{E}$, respectively, then*

$$\mu_i \le \lambda_{i+1}, \qquad \forall\, i \in [n-1], \tag{4.2}$$

and

$$\lambda_{i-1} \le \mu_i, \qquad \forall\, i \in [2, n]. \tag{4.3}$$

Equality holds if and only if $\mathbf{E}$ is singular and there exists a nonzero vector $\mathbf{x}$ satisfying

$$\mathbf{A}\mathbf{x} = \lambda_{i+1}\mathbf{x}, \qquad \mathbf{E}\mathbf{x} = \mathbf{0}_n, \qquad (\mathbf{A}+\mathbf{E})\mathbf{x} = \mu_i \mathbf{x}, \tag{4.4}$$

in inequality (4.2)*, and*

$$\mathbf{A}\mathbf{x} = \lambda_{i-1}\mathbf{x}, \qquad \mathbf{E}\mathbf{x} = \mathbf{0}_n, \qquad (\mathbf{A}+\mathbf{E})\mathbf{x} = \mu_i \mathbf{x}, \tag{4.5}$$

in inequality (4.3).

We observe that Proposition 4.3 immediately implies, for both $\mathbf{A}$ and $\mathbf{A}+\mathbf{E}$, that

$$\lambda_1 = \lambda_2 = \cdots = \lambda_k \;\Rightarrow\; \mu_1 = \mu_2 = \cdots = \mu_{k-1}, \tag{4.6}$$

and

$$\mu_1 = \mu_2 = \cdots = \mu_{k-1} < \mu_k \;\Rightarrow\; \lambda_1 = \lambda_2 = \cdots = \lambda_{k-2} < \lambda_{k+1}. \tag{4.7}$$

By (4.6) and the symmetry of the matrices $\mathbf{A}$ and $\mathbf{A}+\mathbf{E}$, it follows that, whenever $\lambda_1 = \cdots = \lambda_k$ holds true, there exist $k-1$ mutually orthogonal eigenvectors of $\mathbf{A}$ satisfying (4.4) and (4.5). Then, the following theorem holds:

THEOREM 4.4. *Let $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ be the PLDM associated with a UBT $T \in \mathcal{U}_n$, with $n \geq 4$, and let T' be the UBT obtained from T by contracting one of its cherries. Let $\mathbf{2}^{-\boldsymbol{\tau}'}$ be the PLDM associated with T', and denote by $m_g\big(-\frac{1}{4}\big)$ and $m'_g\big(-\frac{1}{4}\big)$ the geometric multiplicities of the eigenvalue $-\frac{1}{4}$ in $\mathbf{2}^{-\boldsymbol{\tau}}$ and $\mathbf{2}^{-\boldsymbol{\tau}'}$, respectively. Then*

$$m'_g\left(-\frac{1}{4}\right) \geq m_g\left(-\frac{1}{4}\right) - 1. \tag{4.8}$$

Moreover, the inequality is strict whenever T' has the same number of cherries as T.

Proof. Without loss of generality, assume that $\langle n-1, n\rangle$ is a cherry of T and that T' is obtained from T by contracting precisely this cherry. By Proposition 3.1, the eigenspace associated with $\langle n-1, n\rangle$ is generated by $\mathbf{x}_{\langle n-1,n\rangle} = (\mathbf{e}_{n-1} - \mathbf{e}_n)$. In addition, by (4.1), we have that

$$\mathbf{2}^{-\boldsymbol{\tau}'} = \mathbf{R}^\top\big(\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}\big)\mathbf{R},$$

where $\mathbf{R}$ is such that $\mathbf{R}_{\cdot j} = \mathbf{e}_j$ for $j \in [1,\, n-1]$, and $\mathbf{E} = \mathbf{e}_{n-1}(\mathbf{2}^{-\boldsymbol{\tau}}_{n-1\cdot})^\top + \mathbf{2}^{-\boldsymbol{\tau}}_{n-1\cdot}\mathbf{e}^\top_{n-1}$ is a symmetric singular matrix with exactly one positive eigenvalue and one negative eigenvalue, as observed above. By Proposition 3.1, there are $m_g\big(-\frac{1}{4}\big) - 1$ distinct eigenvectors of $\mathbf{2}^{-\boldsymbol{\tau}}$ that are orthogonal to $\mathbf{x}_{\langle n-1,n\rangle}$ and that are associated with the eigenvalue $-1/4$. Denote by $\mathbf{x}$ any of such eigenvectors. Then, by Proposition 4.3, namely (4.5), $\mathbf{x} \in \ker(\mathbf{E})$, which, in turn, implies that $x_{n-1} = 0$. As $\mathbf{x}^\top_{\langle n-1,n\rangle}\mathbf{x} = 0$, then we also have $x_n = 0$. Therefore, $\mathbf{x}$ has the form $(x_1, x_2, \ldots, x_{n-2}, 0, 0)^\top$.

Now, consider the vector $\bar{\mathbf{x}} = (x_1, x_2, \ldots, x_{n-2}, 0)^\top \in \mathbb{R}^{n-1}$. By definition, $\bar{\mathbf{x}}$ and $\mathbf{x}$ are related by $\bar{\mathbf{x}} = \mathbf{R}^\top\mathbf{x}$ and $\mathbf{x} = \mathbf{R}\bar{\mathbf{x}}$. In addition, the following chain of equalities holds:

$$\mathbf{2}^{-\boldsymbol{\tau}'}\bar{\mathbf{x}} = \mathbf{R}^\top\big(\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}\big)\mathbf{R}\bar{\mathbf{x}} = \mathbf{R}^\top\big(\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}\big)\mathbf{x} = -\frac{1}{4}\mathbf{R}^\top\mathbf{x} = -\frac{1}{4}\bar{\mathbf{x}}. \tag{4.9}$$

Then, $\bar{\mathbf{x}}$ is an eigenvector of $\mathbf{2}^{-\boldsymbol{\tau}'}$ associated with the eigenvalue $-1/4$. It is worth noting that there are $m_g\big(-\frac{1}{4}\big) - 1$ pairwise orthogonal vectors having the same form of $\bar{\mathbf{x}}$. Denote by S the subspace generated by such vectors. As $\dim(S) = m_g\big(-\frac{1}{4}\big) - 1$, inequality (4.8) follows.

Observe that the multiplicity of the eigenvalue $-1/4$ of $\mathbf{2}^{-\boldsymbol{\tau}'}$ satisfies, by construction, $m'_g(-\frac{1}{4}) \leq m_g(-\frac{1}{4})$. Then, to complete the proof, it is sufficient to show that $m'_g(-\frac{1}{4}) = m_g(-\frac{1}{4})$ when T' has the same number of cherries of T. This case occurs when for exactly one leaf of T, that without loss of generality can be labeled with $n-2$, it holds that $\tau_{n-2,n-1} = \tau_{n-2,n} = 3$. Recall that T' is obtained as a contraction of T with respect to the cherry $\langle n-1, n\rangle$; the parent vertex of $\langle n-1, n\rangle$ in T is, by definition, a leaf labeled $n-1$ in T'. Therefore, the last two rows of $\mathbf{2}^{-\boldsymbol{\tau}'}$ are twins and correspond to the cherry $\langle n-2, n-1\rangle$ in T'. This implies that the vector

$$\bar{\mathbf{x}}_{\langle n-2,n-1\rangle} = \big(0_{[n-3]}, 1, -1\big)^\top \in \mathbb{R}^{n-1}$$

is also an eigenvector of $\mathbf{2}^{-\boldsymbol{\tau}'}$ associated with the eigenvalue $-1/4$. As $\bar{\mathbf{x}}_{\langle n-2,n-1\rangle}$ does not belong to the subspace S, we conclude that $m'_g\big(-\frac{1}{4}\big) = m_g\big(-\frac{1}{4}\big)$, and the statement follows. ∎

We now establish that $-1/4$ is the minimum eigenvalue of every PLDM in Θ_n. We begin with the following preliminary proposition.

Proposition 4.5. *Let $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ be a PLDM associated with a UBT $T \in \mathcal{U}_n$, with $n \geq 4$, and let T' be obtained from T by contracting one of its cherries. Let $\mathbf{2}^{-\boldsymbol{\tau}'}$ denote the PLDM associated with T'. If the smallest eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}}$ is $-1/4$, then the smallest eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}'}$ is also $-1/4$.*

Proof. By (4.1), we can write $\mathbf{2}^{-\boldsymbol{\tau}'}$ as

$$\mathbf{2}^{-\boldsymbol{\tau}'} = \mathbf{R}^{\top}\big(\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}\big)\mathbf{R},$$

where $\mathbf{E}$ and $\mathbf{R}$ are defined as in the proof of Theorem 4.4. Denote by $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ the eigenvalues of $\mathbf{2}^{-\boldsymbol{\tau}}$ and by $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_n$ those of $\mathbf{2}^{-\boldsymbol{\tau}} + \mathbf{E}$. Since $\mathbf{2}^{-\boldsymbol{\tau}}$ is the PLDM of a UBT, Proposition 3.1 implies that $-1/4$ is an eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}}$ with multiplicity at least two, hence $\lambda_1 = \lambda_2 = -\frac{1}{4}$. Proposition 4.3, together with the arguments used in the proof of Theorem 4.4, yield $\mu_1 = -\frac{1}{4}$. Finally, as $\mathbf{R}$ is an orthogonal matrix, by Proposition 4.2 the smallest eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}'}$ is greater than or equal to μ_1 and because $\mathbf{2}^{-\boldsymbol{\tau}'}$ is the PLDM of a UBT, by Proposition 3.1 the inequality holds as an equality, and the statement follows. ∎

In light of the propositions above, the following theorem holds:

Theorem 4.6. *Let $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ denote a PLDM of a UBT $T \in \mathcal{U}_n$ with c cherries and $n \geq 4$, then:*

(i) The smallest eigenvalue of $\mathbf{2}^{-\boldsymbol{\tau}}$ is $-1/4$.

(ii) The multiplicity of the eigenvalue $-1/4$ of $\mathbf{2}^{-\boldsymbol{\tau}}$ is equal to c.

Proof. (i). A UBT T can be seen as the result of a sequence σ of contractions carried out on a most-balanced UBT $\hat{T}$ having $n_{\hat{T}} = 3 \cdot 2^{\lceil \frac{n}{3} \rceil}$ leaves. Then, the statement follows by observing that, by Theorem 3.8, the smallest eigenvalue of the PLDM of $\hat{T}$ is $-1/4$ and, by Proposition 4.5, $-1/4$ is the smallest eigenvalue in the spectrum of the PLDM of any UBT in the sequence σ.

(ii). First, observe that a UBT T can always be reduced to a caterpillar $\bar{T}$ with 4 leaves (and 2 cherries) through a sequence σ of at least $c-2$ contractions, each corresponding to a UBT, whose number of cherries is either equal to, or exactly one less than, that of the UBT at the previous step. Consider now the PLDMs associated with the UBTs in σ. By Proposition 3.1, the multiplicity of the eigenvalue $-1/4$ of each of these PLDMs is at least equal to the number of cherries in the corresponding UBTs. Moreover, by Theorem 4.4, the multiplicity of the eigenvalue $-1/4$ may decrease by at most one after a contraction. In addition, observe that such a decrement may only occur when the contraction reduces the number of cherries with respect to the UBT at the previous step in σ. Finally, by Proposition 3.12, the multiplicity of the eigenvalue $-1/4$ for the PLDM of $\bar{T}$ is equal to 2. Since there must be exactly $c-2$ contractions in σ that decrease the number of cherries by one, it follows that the multiplicity of the eigenvalue $-1/4$ decreases exactly $c-2$ times in σ. Therefore,

$$m_g(-\tfrac{1}{4}) = 2 + (c-2) = c,$$

and the statement follows. ∎

Determining whether Theorem 4.6 can be generalized so as to establish a similar relation between higher-order cherries and their associated eigenvalues (given by Theorem 3.6) is at present an open problem.

We now show that, in the case of PLDMs of caterpillars, $-1/4$ is the only eigenvalue less than or equal to $-1/6$.

PROPOSITION 4.7. *Let $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ be the PLDM associated with a caterpillar $T \in \mathcal{U}_n$, with $n \geq 4$, and let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of $\mathbf{2}^{-\boldsymbol{\tau}}$. Then*

$$-\frac{1}{6} < \lambda_i < \frac{1}{2} \qquad \forall\, i \in [3, n-1].$$

Proof. By Theorem 3.12,

$$\lambda_i = \frac{1}{4}\left(\frac{3}{5 - 4\cos\alpha_i} - 1\right) \qquad \forall\, i \in [3, n-1].$$

As $-1 < \cos\alpha_i < 1$ holds true for all $i \in [3, n-1]$, the statement follows. ◻

Determining whether Proposition 4.7 generalizes to any UBT remains open.

We conclude this section by establishing several extremal properties of certain eigenvalues of the PLDMs associated with caterpillars. In particular, we show that both $-1/6$ and $1/2$ are accumulation points in the spectrum of PLDMs, and we determine the asymptotic ratio of the corresponding edge spectral densities.

By Theorem 3.12, the PLDM $\mathbf{C}(n)$ associated with a caterpillar $T \in \mathcal{U}_n$ has the following distinct eigenvalues:

$$\lambda_i = \frac{1}{4}\left(\frac{3}{5 - 4\cos(\alpha_i)} - 1\right) \qquad \forall\, i \in [2, n-1],$$

where α_i satisfies $\frac{(n-2-i)\pi}{n-3} < \alpha_i \leq \frac{(n-1-i)\pi}{n-3}$ for each $i \in [2, n-2]$. Each of such eigenvalues has multiplicity one. For $0 < y < \frac{2}{3}$, define the *empirical cumulative spectral distribution function*

$$f_n(y) = \frac{1}{n}\left|\left\{\lambda : \frac{1}{2} - y \leq \lambda \leq \frac{1}{2}\right\}\right|,$$

i.e., the fraction of the eigenvalues of $\mathbf{C}(n)$ lying in the interval $\left[\frac{1}{2} - y, \frac{1}{2}\right]$. Then, f_n satisfies the following two properties:

$$\begin{aligned}
&f_n(y) \leq f_{n+1}(y),\\
&\frac{(n-2)\alpha(y)}{n\pi} - \frac{1}{n} < f_n(y) < \frac{(n-2)\alpha(y)}{n\pi} + \frac{1}{n},
\end{aligned}$$

where $\alpha(y) = \arccos\left(\frac{5y-3}{4y-3}\right)$. Now, define

$$f(y) = \frac{1}{\pi}\arccos\left(\frac{5y-3}{4y-3}\right).$$

Observe that $f(y)$ provides a uniform approximation to $f_n(y)$, because

$$\sup_{0<y<\frac{2}{3}} |f_n(y) - f(y)| < \frac{4}{n},$$

and therefore

$$\lim_{n\to\infty} f_n(y) = \frac{1}{\pi}\arccos\left(\frac{5y-3}{4y-3}\right).$$

Now, consider the quantity

$$\Delta_n = \frac{\pi g(n)}{n-3},$$

where $g(n) = o(n^{\beta})$, for some $0 < \beta < 1$, and define the *local spectral density* at scale Δ_n by

$$s_n(y) = \frac{f_n(y+\Delta_n) - f_n(y)}{\Delta_n},$$

which measures the fraction of eigenvalues per unit variation in y. Since f_n is a step function, its local behavior is meaningfully captured only over intervals containing a nonconstant number of steps as n grows. In addition, note that $\Delta_n \to 0$ as $n \to \infty$, while the corresponding neighborhood still contains an increasing number of spectral steps.

We are particularly interested in the behavior of $s_n(y)$ when $y = 0$, corresponding to eigenvalues close to $\lambda = \frac{1}{2}$, and when $y = \frac{2}{3} - \Delta_n$, corresponding to eigenvalues close to $\lambda = -\frac{1}{6}$. In these two cases, we have

$$s_n(0) = \frac{f_n(\Delta_n)}{\Delta_n} \qquad \text{and} \qquad s_n\left(\frac{2}{3} - \Delta_n\right) = \frac{1 - f_n\left(\frac{2}{3} - \Delta_n\right)}{\Delta_n}.$$

It is immediate to verify that

$$\lim_{n\to\infty} s_n(0) = \infty \qquad \text{and} \qquad \lim_{n\to\infty} s_n\left(\frac{2}{3} - \Delta_n\right) = \infty.$$

In addition,

$$\lim_{n\to\infty} \frac{s_n(0)}{s_n\left(\frac{2}{3} - \Delta_n\right)} = \frac{1}{9}.$$

This shows that, in the case of PLDMs of caterpillars, the two accumulation points 1/2 and −1/6 are characterized by different asymptotic spectral densities: in particular, the local density near 1/2 is asymptotically one ninth of that near −1/6.

5. Open questions and final remarks.

The results presented in the previous sections allow us to draw some preliminary conclusions, as well as to identify a number of open questions that are supported by experimental evidence.

Open questions. A first open question concerns the range of the eigenvalues of $\mathbf{2}^{-\boldsymbol{\tau}}$. In Section 4, we showed that $-1/6$ and $1/2$ are sharp bounds for all eigenvalues other than $-1/4$ in the case of caterpillars. This suggests the following extension to general UBTs.

CONJECTURE 5.1. *Let $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ be the PLDM associated with a UBT $T \in \mathcal{U}_n$, with $n \geq 4$, and let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ denote the eigenvalues of $\mathbf{2}^{-\boldsymbol{\tau}}$. Then*

$$-\frac{1}{6} < \lambda_i < \frac{1}{2} \qquad \forall\, i \in [3, n-1].$$

Conjecture 5.1 is supported by exhaustive computations for $n \leq 15$ and by experiments on 500 randomly sampled UBT topologies for each $n \in [16, 250]$. Code and numerical results are available at https://github.com/RobRcx/Spectral-Structure-of-Path-Length-Density-Matrices-of-Unrooted-Binary-Trees.

n	Cospectral pairs	Structures	Percentage	n	Cospectral pairs	Structures	Percentage
14	1	135	0.7407%	21	14	23846	0.0587%
15	1	265	0.3774%	22	18	52233	0.0345%
16	0	552	0.0000%	23	18	114796	0.0157%
17	0	1132	0.0000%	24	46	254371	0.0181%
18	6	2410	0.2490%	25	63	565734	0.0111%
19	4	5098	0.0785%	26	147	1265579	0.0116%
20	6	11020	0.0544%	27	285	2841632	0.0100%

TABLE 1

Cospectral pairs among UBT structures, with relative percentages, for $n \in [14, 27]$. There are no cospectral pairs for $n \in [3, 13]$.

A second question concerns spectral distinguishability, which fits into a broader line of research on related phenomena in trees and similar structures; see, for example, [7] for recent results on strong cospectrality in trees. As discussed in Section 3 and detailed in Appendix B, the spectrum of a PLDM is not, in general, sufficient to distinguish UBT structures, since cospectral pairs already occur for $n = 14$. Consistently with related non-uniqueness phenomena observed for other matrix encodings of trees [14], we expect the number of cospectral pairs of distinct PLDMs in Θ_n to grow unboundedly with n.

At the same time, Table 1 shows that, for $n \in [14, 27]$, the number of cospectral pairs remains very small relative to the total number of UBT structures. Thus, although cospectrality appears already for relatively small values of n, the spectrum still seems to retain substantial discriminating power for moderate values of n, and may therefore serve as a useful approximate invariant in applications. In light of these observations, we are currently investigating the asymptotic behavior of cospectrality in Θ_n in greater detail.

Final remarks. The results of this article establish a coherent connection between the spectrum of a PLDM and the combinatorial structure of the underlying UBT. The fundamental most-balanced UBT and the fundamental caterpillar are spectrally rigid, whereas for more general UBTs only certain subtree configurations can be detected through specific eigenvalues of the associated PLDM. This viewpoint also suggests directions for future research on the spectra of structured graph families, in line with recent work such as [6], which shows that strong structural symmetry can enforce highly constrained spectral behavior.

At the same time, the spectrum is not a complete invariant: as in the case of Laplacian matrices of rooted binary trees, there exist non-isomorphic UBTs in $\mathcal{U}_n$ whose associated PLDMs are cospectral. Thus, the spectrum should be viewed not as a full encoding of a tree, but rather as a compact descriptor that can be combined with additional structural information. In this direction, Theorem 2.1 provides an algebraic characterization of PLDMs through entrywise conditions for $n \geq 3$, while additional characterizations are available for $3 \leq n \leq 11$ [3]. On the spectral side, Theorems 3.8 and 3.12 show that the fundamental most-balanced UBT and the fundamental caterpillar exhibit highly rigid eigenvalue patterns. Taken together, these results suggest that, at least for structured families of UBTs, spectral information combined with consistency conditions such as Kraft's equalities may provide a compact description of the corresponding PLDMs among rational matrices of order n.

This viewpoint is also relevant to optimization. In the *Balanced Minimum Evolu-*

tion Problem (BMEP), given a symmetric real-valued distance matrix $\mathbf{D} = [d_{ij}]_{i,j\in[n]}$, one seeks the UBT $T \in \mathcal{U}_n$ whose associated PLDM $\mathbf{2}^{-\boldsymbol{\tau}} \in \Theta_n$ minimizes

$$L(\mathbf{2}^{-\boldsymbol{\tau}}) = \sum_{i,j\in[n],\, i\neq j} d_{ij}\, 2^{-\tau_{ij}}.$$

Understanding which structural features of a tree are encoded by the spectrum of its PLDM may help identify classes of instances for which optimal solutions are spectrally rigid, or at least stable under perturbations of $\mathbf{D}$. A first indication in this direction is provided by the classical eigenvalue inequalities [12]

$$\sum_{i=1}^{n} \lambda_i(\mathbf{D})\, \lambda_{n+1-i}(\mathbf{2}^{-\boldsymbol{\tau}}) \;\leq\; \mathrm{tr}(\mathbf{D}\, \mathbf{2}^{-\boldsymbol{\tau}}) \;\leq\; \sum_{i=1}^{n} \lambda_i(\mathbf{D})\, \lambda_i(\mathbf{2}^{-\boldsymbol{\tau}}),$$

which suggest that spectral information may help compare candidate PLDMs in ways relevant to the analysis of the BMEP.

Finally, several results in this article appear natural candidates for extension to more general classes of trees; see, for example, Proposition 3.1 and Theorem 3.6. In particular, determining whether the connection between specific eigenvalues and subtree configurations extends beyond the binary case remains an interesting direction for future work.

Acknowledgments. The first author acknowledges support from the Belgian National Fund for Scientific Research (FNRS), via the grant FNRS PDR 40007831, and the Fondation Louvain, via the grant "COALESCENS".

References.

[1] T. Araújo Lima and M. A. M. de Aguiar, *Laplacian matrices for extremely balanced and unbalanced phylogenetic trees*, 2020. arXiv:2008.12866.
[2] A. E. Brouwer and W. H. Haemers, *Spectra of Graphs*, Springer, New York, NY, 2012.
[3] D. Catanzaro, R. Pesenti, and R. Ronco, *Characterizing path-length matrices of unrooted binary trees*, LIDAM Discussion Paper CORE 2025/28, UCLouvain, Center for Operations Research and Econometrics (CORE), 2025. Currently under review in Discrete Optimization.
[4] D. Catanzaro, R. Pesenti, A. Sapucaia, and L. Wolsey, *Optimizing over path-length matrices of unrooted binary trees*, Mathematical Programming, 215 (2026), p. 453–505.
[5] F. R. K. Chung, *Spectral Graph Theory*, vol. 92 of CBMS Regional Conference Series in Mathematics, American Mathematical Society, Providence, RI, 1997.
[6] G. Coutinho, E. Juliano, and T. Jung Spier, *The spectrum of symmetric decorated paths*, Linear Algebra and its Applications, 711 (2025), pp. 17–39.
[7] G. Coutinho, E. Juliano, and T. J. Spier, *Strong cospectrality in trees*, Algebraic Combinatorics, 6 (2023), pp. 955–963.
[8] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Application*, vol. 87 of Pure and Applied Mathematics, Academic Press, New York, NY, 1980.
[9] C. M. da Fonseca and J. Petronilho, *Explicit inverses of some tridiagonal matrices*, Linear Algebra and its Applications, 325 (2001), pp. 7–21.
[10] Z. Du and C. M. da Fonseca, *New families of trees determined by their spectra*, Linear Algebra and its Applications, 722 (2025), pp. 101–113.

[11] G. Fikioris, *Spectral properties of Kac–Murdock–Szegö matrices with a complex parameter*, Linear Algebra and its Applications, 553 (2018), pp. 182–210.
[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 2nd ed., 2013.
[13] M. Kac, W. L. Murdock, and G. Szegö, *On the eigen-values of certain hermitian forms*, Journal of Rational Mechanics and Analysis, 2 (1953), pp. 767–800.
[14] F. A. Matsen and S. N. Evans, *Ubiquity of synonymity: Almost all large binary trees are not uniquely identified by their spectra or their immanantal polynomials*, Algorithms for Molecular Biology, 7 (2012), p. 14.
[15] B. D. McKay, *On the spectral characterisation of trees*, Ars Combinatoria, 3 (1977), pp. 219–232.
[16] OEIS Foundation Inc., *The On-Line Encyclopedia of Integer Sequences*, Published electronically at https://oeis.org, 2026.
[17] D. S. Parker and P. Ram, *The construction of Huffman codes is a submodular ("convex") optimization problem over a lattice of binary trees*, SIAM Journal on Computing, 28 (1999), pp. 1875–1905.
[18] A. J. Schwenk, *Almost all trees are cospectral*, in New Directions in the Theory of Graphs, F. Harary, ed., Academic Press, New York, NY, 1973, pp. 275–307.
[19] L. S. L. Tan, *Explicit inverse of tridiagonal matrix with applications in autoregressive modelling*, IMA Journal of Applied Mathematics, 84 (2019), pp. 679–695.

Appendix A. Proof of Theorem 3.12. Suppose, without loss of generality, that $\mathbf{C}(n)$ has the following form

$$\mathbf{C}(n) = \begin{pmatrix} 0 & 2^{-2} & 2^{-3} & \dots & 2^{2-n} & 2^{1-n} & 2^{1-n} \\ 2^{-2} & 0 & 2^{-3} & \dots & 2^{2-n} & 2^{1-n} & 2^{1-n} \\ 2^{-3} & 2^{-3} & 0 & \dots & 2^{3-n} & 2^{2-n} & 2^{2-n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 2^{2-n} & 2^{2-n} & 2^{3-n} & \dots & 0 & 2^{-3} & 2^{-3} \\ 2^{1-n} & 2^{1-n} & 2^{2-n} & \dots & 2^{-3} & 0 & 2^{-2} \\ 2^{1-n} & 2^{1-n} & 2^{2-n} & \dots & 2^{-3} & 2^{-2} & 0 \end{pmatrix}. \tag{A.1}$$

By Proposition 3.1, the matrix $\mathbf{C}(n)$ has two eigenvalues $\lambda_1 = \lambda_2 = -1/4$. To determine the values of the remaining $n-2$ eigenvalues, we first observe that, by Proposition 3.3, they coincide with the eigenvalues of the matrix $\mathbf{W}$ and that, by Proposition 3.10, they satisfy

$$\lambda_{i+2} = \frac{1}{4}(\mu_i - 1) \qquad \forall\, i \in [n-2], \tag{A.2}$$

where each μ_i, together with its associated normalized eigenvector $\mathbf{x}_i$, forms an eigenpair $(\mu_i, \mathbf{x}_i)$ of the system

$$\mathbf{SVx}_i = \mu_i \mathbf{x}_i, \tag{A.3}$$

with $\mathbf{V} = \mathrm{diag}(2, 1[n-4], 2)$, and $\mathbf{S} \in \mathbb{R}^{(n-2)\times(n-2)}$ a dyadic matrix, whose inverse is the tridiagonal matrix given in (3.19). Now, consider a specific pair $(\mu_i, \mathbf{x}_i)$ for some $i \in [n-2]$. For notational convenience, we drop the index i and rewrite the above system as

$$\mathbf{Vx} = \mu \mathbf{S}^{-1}\mathbf{x},$$

or, equivalently, as

$$\begin{cases} 6x_1 = \mu\left(4x_1 - 2x_2\right), \\ 3x_k = \mu\left(-2x_{k-1} + 5x_k - 2x_{k+1}\right), \qquad k = 2, \dots, n-3, \\ 6x_{n-2} = \mu\left(4x_{n-2} - 2x_{n-3}\right). \end{cases} \tag{A.4}$$

Note that any solution to (A.4) must satisfy $\mu \neq 0$. Indeed, if $\mu = 0$, the system $\mathbf{SVx} = \mathbf{0}$ has only the trivial solution $\mathbf{x} = \mathbf{0}_{n-2}$, which cannot form an eigenpair. The pair $(\mu, \mathbf{x}) = (3, \mathbf{1}_{n-2})$, instead, satisfies (A.4) and forms a valid eigenpair of this system. Substituting $\mu = 3$ into (A.2) yields $\lambda = \frac{1}{4}(\mu - 1) = \frac{1}{2}$, which, by Proposition 3.1, is the largest eigenvalue of $\mathbf{C}(n)$.

To characterize the remaining $n-3$ eigenpairs, we observe that (A.4) is a discrete Sturm–Liouville eigenvalue problem [12] with boundary conditions provided by the first and the third equations in (A.4). Every nontrivial eigenvector that solves (A.4) can be written, up to a scaling, in the following sinusoidal form

$$\mathbf{x} = \big[\sin(k\alpha + \phi)\big]_{k \in [n-2]}, \tag{A.5}$$

where the parameters (α, ϕ), with $0 \leq \alpha < 2\pi$ and $0 \leq \phi < 2\pi$, are determined by the boundary conditions and yield precisely $n-2$ admissible eigenmodes. For example,

$\mathbf{x} = \mathbf{1}_{n-2}$ corresponds to the pair $(\alpha, \phi) = (0, \pi/2)$. Substituting (A.5) into (A.4), yields

$$\begin{cases} \sin(2\alpha + \phi) = \left(2 - \frac{3}{\mu}\right) \sin(\alpha + \phi) \\ \frac{5\mu - 3}{2\mu} \sin(k\alpha + \phi) = \sin((k+1)\alpha + \phi) + \sin((k-1)\alpha + \phi), \qquad k = 2, \ldots, n-3, \\ \sin((n-3)\alpha + \phi) = \left(2 - \frac{3}{\mu}\right) \sin((n-2)\alpha + \phi). \end{cases}$$

Now observe that, since $\mathbf{x} \neq \mathbf{0}_{n-2}$, we must have $\sin(k\alpha + \phi) \neq 0$ for at least one $k \in [2, n-3]$. Moreover, for a fixed $k \in [2, n-2]$, we can exploit elementary trigonometric identities to rewrite the second equation in the following form

$$\frac{5\mu - 3}{2\mu} \sin(k\alpha + \phi) = \sin((k+1)\alpha + \phi) + \sin((k-1)\alpha + \phi) = 2 \sin(k\alpha + \phi) \cos\alpha$$

which, in turn, can be rewritten as

$$\frac{5\mu - 3}{4\mu} = \cos\alpha$$

or equivalently, as

$$\mu = \frac{3}{5 - 4\cos\alpha}.$$

In addition, denoting by

$$\beta = \alpha + \phi$$

and

$$\beta' = (n-2)\alpha + \phi,$$

we can rewrite the boundary conditions in (A.4) as

$$\begin{cases} \sin(\beta + \alpha) = (4\cos\alpha - 3) \sin\beta \\ \sin(\beta' - \alpha) = (4\cos\alpha - 3) \sin\beta' \end{cases}$$

or equivalently, as

$$\begin{cases} \sin\beta\cos\alpha + \sin\alpha\cos\beta = (4\cos\alpha - 3)\sin\beta \\ \sin\beta'\cos\alpha - \sin\alpha\cos\beta' = (4\cos\alpha - 3)\sin\beta'. \end{cases} \tag{A.6}$$

Observe that $\sin\beta = 0$ and $\sin\beta' = 0$ cannot hold at the same time, otherwise $\mathbf{x} = \mathbf{0}_{n-2}$. Moreover, if $\cos\beta = 0$ and/or $\cos\beta' = 0$, then the only solution compatible with (A.6) is $\cos\alpha = 1$, which corresponds, via(A.5), to the pair $(\mu, \mathbf{x}) = (3, \mathbf{1}_{n-2})$ already found above. Hence, to determine the value of the remaining $n-3$ eigenpairs, we may assume that $\sin\beta \neq 0$, $\sin\beta' \neq 0$, $\cos\beta \neq 0$, and $\cos\beta' \neq 0$.

Now, dividing the equations in (A.6) by $\cos\beta$ and $\cos\beta'$, respectively, yields

$$\begin{cases} \tan\beta = -\dfrac{\sin\alpha}{3(1 - \cos\alpha)}, \\ \tan\beta' = \dfrac{\sin\alpha}{3(1 - \cos\alpha)} \end{cases} \tag{A.7}$$

which implies $\tan\beta = -\tan\beta'$. In addition, by elementary trigonometric identities, we have

$$\tan(\beta'-\beta) = \frac{\tan\beta' - \tan\beta}{1+\tan\beta'\tan\beta} = \frac{2\tan\beta'}{1-\tan^2\beta'} = \tan(2\beta').$$

Since $\beta' - \beta = (n-3)\alpha$, we conclude that

$$\tan\big((n-3)\alpha\big) = \tan(2\beta'). \tag{A.8}$$

Now, observe that (A.8) implies either

$$(n-3)\alpha + l\pi = 2\beta', \quad \forall\, l \in \mathbb{R}, \tag{A.9}$$

or

$$(n-3)\alpha + \left(l+\frac{1}{2}\right)\pi = 2\beta', \quad \forall\, l \in \mathbb{R}, \tag{A.10}$$

If (A.9) holds, then

$$\tan\left(\frac{(n-3)\alpha}{2}\right) = \tan(\beta'). \tag{A.11}$$

Alternatively, if (A.10) holds, then

$$\tan\left(\frac{(n-3)\alpha}{2}\right) = -\cot(\beta'). \tag{A.12}$$

Now, by using the second equation in (A.7) and the cotangent half-angle formula, we have that

$$\tan\beta' = \frac{\sin\alpha}{3(1-\cos\alpha)} = \frac{1}{3}\cot\Big(\frac{\alpha}{2}\Big), \tag{A.13}$$

hence (A.11) and (A.12) are equivalent to (3.22a) and (3.22b), respectively.

In light of this equivalence, to complete the proof we just need to show that (3.22) yields exactly $n-3$ distinct solutions. To this end, consider equation (3.22a). The function $\alpha \mapsto \tan\Big(\frac{(n-3)\alpha}{2}\Big)$ has period $\frac{2\pi}{n-3}$; on each branch of its domain it is strictly increasing, and on each branch where it is positive it takes values from 0 to arbitrarily large. On the other hand, $\alpha \mapsto \frac{1}{3}\cot\big(\frac{\alpha}{2}\big)$ is strictly decreasing on $(0,\pi)$, with values from arbitrarily large to 0. Hence (3.22a) has exactly one solution on each positive branch of $\tan\Big(\frac{(n-3)\alpha}{2}\Big)$ contained in $(0,\pi)$. In addition, $\alpha = \pi$ is also a solution when $n-3$ is even. Thus, (3.22a) has exactly $\left\lceil\frac{n-2}{2}\right\rceil$ distinct solutions in $0 < \alpha \le \pi$. With a similar argument, we conclude that (3.22b) has exactly $\left\lfloor\frac{n-4}{2}\right\rfloor$ distinct solutions in $0 < \alpha < \pi$, since it has exactly one solution on each negative branch of $\tan\Big(\frac{(n-3)\alpha}{2}\Big)$ contained in $(0,\pi)$, except for the last one when $n-3$ is even. Thus, (3.22) yields exactly

$$\left\lceil\frac{n-2}{2}\right\rceil + \left\lfloor\frac{n-4}{2}\right\rfloor = n-3$$

distinct solutions, as claimed, and the statement follows.

$\square$

Appendix B. The spectra of PLDMs do not uniquely determine the structure of UBTs.

As discussed in the introduction, spectral studies of adjacency matrices [18] and path-length matrices [15] of unweighted trees, as well as Laplacian matrices of rooted binary trees [1, 14], have shown that the spectral information contained in these matrices is insufficient to uniquely identify the structure of the encoded trees. In particular, matrices of distinct trees may share the same spectrum. This negative result also extends to PLDMs of UBTs, as illustrated in the following example.

Example B.1. Consider the following distinct PLDMs in $\mathcal{U}_{14}$:

$$\mathbf{2}^{-\boldsymbol{\tau}} = \begin{bmatrix}
0 & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-5} & 2^{-5} & 2^{-3} & 2^{-2} & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-5} & 2^{-8} \\
2^{-8} & 0 & 2^{-2} & 2^{-4} & 2^{-6} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 2^{-4} & 2^{-6} & 2^{-6} & 2^{-7} & 2^{-6} \\
2^{-8} & 2^{-2} & 0 & 2^{-4} & 2^{-6} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 2^{-4} & 2^{-6} & 2^{-6} & 2^{-7} & 2^{-6} \\
2^{-8} & 2^{-4} & 2^{-4} & 0 & 2^{-6} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 2^{-2} & 2^{-6} & 2^{-6} & 2^{-7} & 2^{-6} \\
2^{-8} & 2^{-6} & 2^{-6} & 2^{-6} & 0 & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 2^{-6} & 2^{-2} & 2^{-4} & 2^{-7} & 2^{-4} \\
2^{-5} & 2^{-5} & 2^{-5} & 2^{-5} & 2^{-5} & 0 & 2^{-4} & 2^{-4} & 2^{-5} & 2^{-5} & 2^{-5} & 2^{-5} & 2^{-4} & 2^{-5} \\
2^{-5} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-4} & 0 & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-2} & 2^{-7} \\
2^{-3} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-4} & 2^{-4} & 0 & 2^{-3} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-4} & 2^{-7} \\
2^{-2} & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-5} & 2^{-5} & 2^{-3} & 0 & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-5} & 2^{-8} \\
2^{-8} & 2^{-4} & 2^{-4} & 2^{-2} & 2^{-6} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 0 & 2^{-6} & 2^{-6} & 2^{-7} & 2^{-6} \\
2^{-8} & 2^{-6} & 2^{-6} & 2^{-6} & 2^{-2} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 2^{-6} & 0 & 2^{-4} & 2^{-7} & 2^{-4} \\
2^{-8} & 2^{-6} & 2^{-6} & 2^{-6} & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 2^{-6} & 2^{-4} & 0 & 2^{-7} & 2^{-2} \\
2^{-5} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-4} & 2^{-2} & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-7} & 0 & 2^{-7} \\
2^{-8} & 2^{-6} & 2^{-6} & 2^{-6} & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-8} & 2^{-6} & 2^{-4} & 2^{-2} & 2^{-7} & 0
\end{bmatrix} \tag{B.1}$$

and

$$\mathbf{2}^{-\hat{\boldsymbol{\tau}}} = \begin{bmatrix}
0 & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-4} & 2^{-4} & 2^{-2} & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-6} & 2^{-4} \\
2^{-8} & 0 & 2^{-2} & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} & 2^{-8} & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} \\
2^{-8} & 2^{-2} & 0 & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} & 2^{-8} & 2^{-4} & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} \\
2^{-8} & 2^{-4} & 2^{-4} & 0 & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} & 2^{-8} & 2^{-2} & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} \\
2^{-7} & 2^{-5} & 2^{-5} & 2^{-5} & 0 & 2^{-6} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-5} & 2^{-2} & 2^{-6} & 2^{-5} & 2^{-7} \\
2^{-7} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-6} & 0 & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-6} & 2^{-2} & 2^{-3} & 2^{-7} \\
2^{-4} & 2^{-6} & 2^{-6} & 2^{-6} & 2^{-5} & 2^{-5} & 0 & 2^{-4} & 2^{-4} & 2^{-6} & 2^{-5} & 2^{-5} & 2^{-4} & 2^{-4} \\
2^{-4} & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-4} & 0 & 2^{-4} & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-6} & 2^{-2} \\
2^{-2} & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-4} & 2^{-4} & 0 & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-6} & 2^{-4} \\
2^{-8} & 2^{-4} & 2^{-4} & 2^{-2} & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} & 2^{-8} & 0 & 2^{-5} & 2^{-7} & 2^{-6} & 2^{-8} \\
2^{-7} & 2^{-5} & 2^{-5} & 2^{-5} & 2^{-2} & 2^{-6} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-5} & 0 & 2^{-6} & 2^{-5} & 2^{-7} \\
2^{-7} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-6} & 2^{-2} & 2^{-5} & 2^{-7} & 2^{-7} & 2^{-7} & 2^{-6} & 0 & 2^{-3} & 2^{-7} \\
2^{-6} & 2^{-6} & 2^{-6} & 2^{-6} & 2^{-5} & 2^{-3} & 2^{-4} & 2^{-6} & 2^{-6} & 2^{-6} & 2^{-5} & 2^{-3} & 0 & 2^{-6} \\
2^{-4} & 2^{-8} & 2^{-8} & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-4} & 2^{-2} & 2^{-4} & 2^{-8} & 2^{-7} & 2^{-7} & 2^{-6} & 0
\end{bmatrix}, \tag{B.2}$$

and observe that they encode the distinct UBTs shown on the left and the right of Figure 8, respectively. Note, however, that $\mathbf{2}^{-\boldsymbol{\tau}}$ and $\mathbf{2}^{-\hat{\boldsymbol{\tau}}}$ share the same spectrum (rounded to 12 decimal digits):

$$\begin{aligned}
\{\, & 0.500000000000_{[1]}, \quad 0.382324754518_{[1]}, \quad 0.312500000000_{[1]}, \quad 0.199404451781_{[1]}, \\
& 0.125000000000_{[2]}, \ -0.037509087205_{[1]}, \ -0.106720119095_{[1]}, \ -0.250000000000_{[6]} \,\}
\end{aligned}$$

where subscripts denote multiplicities. The information contained in the spectra of these PLDMs is therefore insufficient to distinguish the structure of the encoded UBTs.

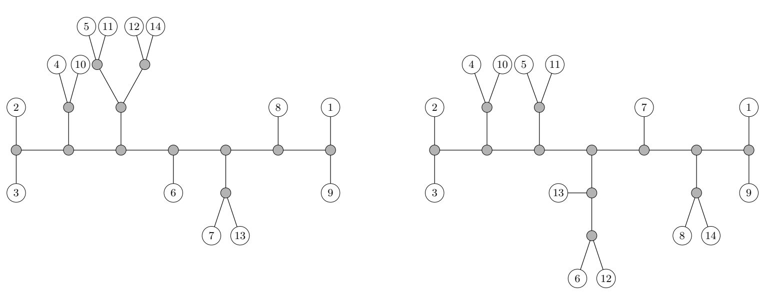


FIG. 8. *The UBTs associated with the PLDMs* (B.1) *(on the left) and* (B.2) *(on the right).*