The Chromatic Symmetric Function for Unicyclic Graphs

Aram Bingham[a] Lisa Johnston[b] Colin Lawson[c] Rosa Orellana[d]
Jianping Pan[e] Chelsea Sato[f]

Submitted: Jan 20, 2026; Accepted: May 28, 2026; Published: Jul 17, 2026

Abstract

Motivated by the question of which structural properties of a graph can be recovered from the chromatic symmetric function (CSF), we study the CSF of connected unicyclic graphs. While it is known that there can be non-isomorphic unicyclic graphs with the same CSF, we find experimentally that such examples are rare for graphs with up to 17 vertices. In fact, in many cases we can recover data such as the number of leaves, number of internal edges, cycle size, and number of attached non-trivial trees, by extending known results for trees to unicyclic graphs. These results are obtained by analyzing the CSF of a connected unicyclic graph in the *star-basis* using the deletion-near-contraction (DNC) relation developed by Aliste-Prieto, Orellana and Zamora, and computing the "leading" partition, its coefficient, as well as coefficients indexed by hook partitions. We also give explicit formulas for star-expansions of several classes of graphs, developing methods for extracting coefficients using structural properties of the graph.

Mathematics Subject Classifications: 05E05, 05C60, 05C38

1 Introduction

The chromatic symmetric function $\mathbf{X}_G$ of a graph G was introduced in [14] by Stanley as a symmetric function generalization of the chromatic polynomial of a graph. Specifically,

[a] Departamento de Matemáticas, Universidad de Chile, 3425 Las Palmeras, Ñuñoa, Santiago, Chile (`aram@matmor.unam.mx`.)
[b] Department of Mathematics, University of California, One Shields Avenue, Davis, CA, U.S.A. (`lisjohnston@ucdavis.edu`).
[c] Department of Mathematics & Statistics, Stephen F. Austin State University, Nacogdoches, TX, USA (`Colin.Lawson@sfasu.edu`).
[d] Mathematics Department, Dartmouth College, Hanover, NH, U.S.A. (`Rosa.C.Orellana@dartmouth.edu`).
[e] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, U.S.A. (`jianping.pan@asu.edu`).
[f] Department of Mathematics, Syracuse University, Syracuse NY, U.S.A.
Current address: Department of Mathematics, Willamette University, Salem, OR, U.S.A. (`cmsato@willamette.edu`).

for a graph G with vertex set $V = \{v_1, \dots, v_n\}$, $\mathbf{X}_G$ is defined, for commuting variables $x_1, x_2, \dots$, by

$$\mathbf{X}_G = \sum_{\kappa} x_{\kappa(v_1)} x_{\kappa(v_2)} \dots x_{\kappa(v_n)},$$

where $\kappa : V \to \mathbb{N}$ ranges over all proper colorings of the vertices of G. Setting the first k variables to 1 and all other variables to zero indeed recovers the value $\chi_G(k)$ of the chromatic polynomial of G. A significant body of research on $\mathbf{X}_G$ has focused on Stanley's "tree isomorphism problem", which asks whether the chromatic symmetric function distinguishes non-isomorphic trees. While this question remains open in general, the chromatic symmetric function has been shown to distinguish various subclasses of trees (see, e.g., [2, 5, 8, 10, 11, 18]), as well as for all trees up to 29 vertices [9].

Many key structural properties of a graph can be recovered from $\mathbf{X}_G$ by examining its coefficients when expanded with respect to various bases of symmetric functions. For example, expanding $\mathbf{X}_G$ in the power sum symmetric function basis reveals structural properties such as the number of vertices, edges, girth and number of connected components [11], the degree sequence and number of leaves for trees [11], and the number of triangles in G [12]. In particular, $\mathbf{X}_G$ can tell us whether G is a connected unicyclic graph [11]. Recently, the chromatic symmetric function has been studied [1, 6] in terms of the *star-basis*, $\{\mathfrak{st}_\lambda \mid \lambda \vdash n\}$, where $\lambda \vdash n$ means λ is an integer partition of n.

In the case that G is a tree, the authors of [6, 7] investigate the expansion $\mathbf{X}_G$ in the star-basis, $\mathbf{X}_G = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$. They focus on the *leading partition* λ_{lead} of $\mathbf{X}_G$, defined as the smallest partition in lexicographic order appearing in $\mathbf{X}_G$ with non-zero coefficient. The authors additionally provided formulas for the leading coefficient $c_{\lambda_{\text{lead}}}$, which encodes the degrees of the "deep" vertices in G, and for the hook coefficients which determine the number of non-leaf ("internal") edges.

In this paper, we analyze the chromatic symmetric function of connected unicyclic graphs in the star-basis. We determine some structural features of these graphs from specific coefficients in their star-basis expansion. Our first main result, Theorem 10, provides an explicit formula for the coefficients c_λ in the star-basis expansion when λ is a hook partition. As an immediate application, we show in Theorem 11 that the size of the cycle in any unicyclic graph can be directly determined from the coefficient $c_{(n)}$. Next, in Theorem 32, we determine the leading partition λ_{lead} for unicyclic graphs. Using λ_{lead}, we show in Theorem 43 how to compute the number of leaves in the graph. We also give formulas for the leading coefficient $c_{\lambda_{\text{lead}}}$ for any unicyclic graph.

With our results on the hook coefficients, leading partition, and the leading coefficient, we distinguish several classes of unicyclic graphs based on their CSF. In Theorem 12, we show that unicyclic graphs with different cycle sizes must have differing CSFs. Finally, in Theorem 44 we prove that *cuttlefish* graphs, which are a certain subclass of *squids*, are distinguished from among all connected unicyclic graphs.

We also present computational data and examples of non-isomorphic graphs with a 4-cycle that share the same CSF in Figures 15 and 16. This demonstrates that for bipartite graphs, the CSF alone does not necessarily distinguish non-isomorphic graphs. As a result, we establish that the tree isomorphism problem is independent of the bipartite property in trees.

Previous work has been done on computing the chromatic symmetric function of graphs with cycles. For example, see [4, 11, 12, 17]. We note that unicyclic graphs are not distinguished by their chromatic symmetric function [12, 14]. However, by analyzing the coefficients and partitions appearing in $\mathbf{X}_G$, we gain insight into when two unicyclic graphs can or cannot be isomorphic.

In Section 7, we give an explicit formula for the coefficient $c_{(n)}$ in $\mathbf{X}_G$ when G is a connected bicyclic graph. This coefficient depends only on the sizes of the cycles in the graph and the number of edges shared by the two cycles. Additionally, we prove several results involving the sums of coefficients in $\mathbf{X}_G$ indexed by partitions of the same length, and use these results to prove that we can recover the sizes of the cycles in connected bicyclic graphs.

This paper is organized as follows. In Section 2, we provide the necessary background. In Section 3, we describe the coefficients for hook partitions. In Section 4, we give formulas for the star-expansion of special families of graphs. In Sections 5 and 6, we present results on the leading partition and leading coefficient. We use the result in Section 5 to prove that cuttlefish are distinguished (and reconstructible) from among all connected unicyclic graphs. In Section 7, we present results on bicyclic graphs. Finally in Section 8, we give examples and data.

Acknowledgments

This material is based on work supported by the National Science Foundation under Grant Number DMS-1916439 while the authors were participating in the Mathematics Research Communities (MRC) 2024 Summer Conference at Beaver Hollow Conference Center in Java Center, New York. We extend our thanks to the American Mathematical Society for their support. A. Bingham was partially supported by the grant DST/INT/RUS/RSF/P-41/2021 from the Department of Science & Technology, Govt. of India, and FONDECYT-ANID grant 3250472. L. Johnston was partially supported by NSF grant DMS-2053350. R. Orellana was partially supported by NSF grant DMS-2153998. This material is also based upon work supported by the National Science Foundation, while A. Bingham, R. Orellana and J. Pan were in residence at the ICERM semester program "Categorification and Computation in Algebraic Combinatorics" in Fall 2025.

2 Definitions and preliminaries

2.1 Graph theory

In this paper, we utilize standard definitions and fundamental concepts from graph theory. For an overview of graph theory background, refer to, for example, [16]. For convenience and to establish the necessary notation for this paper, we recall some of the basic concepts here.

Graph basics

A *graph* G is an ordered pair (V, E), where $V = V(G)$ is the set of vertices of G and $E = E(G)$ is its set of edges. The number of vertices in V is the *order* of G, denoted by $|V|$. The *degree* of a vertex v is the number of edges incident to it, and is denoted by $\deg(v)$. The *degree sequence* of a graph is the list of the degrees of all of its vertices, arranged in non-increasing order.

A graph G is said to be *simple* if G contains no loops or multiple edges between two vertices. All graphs in this paper are assumed to be simple, unless otherwise stated. A *path* in a graph G is a sequence of distinct vertices $v_1, \ldots, v_n$ such that $v_i v_{i+1}$ is an edge in $E(G)$ for each $1 \leqslant i \leqslant n-1$. A path on n vertices in a graph is denoted by P_n. A graph G is *connected* provided that there is a path between any two vertices in $V(G)$. A *cycle* is a graph with an equal number of vertices and edges whose vertices can be placed around a circle so that two vertices are adjacent if and only if they appear consecutively along the circle.

We denote by C_n a cycle on n vertices in a graph. A graph is a *tree* if there is a unique path between any two vertices. Equivalently, a tree is a connected graph that contains no cycles. A *rooted tree* is a tree with a designated root vertex. We call a tree comprised of a single vertex a *trivial tree*, while any tree with more than one vertex is a *non-trivial tree*. A graph such that each of its connected components is a tree is called a *forest*.

A *unicyclic graph* is a graph that contains exactly one cycle. Connected unicyclic graphs can also be characterized as simple graphs whose number of edges equals the number of vertices. In this paper, we consider connected unicyclic graphs G constructed from c rooted trees $T_1, \ldots, T_c$ by adding edges between the roots of consecutive trees T_i and T_{i+1} for $1 \leqslant i \leqslant c-1$, and connecting the root of T_c with the root of T_1, creating a cycle of size c in the graph G with vertices corresponding to the roots of the trees. We also call such a graph a *c-unicyclic* graph. We use r to denote the number of non-trivial rooted trees. See Figure 1 for an illustration.

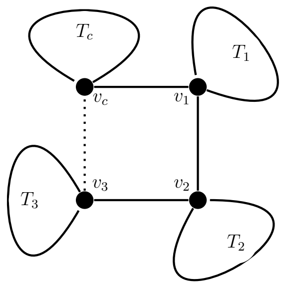

Figure 1: A connected unicyclic graph with a c-cycle and c (potentially trivial) rooted trees.

For example, the graph on the left of Figure 2 is a unicyclic graph constructed from 4 rooted trees with roots v_1, v_2, v_3 and v_4 where v_4 is the root of a trivial tree. In this case, we have $r = 3$.

Vertex and edge types

We now define specific types of vertices and edges in a graph that will be used throughout the paper. An *internal vertex* is a vertex with degree at least two. A *leaf* is a vertex of degree one. An internal vertex is a *deep vertex* if it has no adjacent leaves. An *internal* *edge* is an edge such that both of its endpoints are internal vertices, while a *leaf-edge* is an edge such that at least one of its endpoints is a leaf. For a graph G, we denote by $I(G)$ the set of all internal edges of G. The connected components of the graph $G \setminus I(G)$ are referred to as the *leaf components*. For a connected unicyclic graph G with c rooted trees (as constructed above), we say vertex v is a *sprout* if it is a deep vertex on the c-cycle that is of degree at least three.

Example 1. Let G be the graph in the left of Figure 2 on 14 vertices. The internal vertices are $v_1, v_2, v_3, v_4, v_5, v_{11}, v_{12}$, and the deep vertices are v_3, v_4, v_{11}. The only sprout is v_3.

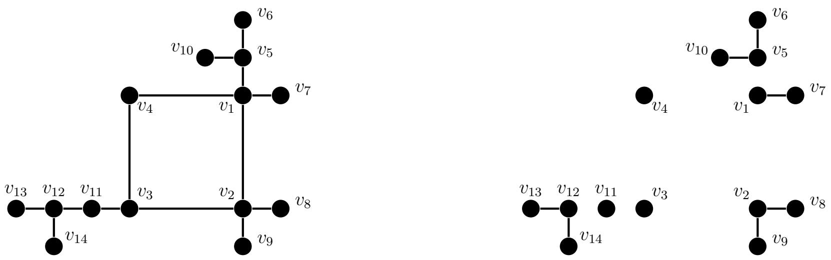

Figure 2: Left: A unicyclic graph G on 14 vertices; right: $G \setminus I(G)$.

2.2 Symmetric functions

A *partition* is a sequence $\lambda = (\lambda_1, \lambda_2 \dots, \lambda_k)$ of positive integers such that $\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_k$. We say that λ is a partition of n, denoted $\lambda \vdash n$ or $|\lambda| = n$, if $\sum_i \lambda_i = n$. Each λ_i is a *part*, and the number of parts is the *length* of λ, denoted $\ell(\lambda)$. If λ has exactly m_i parts equal to i for $1 \leqslant i \leqslant n$, we may also denote λ by $\lambda = (1^{m_1} 2^{m_2} \dots n^{m_n})$. A *hook partition* of n is a partition of the form $\lambda = (n - m_1, 1^{m_1})$, including the case where $m_1 = 0$. The total ordering on partitions of n that we use throughout the paper is the *lexicographic order*: $\mu \leqslant \lambda$ if $\mu = \lambda$ or $\mu_i = \lambda_i$ for $1 \leqslant i < j$ and $\mu_j < \lambda_j$ for some $1 \leqslant j \leqslant \ell(\mu)$.

Algebra of symmetric functions

A formal power series f in a countably infinite set of commuting variables $x_1, x_2, \dots$ is called a *symmetric function* if it has bounded degree and is invariant under the action of any permutation w of the positive integers on the indices of the variables, that is

$$f(x_1, x_2, \dots) = f(x_{w(1)}, x_{w(2)}, \dots).$$

Let Λ^n be the vector space of all homogeneous symmetric functions of degree n. The *r-th elementary symmetric function* e_r is defined by

$$e_r := \sum_{i_1 < i_2 < \cdots < i_r} x_{i_1} x_{i_2} \cdots x_{i_r}.$$

For a partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$, the *elementary symmetric function* corresponding to λ is given by

$$e_\lambda := e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_k}$$

forms a basis of Λ^n, called the *elementary basis*. Using the elementary symmetric functions, we can define the graded $\mathbb{Q}$-algebra of symmetric functions, denoted by Λ, as

$$\Lambda = \Lambda^0 \oplus \Lambda^1 \oplus \Lambda^2 \oplus \cdots\,, \text{ where } \Lambda^n := \mathrm{Span}_{\mathbb{Q}}\{e_\lambda \mid \lambda \vdash n\} \text{ for } n \geqslant 1\,,$$

and $\Lambda^0 := \mathbb{Q}$.

The chromatic symmetric function

Let G be a finite graph with vertex set $V = \{v_1, \dots, v_n\}$. A *proper coloring* of G is a function $\kappa : V \to \mathbb{N}$ such that $\kappa(v_i) \neq \kappa(v_j)$ whenever v_iv_j is an edge of G with $1 \leqslant i < j \leqslant n$. The *chromatic symmetric function* (CSF) [14] of G is defined by

$$\mathbf{X}_G = \sum_{\kappa} x_{\kappa(v_1)} x_{\kappa(v_2)} \dots x_{\kappa(v_n)},$$

where the sum is over all proper colorings κ of G. The function $\mathbf{X}_G$ is a homogeneous symmetric function of degree n, where n is the order of G. In the case that G is a disjoint union of two subgraphs, $G = H_1 \sqcup H_2$, the chromatic symmetric function of G is $\mathbf{X}_G = \mathbf{X}_{H_1} \cdot \mathbf{X}_{H_2}$. In their work [3], Cho and van Willigenburg give new bases of symmetric functions using the chromatic symmetric function and they proved the following result.

Theorem 2 ([3, Theorem 5])**.** *For any positive integer k, let G_k denote a connected graph with k vertices and let $\{G_k\}_{k\geqslant 1}$ be a family of such graphs. Given a partition $\lambda \vdash n$, of length ℓ, define $G_\lambda = G_{\lambda_1} \sqcup G_{\lambda_2} \sqcup \dots \sqcup G_{\lambda_\ell}$. Then $\{\mathbf{X}_{G_\lambda} \mid \lambda \vdash n\}$ is a basis for Λ^n.*

The star-basis

In this paper, we focus on the star-basis of symmetric functions, which is constructed from the chromatic symmetric functions of star graphs. A *star graph* on k vertices, denoted by St_k, is a tree with $k-1$ vertices of degree 1 and one central vertex of degree $k-1$. We denote the chromatic symmetric function of St_k by $\mathfrak{st}_k := \mathbf{X}_{\mathsf{St}_k}$. For a partition $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_\ell)$, the chromatic symmetric function of the star forest $\mathsf{St}_{\lambda_1} \sqcup \mathsf{St}_{\lambda_2} \sqcup \dots \sqcup \mathsf{St}_{\lambda_\ell}$ is denoted by

$$\mathfrak{st}_\lambda := \mathfrak{st}_{\lambda_1} \mathfrak{st}_{\lambda_2} \cdots \mathfrak{st}_{\lambda_\ell}\,.$$

By Theorem 2, the set $\{\mathfrak{st}_\lambda \mid \lambda \vdash n\}$ is a basis, called the *star-basis*, for the space Λ^n. Hence, we may express $\mathbf{X}_G$ as:

$$\mathbf{X}_G = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda\,.$$

Definition 3. Given a graph G, we define the *leading partition* to be the *smallest* partition in lexicographic order such that $c_{\lambda_{\mathrm{lead}}} \neq 0$, where $\mathbf{X}_G = \sum_\lambda c_\lambda \mathfrak{st}_\lambda$ is the star-expansion of G, which will be denoted $\lambda_{\mathrm{lead}}(\mathbf{X}_G)$. When the context is clear, we will use λ_{lead} for short.

An algorithm is given in [1] for writing the chromatic symmetric function of a graph in the star-basis; this involves the deletion-near-contraction relations, which we recall next.

Given two sequences $a = (a_1, \dots, a_p)$ and $b = (b_1, \dots, b_q)$, let $(a, b) = a \cdot b = (a_1, \dots, a_p, b_1, \dots, b_q)$. Let sort() be the function that orders a sequence into a partition. The next lemma follows from the definitions of leading partition and lexicographic order. We will use it in the proofs in Sections 5 and 6.

Lemma 4. *Let H_1 and H_2 be two graphs. Then we have $\lambda_{\text{lead}}(\mathbf{X}_{H_1 \sqcup H_2}) = \text{sort}(\lambda_{\text{lead}}(\mathbf{X}_{H_1}) \cdot \lambda_{\text{lead}}(\mathbf{X}_{H_2}))$.*

2.3 Deletion-near-contraction relation

Denote by $G \setminus e$ the graph obtained from G by deleting edge e. The *leaf-contraction* operation is denoted $G \odot e$ and yields the graph obtained by contracting edge e and attaching a new leaf to the vertex that results from this contraction, connected by a new leaf-edge ℓ_e. The *dot-contraction* $(G \odot e) \setminus \ell_e$ amounts to contracting edge e and then adding an isolated vertex.

We note that if G contains a 3-cycle and e is an edge on the cycle, then $G \odot e$ and $(G \odot e) \setminus \ell_e$ will produce graphs with multiple edges. To ensure that all graphs are simple, we will delete any duplicate edges in these cases.

Proposition 5 ([1])**.** *[The deletion-near-contraction (DNC) relation] Let G be a simple graph and e be any edge in G. Then*

$$\mathbf{X}_G = \mathbf{X}_{G \setminus e} - \mathbf{X}_{(G \odot e) \setminus \ell_e} + \mathbf{X}_{G \odot e}.$$

Observe that if e is a leaf-edge, then the DNC relation does not provide any new information for $\mathbf{X}_G$, as $G \setminus e \cong (G \odot e) \setminus \ell_e$ and $G \odot e \cong G$. Thus, we will only apply the DNC relation to internal edges of G. Furthermore if e is an internal edge, then the graphs $G \setminus e$, $(G \odot e) \setminus \ell_e$, and $G \odot e$ each have fewer internal edges than G. Note that a simple graph with no internal edges is a star forest. Therefore, by repeatedly applying the DNC relation to internal edges, we can express $\mathbf{X}_G$ as a linear combination of $\mathfrak{st}_\lambda$, where λ is a partition of the number of vertices of G.

We now recall the star-expansion algorithm introduced in [1], which formalizes the recursive construction of $\mathbf{X}_G$ in the star-basis described above.

Algorithm 1: Star-expansion

Input: A simple graph G.
Initialization: Let $\mathcal{T}$ be a tree with root labeled G and no edges.
Iteration: If H is a leaf $\mathcal{T}$ and H has some internal edge e, then extend $\mathcal{T}$ by adding three children to H labeled $H \setminus e, (H \odot e) \setminus \ell_e$, and $H \odot e$. Label the three new edges with $+$ or $-$ according to the corresponding coefficient in the DNC relation. The algorithm terminates when all leaves in $\mathcal{T}$ contain no internal edges.
Output: A rooted tree $\mathcal{T}(G)$ with leaves labeled by star forests.

We call the output of the star-expansion algorithm the *DNC-tree*. Using the DNC-tree, the authors in [1] demonstrated how to write the star-expansion of $\mathbf{X}_G$. We first introduce some notation. Let $\iota(H)$ and $\iota(G)$ denote the number of isolated vertices of the graphs H and G, respectively. Let $\lambda(H)$ be the partition whose parts correspond to the orders of the connected components of H.

Theorem 6 ([1])**.** *For every simple graph G, let $\mathcal{T}(G)$ be a DNC-tree obtained from the star-expansion algorithm, and let $L(\mathcal{T}(G))$ be the multiset of leaf labels of $\mathcal{T}(G)$. Then*

$$\mathbf{X}_G = \sum_{H \in L(\mathcal{T}(G))} (-1)^{\iota(H) - \iota(G)} \mathfrak{st}_{\lambda(H)}.$$

In addition, no cancellations occur in the computation; that is, for any partition λ, all terms $\mathfrak{st}_\lambda$ appear with the same sign.

Example 7. We give an example of the star-expansion algorithm in Figure 3 for the graph G at the top of the DNC-tree. The red edges indicate to which edge the DNC relation is applied. The resulting CSF in the star-basis is

$$\mathbf{X}_G = 2\mathfrak{st}_{(4)} - 2\mathfrak{st}_{(3,1)} + \mathfrak{st}_{(2,2)}.$$

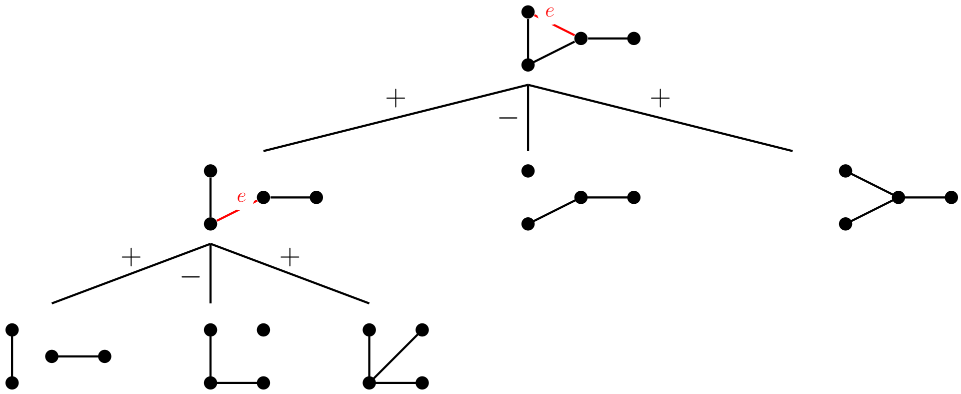

Figure 3: The DNC-tree $\mathcal{T}(G)$.

3 Coefficients for hook partitions

We give the coefficients for hook partitions for all connected unicyclic graphs. First, we give a known fact about the coefficients of hook partitions for trees.

Proposition 8 ([6], Proposition 3.8)**.** *Let T be an n-vertex tree and k be the number of internal edges of T. If $\mathbf{X}_T = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$, then*

$$c_{(n-m_1,1^{m_1})} = (-1)^{m_1}\binom{k}{m_1}.$$

Note that a unicyclic graph with cycle size c always has at least c internal edges. The following lemma is the base case for Theorem 10, which is proved by induction on the size of the cycle.

Lemma 9. *Let G be a connected unicyclic graph on n vertices with one 3-cycle, let $\lambda = (n - m_1, 1^{m_1})$ be a hook partition (note $m_1 \neq n-1$), and let k be the number of internal edges of G. Let $0 \leqslant r \leqslant 3$ be the number of (non-trivial) rooted trees on the 3-cycle. If $\mathbf{X}_G = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$ then*

$$c_{(n-m_1,1^{m_1})} = (-1)^{m_1}\left[(r-1)\binom{k-2}{m_1-1} + 2\binom{k-2}{m_1}\right].$$

Proof. We prove each case for r separately.

0. Observe that when $r = 0$, the graph G is the cycle graph on 3 vertices, C_3, and $\mathbf{X}_{C_3} = 2\mathfrak{st}_{(3)} - 1\mathfrak{st}_{(2,1)}$. This agrees with the formula since $k = 3$.

1. Suppose $r = 1$, and let e be an edge on the cycle that is adjacent to the single rooted tree. The graph $G \setminus e$ is a tree with $k-2$ internal edges, so the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{G\setminus e}$ is $(-1)^{m_1}\binom{k-2}{m_1}$ by Proposition 8. The graph $(G \odot e) \setminus \ell_e$ is the union of an isolated vertex and a tree with $k-3$ internal edges, so the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ is $(-1)^{m_1-1}\binom{k-3}{m_1-1}$. The graph $G \odot e$ is a tree with $k-3$ internal edges, so the $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)}$ is $(-1)^{m_1}\binom{k-3}{m_1}$. Theorem 5 and Pascal's rule imply that in $\mathbf{X}_G$, $c_{(n-m_1,1^{m_1})} = (-1)^{m_1}\left[\binom{k-2}{m_1} + \binom{k-3}{m_1-1} + \binom{k-3}{m_1}\right] = (-1)^{m_1}\left[\binom{k-2}{m_1} + \binom{k-2}{m_1}\right] = (-1)^{m_1}\left[2\binom{k-2}{m_1}\right]$.

2. Suppose $r = 2$, and let e be an edge on the cycle that is adjacent to only a single rooted tree. Following a similar argument as above, the graph $G \setminus e$ is a tree with $k-2$ internal edges, so the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{G\setminus e}$ is $(-1)^{m_1}\binom{k-2}{m_1}$. The graph $(G \odot e) \setminus \ell_e$ is the union of an isolated vertex and a tree with $k-2$ internal edges, so $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ is $(-1)^{m_1-1}\binom{k-2}{m_1-1}$. The graph $G \odot e$ is a tree with $k-2$ internal edges, so $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)}$ is $(-1)^{m_1}\binom{k-2}{m_1}$. Thus in $\mathbf{X}_G$, $c_{(n-m_1,1^{m_1})} = (-1)^{m_1}\left[\binom{k-2}{m_1} + \binom{k-2}{m_1-1} + \binom{k-2}{m_1}\right] = (-1)^{m_1}\left[\binom{k-2}{m_1-1} + 2\binom{k-2}{m_1}\right]$.

3. Let $r = 3$, and let e be any edge on the cycle. The graph $G \setminus e$ is a tree with $k-1$ internal edges, so the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{G\setminus e}$ is $(-1)^{m_1}\binom{k-1}{m_1}$. The graph $(G \odot e) \setminus \ell_e$ is the union of an isolated vertex and a tree with $k-2$ internal edges, so $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ is $(-1)^{m_1-1}\binom{k-2}{m_1-1}$. The graph $G \odot e$ is a tree with $k-2$ internal edges, so $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)}$ is $(-1)^{m_1}\binom{k-2}{m_1}$. Thus in $\mathbf{X}_G$,
$$\begin{aligned} c_{(n-m_1,1^{m_1})} =&(-1)^{m_1}\left[\binom{k-1}{m_1} + \binom{k-2}{m_1-1} + \binom{k-2}{m_1}\right] \\ =&(-1)^{m_1}\left[2\binom{k-2}{m_1-1} + 2\binom{k-2}{m_1}\right]. \end{aligned}$$

□

The following theorem, which we will prove by induction on the cycle size, covers the general case of hook coefficients for connected unicyclic graphs, where Theorem 9 is the base case.

Theorem 10. *Let G be a connected unicyclic graph on n vertices with a single cycle of size $c \geqslant 3$, let $\lambda = (n - m_1, 1^{m_1})$ be a hook partition, k be the number of internal edges of G, and $r \geqslant 0$ be the number of non-trivial rooted trees attached to the cycle in G. If $\mathbf{X}_G = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$, then*
$$c_{(n-m_1,1^{m_1})} = (-1)^{m_1}\left[(r-1)\binom{k-2}{m_1-1} + (c-1)\binom{k-2}{m_1}\right]. \tag{1}$$

Proof. We will split into cases based on the number of rooted trees.

Case 1 ($r \geqslant 1$): We proceed by inducting on the size of the cycle c. The base case when $c = 3$ is Lemma 9. Now suppose that the claim is true for any unicyclic graph with a cycle of size $c = z \geqslant 3$. Let G be a connected graph on n vertices with a single cycle of size $z+1$, $r \geqslant 1$ rooted trees, and k internal edges. We show that the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_G$ is $(-1)^{m_1}\left[(r-1)\binom{k-2}{m_1-1} + z\binom{k-2}{m_1}\right]$. Since $r \geqslant 1$ and the rooted trees can be placed anywhere on the cycle, there is either an edge on the cycle that is adjacent to only a single rooted tree, or there exists no such edge and hence every edge of the cycle is adjacent to two rooted trees. We examine each case separately.

Case 1a: Suppose e is an edge on the cycle that is adjacent to only a single rooted tree. $G \setminus e$ is a tree with $k-2$ internal edges, so the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{G\setminus e}$ is $(-1)^{m_1}\binom{k-2}{m_1}$. Next, $(G \odot e) \setminus \ell_e$ is the union of an isolated vertex with a unicyclic graph with cycle size z, r rooted trees, and $k-1$ internal edges. By our inductive hypothesis, the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ is $(-1)^{m_1-1}\left[(r-1)\binom{k-3}{m_1-2} + (z-1)\binom{k-3}{m_1-1}\right]$. Finally, $G \odot e$ is a unicyclic graph with cycle size z, r rooted trees, and $k-1$ internal edges. Again by our inductive hypothesis, $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{G\odot e}$ is $(-1)^{m_1}\left[(r-1)\binom{k-3}{m_1-1} + (z-1)\binom{k-3}{m_1}\right]$. Thus by the DNC relation and Pascal's rule, the coefficient of $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_G$ is

$$\begin{aligned}
c_{(n-m_1,1^{m_1})} =& (-1)^{m_1}\left[\binom{k-2}{m_1} + (r-1)\left[\binom{k-3}{m_1-2} + \binom{k-3}{m_1-1}\right]\right.\\
&\left. + (z-1)\left[\binom{k-3}{m_1-1} + \binom{k-3}{m_1}\right]\right]\\
=& (-1)^{m_1}\left[\binom{k-2}{m_1} + (r-1)\binom{k-2}{m_1-1} + (z-1)\binom{k-2}{m_1}\right]\\
=& (-1)^{m_1}\left[(r-1)\binom{k-2}{m_1-1} + z\binom{k-2}{m_1}\right].
\end{aligned}$$

Case 1b: Suppose e is an edge on the cycle that is adjacent to two rooted trees. So $r \geqslant 2$. Following a similar argument as above, $G \setminus e$ is a tree with $k-1$ internal edges, so the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{G\setminus e}$ is $(-1)^{m_1}\binom{k-1}{m_1}$. Next, $(G \odot e) \setminus \ell_e$ is the union of an isolated vertex with a unicyclic graph with cycle size z, $r-1$ rooted trees, and $k-1$ internal edges. By our inductive hypothesis, the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ is $(-1)^{m_1-1}\left[(r-2)\binom{k-3}{m_1-2} + (z-1)\binom{k-3}{m_1-1}\right]$. Finally, $G \odot e$ is a unicyclic graph with cycle size z, $r-1$ rooted trees, and $k-1$ internal edges. By our inductive hypothesis, $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_{G\odot e}$ is $(-1)^{m_1}\left[(r-2)\binom{k-3}{m_1-1} + (z-1)\binom{k-3}{m_1}\right]$. Thus the coefficient $c_{(n-m_1,1^{m_1})}$ in $\mathbf{X}_G$ is

$$\begin{aligned}
c_{(n-m_1,1^{m_1})} =& (-1)^{m_1}\left[\binom{k-1}{m_1} + (r-2)\left[\binom{k-3}{m_1-2} + \binom{k-3}{m_1-1}\right]\right.\\
&\left. + (z-1)\left[\binom{k-3}{m_1-1} + \binom{k-3}{m_1}\right]\right]
\end{aligned}$$

$$
\begin{aligned}
&=(-1)^{m_1}\left[\binom{k-1}{m_1}+(r-2)\binom{k-2}{m_1-1}+(z-1)\binom{k-2}{m_1}\right]\\
&=(-1)^{m_1}\left[\binom{k-2}{m_1}+\binom{k-2}{m_1-1}+(r-2)\binom{k-2}{m_1-1}+(z-1)\binom{k-2}{m_1}\right]\\
&=(-1)^{m_1}\left[(r-1)\binom{k-2}{m_1-1}+z\binom{k-2}{m_1}\right].
\end{aligned}
$$

Case 2 ($r = 0$): We again induct on the size of the cycle c. Suppose that for any cycle graph C_z on z vertices where $z \geqslant 3$, the hook coefficients are

$$c_{(z-m_1,1^{m_1})} = (-1)^{m_1}\left[-\binom{z-2}{m_1-1}+(z-1)\binom{z-2}{m_1}\right].$$

This follows from $r = 0$ and $n = k = z$ in this case. Let $G = C_{z+1}$. We show that the hook coefficients in C_{z+1} are

$$c_{(z+1-m_1,1^{m_1})} = (-1)^{m_1}\left[-\binom{z-1}{m_1-1}+z\binom{z-1}{m_1}\right].$$

Let e be any edge on the cycle. Then $G\backslash e$ is a tree with $z-2$ internal edges, so $c_{(z+1-m_1,1^{m_1})}$ in $\mathbf{X}_{G\backslash e}$ is $(-1)^{m_1}\binom{z-2}{m_1}$. Now, $(G \odot e) \setminus \ell_e$ is the union of an isolated vertex with the cycle graph on z vertices. By our inductive hypothesis, the coefficient $c_{(z+1-m_1,1^{m_1})}$ in $\mathbf{X}_{(G\odot e)\backslash \ell_e}$ is $(-1)^{m_1-1}\left[-\binom{z-2}{m_1-2}+(z-1)\binom{z-2}{m_1-1}\right]$. Finally, $G \odot e$ is a unicyclic graph with cycle size z, 1 rooted tree, and z internal edges. By case 1, $c_{(z+1-m_1,1^{m_1})}$ in $\mathbf{X}_{G\odot e}$ is $(-1)^{m_1}\left[(z-1)\binom{z-2}{m_1}\right]$. Thus the coefficient $c_{(z+1-m_1,1^{m_1})}$ in $\mathbf{X}_G$ is

$$
\begin{aligned}
c_{(n-m_1,1^{m_1})} &=(-1)^{m_1}\left[\binom{z-2}{m_1}-\binom{z-2}{m_1-2}+(z-1)\binom{z-2}{m_1-1}+(z-1)\binom{z-2}{m_1}\right]\\
&=(-1)^{m_1}\left[\binom{z-2}{m_1}+\binom{z-2}{m_1-1}-\binom{z-1}{m_1-1}+(z-1)\binom{z-1}{m_1}\right]\\
&=(-1)^{m_1}\left[-\binom{z-1}{m_1-1}+z\binom{z-1}{m_1}\right]. \qquad \square
\end{aligned}
$$

Consequently, given any connected unicyclic graph G, we can determine the coefficients for the elements in the star-basis expansion corresponding to hook partitions. We immediately obtain the following corollaries.

Corollary 11. *Let G be a connected unicyclic graph of order n with a single cycle of size $c \geqslant 3$. If $\mathbf{X}_G = \sum_{\lambda\vdash n} c_\lambda \mathfrak{st}_\lambda$, then*

$$c_{(n)} = c - 1.$$

Corollary 12. *If G_1 and G_2 are connected unicyclic graphs with cycles of different sizes, then $\mathbf{X}_{G_1} \neq \mathbf{X}_{G_2}$.*

In particular, the previous corollary shows that the star-expansion of the chromatic symmetric function of a unicyclic graph can be used to distinguish the cycle C_n from other connected unicyclic graphs of order n. The cycle will be the only connected unicyclic graph with $c_{(n)} = n-1$; alternatively, it will be the only connected unicyclic graph for which $c_{(2,1^{n-2})} \neq 0$, since no other connected unicyclic graph with n vertices satisfies $k = n$.

Next, we show how the longest hook partition encodes the relationship between the number of internal edges and rooted trees.

Proposition 13. *Let G be a connected unicyclic graph of order n, k internal edges, r rooted trees, and a cycle of size $c \geqslant 3$. If $\mathbf{X}_G = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$ and $\lambda^\star = (n-m_1, 1^{m_1})$ is the longest hook partition such that $c_{\lambda^\star} \neq 0$, then*

$$k = \begin{cases} m_1 + 2 & \text{if } r = 0 \text{ or } 1, \\ m_1 + 1 & \text{if } r \geqslant 2. \end{cases}$$

Furthermore, the coefficient $c_{\lambda^\star}$ corresponding to the longest hook partition is

$$c_{\lambda^\star} = \begin{cases} (-1)^{k-2} & \text{if } r = 0, \\ (-1)^{k-2}(c-1) & \text{if } r = 1, \\ (-1)^{k-1}(r-1) & \text{if } r \geqslant 2. \end{cases}$$

Proof. We consider the cases when $r = 0$, $r = 1$, and $r \geqslant 2$ separately. First consider when $r = 0$, that is, G is cycle graph C_n, and $n = k = c$. Then Equation (1) reduces to

$$c_{(n-m_1,1^{m_1})} = (-1)^{m_1}\left[-\binom{k-2}{m_1-1} + (k-1)\binom{k-2}{m_1}\right].$$

The largest m_1 such that $c_{(n-m_1,1^{m_1})} \neq 0$ is $m_1 = k-2$, since the next hook longer than the partition $(n-k+2, 1^{k-2}) = (2, 1^{k-2})$ is (1^k), which is when $m_1 = k$ and hence $c_{(1^k)} = 0$. Next, consider when $r = 1$. Then Equation (1) reduces to

$$c_{(n-m_1,1^{m_1})} = (-1)^{m_1}\left[(c-1)\binom{k-2}{m_1}\right]$$

and we again see that the largest m_1 such that $c_{(n-m_1,1^{m_1})} \neq 0$ is $m_1 = k-2$. Finally, if $r \geqslant 2$, the largest m_1 such that $c_{(n-m_1,1^{m_1})} \neq 0$ is $m_1 = k-1$, as can be seen by examining Equation (1). $\square$

Example 14. Let G be the graph in Figure 4. We have that $\mathbf{X}_G = 3\,\mathfrak{st}_{(8)} - 10\,\mathfrak{st}_{(7,1)} + 4\,\mathfrak{st}_{(6,2)} + 12\,\mathfrak{st}_{(6,1,1)} + 2\mathfrak{st}_{(5,3)} - 10\,\mathfrak{st}_{(5,2,1)} - 6\,\mathfrak{st}_{(5,1,1,1)} + 1\,\mathfrak{st}_{(4,4)} - 5\,\mathfrak{st}_{(4,3,1)} + 2\,\mathfrak{st}_{(4,2,2)} + 8\,\mathfrak{st}_{(4,2,1,1)} + 1\,\mathfrak{st}_{(4,1,1,1,1)} + 2\,\mathfrak{st}_{(3,3,2)} + 2\,\mathfrak{st}_{(3,3,1,1)} - 4\,\mathfrak{st}_{(3,2,2,1)} - 2\,\mathfrak{st}_{(3,2,1,1,1)} + 1\,\mathfrak{st}_{(2,2,2,1,1)}$. The hook coefficients are highlighted in red and match Equation 1 as $n = 8, c = 4, k = 5$, and $r = 2$.

Additionally, the longest hook partition is $\lambda^\star = (4,1,1,1,1)$ and the corresponding coefficient is $c_{\lambda^\star} = (-1)^{k-1}(r-1) = 1$.

Corollary 15. *Let G_1 and G_2 be connected unicyclic graphs on n vertices with k_1 and k_2 internal edges and r_1 and r_2 rooted tree respectively. If $k_1 = k_2$ and $r_1 \neq r_2$ then $\mathbf{X}_{G_1} \neq \mathbf{X}_{G_2}$.*

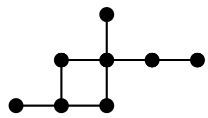

Figure 4: A unicyclic graph G on 8 vertices.

Figure 5: Two graphs with the same hook coefficients: $c_{(8)} = 3$, $c_{(7,1)} = -9$, $c_{(6,1,1)} = 9$, $c_{(5,1^3)} = -3$, and $c_\lambda = 0$ for all other hooks λ.

The coefficient $c_{(n-1,1)}$ can be used to write a quadratic equation for r or k, yielding two possible solutions. We note however, that hook coefficients alone are insufficient for recovering these values exactly. As an example see Figure 5.

On the other hand, the following proposition shows that hook coefficients only fail to distinguish between the $r = 1$ and $r = c$ cases.

Proposition 16. *Let G and H be two connected unicyclic graphs such that the star-expansions of $\mathbf{X}_G$ and $\mathbf{X}_H$ have the same hook coefficients. Let r_G, r_H denote the number of non-trivial rooted trees of each graph, and let k_G, k_H denote the respective number of internal edges. Then, up to interchanging G and H, either:*

1. *$r_G = r_H$ and $k_G = k_H$, or*
2. *$r_G = 1$, $r_H = c$ where c is the size of the cycle in each graph, and $k_G = k_H + 1$.*

Proof. First note that by Theorem 11 both graphs have the same cycle size c. Furthermore, the cycle C_n is distinguished among unicyclic graphs by the coefficient $c_{(n)}$ as the unique graph with $c_{(n)} = n - 1$. Henceforth, we can assume $r_G, r_H \geqslant 1$. It follows from Theorem 13 that $k_G = m+2$ or $k_G = m+1$, where m is the number of 1's in the longest hook appearing in the star-expansion of $\mathbf{X}_G$. We know that if $k_G = k_H$ then $r_G = r_H$ by Theorem 13, which takes care of the first case. Supposing $k_H \neq k_G$, since $\mathbf{X}_H$, has the same hook coefficients it must be the case that $k_H = m + 2$ or $k_H = m + 1$ for the same m. Thus, without loss of generality let $k_G = m + 2$ and $k_H = m + 1$. This implies that $r_G = 1$ and $r_H \geqslant 2$. Moreover, applying the formula of Theorem 10 to the partition $(n - 1, 1)$ we have

$$\begin{aligned} c_{(n-1,1)}(G) = -[(r_G - 1) + (c-1)(k_G - 2)] &= -[(r_H - 1) + (c-1)(k_H - 2)] \\ &= c_{(n-1,1)}(H) \\ r_G + (c-1)m &= r_H + (c-1)(m-1) \end{aligned}$$

so that $r_H - r_G = c - 1$, implying $r_H = c$ as desired. □

As a consequence, there is an algorithm to obtain r from the hook coefficients unless $r = 1$ or c.

Corollary 17. *We can recover r, the number of non-trivial rooted trees attached to the cycle, from the CSF of a connected unicyclic graph whenever $1 < r < c$.*

Proof. From Theorem 13, we know that the number of internal edges is either $m+1$ or $m+2$, where m is the number of 1's appearing in the longest hook in the support of the star-expansion of $\mathbf{X}_G$. As in the proof of that proposition, substituting $m+1$ and $m+2$ for k into

$$c_{(n-1,1)} = -[(r-1) + (c-1)(k-2)]$$

will give two possible values for r whose difference is $c-1$. Unless these values are 1 and c, one of them will be greater than c (which is impossible) or less than 1. If these values are 0 and $c-1$, whether the graph is C_n $(r = 0)$ can be determined by the coefficient of $c_{(n)}$, as discussed after Theorem 12. Otherwise, one of these two values will be negative, which is impossible. Thus, we only get one valid solution when $1 < r < c$ and in that case, we have

$$r = -[(c_{(n-1,1)} - 1) + (c-1)(m-1)]. \qquad \square$$

Remark 18. As shown in Figure 5, the hook coefficients are not enough to determine the number of nontrivial rooted trees attached to the cycle. In fact, Corollary 17 is the best we can do, since there are unicyclic graphs with only one rooted tree attached to a cycle that have the same chromatic symmetric function as another that has c rooted trees attached to the cycle. Here is the smallest example of an infinite family exhibiting this behavior, as described in [12]:

Figure 6: Two graphs with the same chromatic symmetric function where one has one nontrivial rooted tree and the other has three.

4 Paths, cycles, and pans

In this section, we explicitly compute the coefficients of all partitions in the star-expansions of paths, cycles, and cycles with a single leaf, also called *pan graphs*. While paths are not unicyclic, they provide a test case for illustration of the technique. All three types of graphs have the advantage that their internal edges can be labeled with consecutive integers $\{1, 2, \ldots, s\}$ in a natural way. This determines an order in which the star-expansion algorithm is applied, where at each stage we apply the DNC relation to the remaining internal edge with the smallest label.

In this set-up, we may encode a leaf of the DNC-tree for a graph G as a word $\underline{w} = w_1 w_2 \ldots w_s$ in the alphabet $\{L, M, R, X\}$, where the symbol at position i indicates which branch of the DNC-tree we follow when acting on internal edge i. Specifically:

- $w_i = L$ means that when we reach the node of the DNC-tree at which internal edge i is acted upon, we proceed to the graph obtained by deletion of this edge.

- $w_i = M$ means that when acting on internal edge i, we proceed to the graph obtained by dot-contraction.

- $w_i = R$ means that when acting on internal edge i, we proceed to the graph obtained by leaf-contraction.
- $w_i = X$ means that after applying the operations corresponding to $w_1, \dots, w_{i-1}$, the edge with label i is no longer internal so it cannot be acted on. In the DNC-tree, we instead take the next available internal edge and continue the algorithm.

All distinct words that identify valid DNC-tree leaves contribute to the coefficient of some $\mathfrak{st}_\lambda$ in the star-expansion of $\mathbf{X}_G$ in a cancellation-free way. We call a word $\underline{w}$ that produces the star forest St_λ a *λ-word*, and we compute the coefficient c_λ by counting such words. We note that in all of our labelings, internal edges i and $i+1$ will be incident for $1 \leqslant i < s$, which implies that $w_i = L$ will often force $w_{i+1} = X$. On the other hand, the operations $w_i = M$ or $w_i = R$ do not usually impact whether or not w_{i+1} is internal, as we will discuss in the proofs of our results.

Example 19. Let $G = P_7$ where the internal edges are labeled $\{1, 2, 3, 4\}$ from left to right. Then, $MRML$ and $MLXM$ are two valid λ-words for $\lambda = (3, 2, 1, 1)$, implying that the absolute value of the coefficient c_λ in the star-expansion of $\mathbf{X}_{P_7}$ is at least 2.

Definition 20. If we write a partition as $\lambda = (\lambda_1, \dots, \lambda_{\ell(\lambda)})$, let b be the index of the last part satisfying $\lambda_b > 1$. We call the partition $\hat{\lambda} = (\lambda_1, \dots, \lambda_b)$ the *body* of λ and $\tilde{\lambda} = (\lambda_{b+1}, \dots, \lambda_{\ell(\lambda)}) = (1, \dots, 1)$ the *tail*.

Proposition 21. *Let $n \geqslant 4$ and let P_n the path with n vertices. Then*

$$\mathbf{X}_{P_n} = \sum_{\lambda = (1^{m_1} 2^{m_2} \dots n^{m_n}) \vdash n} (-1)^{m_1} \binom{m_2 + \dots + m_n}{m_2, \dots, m_n} \binom{n - 2 - \ell(\lambda) + m_1}{m_1} \mathfrak{st}_\lambda.$$

Proof. Label the internal edges of P_n consecutively from one end to the other by $\{1, \dots, n-3\}$, and apply the star-expansion algorithm to these edges by selecting the edge with smallest possible label at each stage. We analyze each factor that constitutes the coefficient c_λ for $\lambda = (1^{m_1} 2^{m_2} \dots n^{m_n})$ separately.

Signs in the star-expansion algorithm are introduced only when applying dot-contraction. Starting from a connected graph with at least two vertices, dot-contraction is also the only way to obtain parts of size 1 in a star forest produced by the star-expansion algorithm. Then, m_1 must be the number of times M appears in any word that produces St_λ, accounting for the factor of $(-1)^{m_1}$.

We now explain the multinomial coefficient. Observe that any valid λ-word $\underline{w} = w_1 \dots w_{n-3}$ must also satisfy an additional condition: there must be exactly $b - 1 = \ell(\lambda) - m_1 - 1$ symbols w_i with $w_i = L$ in $\underline{w}$, where b is the number of body parts, $\ell(\hat{\lambda})$. This follows because every deletion of an internal edge produces a new connected component of size at least 2, while neither dot-contraction nor leaf-contraction do so.

The presence of M's anywhere in the string $\underline{w}$ does not affect the part sizes of the body partition so we may ignore them for determining these parts. Hence, let $\underline{v} = v_1 \dots v_{n-3-m_1}$ be the word obtained from $\underline{w}$ by omitting all M's.

Then $\underline{v}$ has the form $\underline{v} = R \cdots R$ if no L's appear, and if at least one L appears then

$$\underline{v} = R \cdots RLXR \cdots RLXR \cdots \cdots RLXR \cdots R,$$

where $v_{i_1}, \ldots, v_{i_{b-1}} = L$ for some $1 \leqslant i_1 < i_2 < \cdots < i_{b-1} \leqslant n-3-m_1$. In addition, we must have $v_{i_j+1} = X$ if $i_j < n-3-m_1$, since the next edge will no longer be internal. In particular, it may be that $i_{b-1} = n-3-m_1$ in which case this L is not followed by an X.

By counting the number of symbols between the L's, we obtain a rearrangement (or "permutation") $\pi = (\pi_1, \ldots, \pi_b)$ of the body parts of λ as follows:

$$\overbrace{R\cdots R}^{\pi_1-2}\overbrace{LXR\cdots R}^{\pi_2}\overbrace{LXR\cdots R}^{\pi_3}LXR\cdots\cdots R\overbrace{LXR\cdots R}^{\pi_b-1}.$$

That is, $\pi_b = \pi_1 = n-3-m_1$ if there are no L's, and if at least one L appears then

- $\pi_1 = i_1 + 1$,
- $\pi_j = i_j - i_{j-1}$ for $2 \leqslant j \leqslant b-1$ (when $b \geqslant 3$),
- and $\pi_b = n-1-m_1-i_{b-1}$.

To account for π_b, for example, observe that if the last deletion L occurs at internal edge s, there is still a path on $n-1-s$ vertices in the connected component containing any remaining internal edges. The number of vertices π_b that remain in this last connected component is exactly $n-1-m_1-i_{b-1}$, so that in particular if $i_{b-1} = n-3-m_1$, $\pi_b = 2$. Determination of the other parts resulting from the word $\underline{v}$ is similar, noting that each R will grow some part π_j by 1 by definition of leaf-contraction.

Thus, a valid λ-word $\underline{w}$ must satisfy the property that the subword $\underline{v}$ produces a rearrangement π of the body parts of λ in this way. Since there is a unique string $\underline{v}$ for each π, we count $\binom{m_2+\cdots+m_n}{m_2,\ldots,m_n}$ possibilities.

Once the body part rearrangement is determined, a valid λ-word can be completed by inserting m_1 symbols M into the word $\underline{v}$ so as to recover a word $\underline{w}$ of length $n-3$.

To count the possibilities here, consider the set of valid DNC-tree words with $b-1$ $L's$ as the disjoint union of two sets.

1. W_1 consisting of words of length $n-3-(b-1)$ in $\{LX, M, R\}$, where LX counts as one symbol and is used $b-1$ times.
2. W_2 consisting of words of length $n-3-(b-1)$ in $\{LX, M, R\}$, where LX is used $b-2$ times, appended by the symbol L.

In either case, we choose which m_1 of the $n-3-(b-1) = n-2-(\ell(\lambda)-m_1)$ symbols should be M's, which can be done in $\binom{n-2-\ell(\lambda)+m_1}{m_1}$ many ways. Once this is done, the rest of the symbols are determined by the choice of body part rearrangement and associated word $\underline{v}$, so multiplying the binomial coefficient with the multinomial $\binom{m_2+\cdots+m_n}{m_2,\ldots,m_n}$ gives the number of DNC leaves with star forest St_λ. $\square$

Now we handle the cycle C_n, which is, in some sense, the primordial connected unicyclic graph and already gives some indications of how this class differs from trees. As we discussed in the previous section, the coefficients c_λ for hook partitions λ behave differently than non-hook coefficients.

Proposition 22. *Let C_n be the cycle on n vertices for $n \geqslant 3$, and let $\mathbf{X}_{C_n} = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$. Then for $\lambda = (1^{m_1} 2^{m_2} \dots n^{m_n})$ we have*

$$c_\lambda = \begin{cases} (-1)^{m_1} \left[-\binom{n-2}{m_1-1} + (n-1)\binom{n-2}{m_1} \right] & \text{if } \lambda = (n-m_1, 1^{m_1}),\ m_1 \neq n-1, \\ \\ (-1)^{m_1} \frac{n}{\ell(\lambda)-m_1} \binom{m_2+\dots+m_n}{m_2,\dots,m_n} \binom{n-\ell(\lambda)+m_1-1}{m_1} & \text{else.} \end{cases}$$

Proof. Again label the internal edges $\{1, 2, \dots, n\}$ around the cycle and apply DNC relations to the internal edge with smallest label at each stage. As before, each leaf of the DNC-tree is determined by a word in $\{L, M, R, X\}$ of length n which we denote $\underline{w} = w_1 \dots w_n$. The sign $(-1)^{m_1}$ is always given by the number of M's in a λ-word, since both isolated vertices in St_λ and signs arise only from dot-contraction. We argue for the remainder of each expression separately.

Hook case: The desired equation immediately follows from Theorem 10, as $c = n = k$ and $r = 0$ in this case.

Non-hook case: As above, let $\underline{v} = v_1 \dots v_{n-m_1}$ be the word obtained from a λ-word $\underline{w}$ by omitting all M's. We will count the contributions from $v_1 = L$ and $v_1 \neq L$ separately.

For each body part rearrangement π of $\hat{\lambda} = (\lambda_1, \dots, \lambda_b)$, there is a unique corresponding subword $\underline{v}$ yielding π for which $v_1 = L$. That is, no R appears before the first L so this word has the form

$$\overbrace{LXR \cdots R}^{\pi_1} \overbrace{LXR \cdots R}^{\pi_2} LXR \cdots \cdots R \overbrace{LXR \cdots RX}^{\pi_b}.$$

To reconstruct a valid λ-word $\underline{w}$ with this $\underline{v}$, we choose where to place m_1 M's among $n - b - 1$ slots since the last symbol is $w_n = X$ here. Thus, these words contribute

$$\binom{m_2 + \dots + m_n}{m_2, \dots, m_n} \binom{n - \ell(\lambda) + m_1 - 1}{m_1} \tag{2}$$

leaves to c_λ.

Now we consider λ-words with $v_1 \neq L$; that is, those $\underline{w}$ for which the subword $\underline{v}$ contains an R before the first L. Here, we are free to choose where to place the M's from any of the $n - b$ slots corresponding to letters in the alphabet $\{LX, M, R\}$. The subword $\underline{v}$ can be constructed as

$$\overbrace{R \cdots R}^{p} \overbrace{LXR \cdots R}^{\pi_1} \overbrace{LXR \cdots R}^{\pi_2} LXR \cdots \cdots R \overbrace{LXR \cdots R}^{\pi_{b-1}} \overbrace{LXR \cdots R}^{\pi_b - p},$$

so that for any body part rearrangement ending in a given $\pi_b = j$, there are $j - 1$ possible subwords yielding the same π. Thus, this gives a contribution of

$$\left(\sum_{j=2}^{n} (j-1) \binom{m_2 + \dots + m_n - 1}{m_2, \dots, m_j - 1, \dots, m_n} \right) \binom{n - \ell(\lambda) + m_1}{m_1}.$$

Since

$$\sum_{j=2}^{n} (j-1) \binom{m_2 + \dots + m_n - 1}{m_2, \dots, m_j - 1, \dots, m_n} = \sum_{j=2}^{n} \frac{(j-1) m_j}{m_2 + \dots + m_n} \binom{m_2 + \dots + m_n}{m_2, \dots, m_n}$$

$$= \frac{1}{m_2 + \cdots + m_n} \left(\sum_{j=2}^{n} (j-1) m_j \right) \binom{m_2 + \cdots + m_n}{m_2, \ldots, m_n}$$

$$= \frac{n - \ell(\lambda)}{\ell(\lambda) - m_1} \binom{m_2 + \cdots + m_n}{m_2, \ldots, m_n}.$$

Adding with (2) and simplifying gives the result. $\square$

Example 23. By Theorem 22, the CSF of the 5-cycle C_5 is

$$4\,\mathfrak{st}_{(5)} - 11\,\mathfrak{st}_{(4,1)} + 5\,\mathfrak{st}_{(3,2)} + 9\,\mathfrak{st}_{(3,1,1)} - 5\,\mathfrak{st}_{(2,2,1)} - 1\,\mathfrak{st}_{(2,1,1,1)}.$$

For instance, the hook partition $(4,1)$ has $m_1 = 1$, so $c_{(4,1)}$ is

$$(-1)\left[-\binom{3}{0} + 4\binom{3}{1}\right] = -11.$$

For the non-hook partition $(3,2)$, we have $m_1 = 0$, $m_2 = 1$, $m_3 = 1$, and $\ell(3,2) = 2$, so $c_{(3,2)}$ is

$$\frac{5}{2}\binom{2}{1,1}\binom{2}{0} = 5.$$

Similarly, for $(2,2,1)$ we have $m_1 = 1$, $m_2 = 2$, and $\ell(2,2,1) = 3$, so $c_{(2,2,1)}$ is

$$(-1)\frac{5}{2}\binom{2}{2}\binom{2}{1} = -5.$$

We highlight a few features of the formula for the star-expansion of $\mathbf{X}_{C_n}$. If $m_1 = 0$ and $\ell(\lambda) \geqslant 2$, we can write

$$c_\lambda = \frac{n(\ell(\lambda)-1)!}{\prod_{i=2}^{n} m_i!}.$$

This shows that $c_\lambda = n$ for any two-part partition with $\lambda_1 > \lambda_2 \geqslant 2$, or $c_{(n/2,n/2)} = n/2$ if n is even. We also see that the leading partition of $\mathbf{X}_{C_n}$ is $(2, 1^{n-2})$, since the only smaller partition is (1^n) which is seen to have coefficient 0 by our formula. The coefficient of $(2, 1^{n-2})$ is always $(-1)^{n-2}$.

Remark 24. When $n \geqslant 6$, for $\mathbf{X}_{C_n}$ our formula gives $c_{(3,3,1^{n-6})} = \binom{n-1}{3}\frac{n+2}{3}$. This sequence begins $0, 3, 14, 40, 90, 175, \ldots$ (OEIS A117662) and is equal to the number of certain pairs of non-crossing Dyck paths, or also exterior intersections of diagonals of an irregular n-gon.

Next we look at the CSF for the unicyclic graph consisting of one $(n-1)$-cycle and one leaf incident to one of the vertices of the cycle, also called the "pan" on n vertices.

Example 25. The pan graph for $n = 6$ and its chromatic symmetric function:

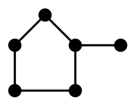

Figure 7: A pan graph

$$4\,\mathfrak{st}_{(6)} - 12\,\mathfrak{st}_{(5,1)} + 4\,\mathfrak{st}_{(4,2)} + 12\,\mathfrak{st}_{(4,1,1)} + 2\,\mathfrak{st}_{(3,3)} - 9\,\mathfrak{st}_{(3,2,1)} - 4\,\mathfrak{st}_{(3,1,1,1)} + \mathfrak{st}_{(2,2,2)} + 3\,\mathfrak{st}_{(2,2,1,1)}.$$

Proposition 26. *If G is the pan graph on n vertices and $\mathbf{X}_G = \sum_\lambda c_\lambda \mathfrak{st}_\lambda$, then for $\lambda = (1^{m_1}2^{m_2}\dots n^{m_n})$ we have*

$$c_\lambda = \begin{cases} (-1)^{m_1}(n-2)\binom{n-3}{m_1}, & \text{if } \lambda = (n-m_1, 1^{m_1}), \quad m_1 \neq n-1, \\ \\ (-1)^{m_1}\frac{n-\ell(\lambda)}{\ell(\lambda)-m_1}\binom{m_2+\dots+m_n}{m_2,\dots,m_n}\binom{n-\ell(\lambda)+m_1-1}{m_1} & \text{else.} \end{cases}$$

Proof. Orient the cycle so that we can label its internal edges from 1 to $n-1$ starting at one edge that is adjacent to the leaf-edge and ending at the other such edge. We again proceed to apply DNC relations to each internal edge in increasing order of the labels. Denote the word in $\{L, M, R, X\}$ that determines a DNC leaf by $\underline{w} = w_1 \dots w_{n-1}$. The sign is given by the number of M's in $\underline{w}$ as in the path and cycle cases.

Hook case: We immediately obtain the desired equation by Theorem 10 as $r = 1$ and $c = k = n-1$ in this case.

Non-hook case: Here, any λ-word must contain at least two L's. We again consider the subword $\underline{v}$ obtained by deleting all M's from a valid λ-word. Such $\underline{v}$ then has the form

$$\overbrace{R\cdots R}^{p}\overbrace{LXR\cdots R}^{\pi_1}\overbrace{LXR\cdots R}^{\pi_2}LXR\cdots\cdots R\overbrace{LXR\cdots R}^{\pi_{b-1}}\overbrace{LXR\cdots R}^{\pi_b-p-1}. \tag{3}$$

Let e be the leaf-edge of the initial pan graph. In contrast with the cycle case, having no R's before the first L no longer forces $w_{n-1} = X$ because of the presence of e. Thus, we can consider all words $\underline{v}$ of the form described in (3) together, and each one determines a rearrangement π of $\hat{\lambda} = (\lambda_1, \dots, \lambda_b)$ as indicated.

Notice that there are $\pi_b - 1$ ways of constructing a subword $\underline{v}$ that gives the same π by just cyclically shifting $\underline{v}$ to vary p, the number of initial R's appearing. Thus, for a body part permutation ending with $\pi_b = j$, we have $j-1$ choices of corresponding subword $\underline{v}$. This yields

$$\sum_{j=2}^{n}(j-1)\binom{m_2+\dots+m_n-1}{m_2,\dots,m_j-1,\dots m_n} = \frac{n-\ell(\lambda)}{\ell(\lambda)-m_1}\binom{m_2+\dots+m_n}{m_2,\dots,m_n}$$

total subword choices for each body partition $\dot{\lambda} = (\lambda_1, \dots, \lambda_b)$. Once the subword $\underline{v}$ is chosen, we can insert m_1 M's by choosing their positions from among $n - \ell(\lambda) + m_1 - 1$ slots, and completing the rest with the symbols from $\underline{v}$ to obtain a word of length $n-b-1$ in the alphabet $\{LX, M, R\}$. This accounts for the binomial coefficient in the formula, and multiplying with the number of options for $\underline{v}$ finishes the proof. □

We expect that the techniques of this section can be extended to other families of connected unicyclic graphs. For example, a *tadpole* is a unicyclic graph obtained by attaching a path to a single vertex on the cycle, so the pan graph is a special case where the attached path has length 1. A *squid* is a connected unicyclic graph with exactly one vertex of degree greater than 2. Alternatively one can think of a squid as a unicyclic graph with one non-trivial rooted tree which consists of some number of paths (called *tentacles*) emanating from the root. These families still have a natural ordering of their internal edges along the cycle and attached paths, so the corresponding DNC-tree leaves can, in principle, be encoded by words similar to the $\{L, M, R, X\}$ words used above.

However, for general unicyclic graphs, this technique may not immediately extend. In general, the rooted trees attached to the cycle may branch in many places, there is no single natural ordering of the internal edges, and the dependencies within the DNC-tree become more difficult to control.

5 Leading partitions for unicyclic graphs

In this section, we describe the leading partition of the chromatic symmetric function of any connected unicyclic graph G on n vertices, generalizing a result of Gonzalez, Orellana, and Tomba in [6, Theorem 4.15]. In the case where there are no non-trivial rooted trees attached to the cycle (i.e., $r = 0$), the leading partition is $(2, 1^{n-2})$ as discussed before Theorem 24. In this section we consider the case where $r \geqslant 1$. We first introduce some notation which we will use throughout this section.

Recall sort() is the function that takes in a sequence of non-negative integers and outputs it as a partition. Recall that $I(G)$ is the set of internal edges of the graph G and the connected components of $G \setminus I(G)$ are called the *leaf components* of G.

Definition 27. Notice that $G \setminus I(G)$ is a star forest since every remaining edge is incident with a leaf. We define $\lambda_{\mathrm{LC}}(G)$ to be the partition whose parts are the orders of the connected components of the resulting star forest in $G \setminus I(G)$.

Recall that λ_{lead} was defined in Theorem 3 as the smallest partition, λ, in lexicographic order such that $c_\lambda \neq 0$ in $\mathbf{X}_G$.

Theorem 28 ([6, Theorem 4.15])**.** *Let F be a forest. Then $\lambda_{\mathrm{lead}}(\mathbf{X}_F) = \lambda_{\mathrm{LC}}(F)$.*

Example 29. Let F be the forest in Figure 8. After removing the only internal edge e indicated in red, we get $\lambda_{\mathrm{lead}}(\mathbf{X}_F) = \lambda_{\mathrm{LC}}(F) = (3, 2, 1)$.

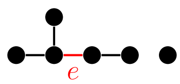

Figure 8: A forest on 6 vertices.

Notation 30. Let G be a connected c-unicyclic graph of order n. We label vertices in the c-cycle with $v_1, v_2, \ldots, v_c$ such that $v_i v_{i+1}$ is an edge for $1 \leqslant i \leqslant c-1$, as well as $v_c v_1$. Furthermore, each v_i is the root of some rooted tree T_i (which could be trivial). Let $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_c)$ be the sequence of positive integers where each part λ_i equals 1 plus the number of leaves attached to v_i in G for $1 \leqslant i \leqslant c$. Let $\mu^{(i)}$ be the partition obtained from the leaf components of T_i after removing v_i and the leaves incident to v_i. Let $\mu = (\mu^{(1)} \cdot \mu^{(2)} \cdot \cdots \cdot \mu^{(c)})$.

Note that for a connected c-unicyclic graph G we have $\mathrm{sort}(\lambda \cdot \mu) = \lambda_{\mathrm{LC}}(G)$. We will see that this will give the leading partition for $\mathbf{X}_G$ when there are at least two non-trivial rooted trees. When there is only one non-trivial rooted tree, we need to make a minor change to this partition to get the leading partition for $\mathbf{X}_G$.

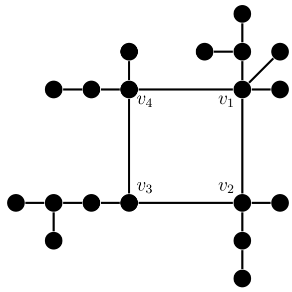

Figure 9: A graph with a 4-cycle.

Example 31. Consider the graph in Figure 9. We have $\lambda = (3,2,1,2)$, $\mu^{(1)} = (3)$, $\mu^{(2)} = (2)$, $\mu^{(3)} = (3,1)$ and $\mu^{(4)} = (2)$. Hence, $\mu = (3,2,3,1,2)$ and $\lambda_{\mathrm{LC}}(G) = \mathrm{sort}(\lambda\cdot\mu) = (3,3,3,2,2,2,2,1,1)$.

Our main goal for this section is to prove the following.

Theorem 32. *Let G be a unicyclic graph and $r \geqslant 1$ denote the number of non-trivial rooted trees attached to the cycle in G. Let $\lambda_{\mathrm{lead}} = \lambda_{\mathrm{lead}}(\mathbf{X}_G)$.*

1. *If $r = 1$ then*
$$\lambda_{\mathrm{lead}} = \lambda_{\mathrm{LC}}(G \setminus e)$$
where e is chosen to be an edge in the cycle incident to the root of the non-trivial tree.

2. *If $r \geqslant 2$ then*
$$\lambda_{\mathrm{lead}} = \lambda_{\mathrm{LC}}(G).$$

Note that due to Theorem 4, it suffices to prove Theorem 32 for connected unicyclic graphs.

5.1 One non-trivial rooted tree ($r = 1$)

We will prove the $r = 1$ case of Theorem 32(1) by induction on the size of the cycle. First we consider the base case, when G contains a triangle.

Lemma 33. *When $c = 3$, Theorem 32(1) holds.*

Proof. Define $\lambda_1, \lambda_2, \lambda_3, \mu$ as in Notation 30. First, we note that if G has only one non-trivial rooted tree, say T_1 with root v_1, $\lambda_{\mathrm{LC}}(G \setminus e) = \mathrm{sort}(\lambda_1, \mu, 2)$ where $e = v_1v_2$.

The proof follows from a straightforward computation using the DNC relation on e and Theorem 28. As $G \setminus e, G \odot e$ and $(G \odot e) \setminus \ell_e$ are all forests, Theorem 28 applies.

$$\begin{aligned}
\lambda_{\mathrm{lead}}(\mathbf{X}_{G\setminus e}) &= \lambda_{\mathrm{LC}}(G \setminus e) = \mathrm{sort}(\lambda_1, \mu, 2)\\
\lambda_{\mathrm{lead}}(\mathbf{X}_{G\odot e}) &= \mathrm{sort}(\lambda_1 + 2, \mu)\\
\lambda_{\mathrm{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) &= \mathrm{sort}(\lambda_1 + 1, \mu, 1).
\end{aligned}$$

Since $\mathrm{sort}(\lambda_1, \mu, 2) \leqslant \mathrm{sort}(\lambda_1 + 1, \mu, 1) < \mathrm{sort}(\lambda_1 + 2, \mu)$, we have completed the proof. $\square$

Example 34. Let G be the graph in Figure 10. We have $\lambda_1 = \lambda_2 = \lambda_3 = 1$ and $\mu = (3)$. Let $e = v_1v_2$. We compute:

$$\lambda_{\text{lead}}(\mathbf{X}_{G\setminus e}) = (3,2,1)\,,\ \lambda_{\text{lead}}(\mathbf{X}_{G\odot e}) = (3,3)\,,\ \lambda_{\text{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) = (3,2,1)\,.$$

Therefore $\lambda_{\text{lead}}(\mathbf{X}_G) = (3,2,1) = \lambda_{\text{LC}}(G \setminus e)$.

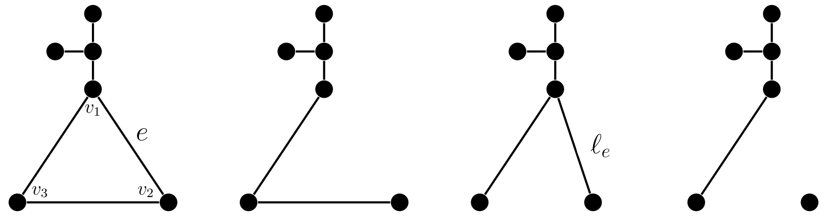

Figure 10: Left to right: G, $G\setminus e$, $G\odot e$, $(G\odot e)\setminus \ell_e$.

Proposition 35. *Theorem 32(1) is true for all $c \geqslant 3$.*

Proof. Define λ, μ as in Notation 30. We prove the statement by induction on c. The base case is proven in Theorem 33. Suppose the statement holds for any c-unicyclic graph with one non-trivial rooted tree where $c \geqslant 3$. Let G be a $(c+1)$-unicyclic graph with one non-trivial rooted tree, say T_1, so that $\lambda_2 = \lambda_3 = \cdots = \lambda_{c+1} = 1$. Let $e = v_1v_2$.

First we compute $\lambda_{\text{LC}}(G \setminus e) = \text{sort}(\lambda_1, \lambda_2+\lambda_3, \ldots, \lambda_{c+1}, \mu) = \text{sort}(\lambda_1, 2, 1^{c-2}, \mu)$. Consider the three graphs that we will obtain by applying the DNC relation on the edge e in G. By Theorem 28, we have $\lambda_{\text{lead}}(\mathbf{X}_{G\setminus e}) = \text{sort}(\lambda_1, 2, 1^{c-2}, \mu) = \lambda_{\text{LC}}(G\setminus e)$.

Since $G \odot e$ is c-unicyclic with exactly one rooted tree, we can apply the inductive hypothesis to compute $\lambda_{\text{lead}}(\mathbf{X}_{G\odot e})$. Using $e' = v_1v_3$, we get

$$\lambda_{\text{lead}}(\mathbf{X}_{G\odot e}) = \lambda_{\text{LC}}((G\odot e)\setminus e') = \text{sort}(\lambda_1+1, 2, 1^{c-3}, \mu),$$

as a new leaf becomes incident to v_1 as a result of the leaf-contraction. Similarly, we can compute

$$\lambda_{\text{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) = \lambda_{\text{LC}}(((G\odot e)\setminus \ell_e)\setminus e') = \text{sort}(\lambda_1, 2, 1^{c-2}, \mu),$$

where the additional 1 arises from the dot-contraction.

Since $\lambda_{\text{lead}}(\mathbf{X}_{G\setminus e}) = \lambda_{\text{LC}}(G\setminus e) = \text{sort}(\lambda_1, 2, 1^{c-2}, \mu) < \text{sort}(\lambda_1+1, 2, 1^{c-3}, \mu) = \lambda_{\text{lead}}(\mathbf{X}_{G\odot e})$, we have completed the proof. $\square$

The following corollaries are evident from the proof given above.

Corollary 36. *When we apply DNC relation on a graph satisfying the condition in Theorem 32(1), then the leading partition does not come from leaf-contraction when the edge is incident to a non-trivial rooted tree.*

Corollary 37. *When G is a graph satisfying the condition in Theorem 32(1), then 2 must be a part in $\lambda_{\text{lead}}(\mathbf{X}_G)$.*

5.2 More than one non-trivial rooted tree ($r \geqslant 2$)

Here we assume we have more than one non-trivial tree attached at distinct cycle vertices.

Lemma 38. *When $c = 3$, Theorem 32(2) holds.*

Proof. Since $r \geqslant 2$, without loss of generality we may assume T_1 and T_2 are non-trivial trees, and let $e = v_1v_2$.

First note $\lambda_{\mathrm{LC}}(G) = \mathrm{sort}(\lambda_1, \lambda_2, \lambda_3, \mu)$. Since in this case e is adjacent to both T_1 and T_2, this is also equal to $\lambda_{\mathrm{lead}}(\mathbf{X}_{G\setminus e})$. We apply the DNC relation on the edge e of G.

Next we compute the leading partitions in the other two terms in the DNC relation. Notice that after leaf-contracting, the vertices of the contracted edge are identified to a single vertex, say v_0, with $\lambda_1 + \lambda_2 - 1$ leaves. In addition, $G \odot e$ is now a tree, making $\lambda_{\mathrm{lead}}(\mathbf{X}_{G\odot e}) = \lambda_{\mathrm{LC}}(G \odot e)$. If T_3 is a trivial tree, then $\lambda_3 = 1$ and will become a leaf incident to v_0 during the leaf-contraction and thus $\lambda_{\mathrm{lead}}(\mathbf{X}_{G\odot e}) = \mathrm{sort}(\lambda_1 + \lambda_2 + \lambda_3, \mu)$. If instead, T_3 is not a trivial tree, then the edge v_0v_3 is internal and λ_3 will stay as its own leaf component in $\lambda_{\mathrm{LC}}(G \odot e)$. Therefore we have:

$$\lambda_{\mathrm{lead}}(\mathbf{X}_{G\odot e}) = \begin{cases} \mathrm{sort}(\lambda_1 + \lambda_2 + \lambda_3, \mu) & \text{if } T_3 \text{ is trivial,} \\ \mathrm{sort}(\lambda_1 + \lambda_2, \lambda_3, \mu) & \text{if } T_3 \text{ is not trivial.} \end{cases}$$

Following a similar discussion, we can compute $\lambda_{\mathrm{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e})$:

$$\lambda_{\mathrm{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) = \begin{cases} \mathrm{sort}(\lambda_1 + \lambda_2 + \lambda_3 - 1, 1, \mu) & \text{if } T_3 \text{ is trivial,} \\ \mathrm{sort}(\lambda_1 + \lambda_2 - 1, 1, \lambda_3, \mu) & \text{if } T_3 \text{ is not trivial.} \end{cases}$$

In both cases, we have $\lambda_{\mathrm{LC}}(G) = \lambda_{\mathrm{lead}}(\mathbf{X}_{G\setminus e}) \leqslant \lambda_{\mathrm{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) < \lambda_{\mathrm{lead}}(\mathbf{X}_{G\odot e})$, which proves the statement. ☐

Example 39. Let G be the graph in Figure 11. We have $\lambda_1 = 1, \lambda_2 = 3, \lambda_3 = 1$ and $\mu = (3, 2)$. In addition T_3 is a trivial rooted-tree at v_3. Let $e = v_1v_2$. We compute:

$$\lambda_{\mathrm{lead}}(\mathbf{X}_{G\setminus e}) = (3,3,2,1,1)\,,\ \lambda_{\mathrm{lead}}(\mathbf{X}_{G\odot e}) = (5,3,2)\,,\ \lambda_{\mathrm{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) = (4,3,2,1)\,.$$

Therefore, $\lambda_{\mathrm{lead}}(\mathbf{X}_G) = (3,3,2,1,1) = \lambda_{\mathrm{LC}}(G)$.

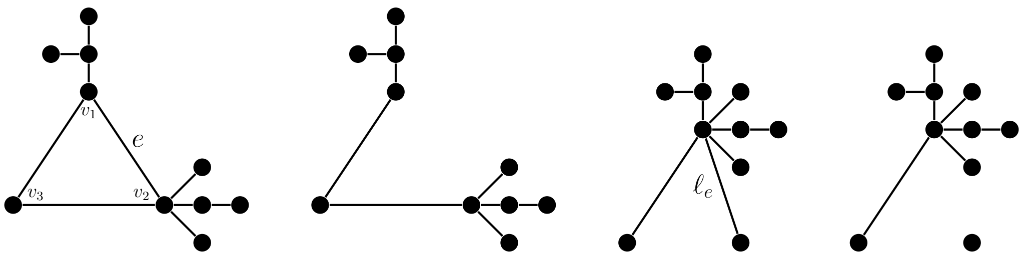

Figure 11: Left to right: $G,\ G \setminus e, G \odot e,\ (G \odot e) \setminus \ell_e$.

Proposition 40. *Theorem 32(2) is true for all $c \geqslant 3$.*

Proof. We prove the statement by induction on c. The base case is proven in Theorem 38. Suppose the statement holds for any c-unicyclic graph with at least 2 non-trivial rooted trees for some $c \geqslant 3$. Let G be a $(c+1)$-unicyclic graph with at least two non-trivial rooted trees. Define $\lambda_1, \ldots, \lambda_{c+1}$, and μ as in Notation 30.

We will apply the DNC relation on the edge $e = v_1 v_2$, where T_1 is a non-trivial tree. Note that $\lambda_{\mathrm{LC}}(G) = \mathrm{sort}(\lambda_1, \ldots, \lambda_{c+1}, \mu)$. In addition, we see that in $G \odot e$ (resp. $(G \odot e) \setminus \ell_e$), the vertices of the contracted edge are identified to a single vertex, say v_0, with $\lambda_1 + \lambda_2 - 1$ (resp. $\lambda_1 + \lambda_2 - 2$) leaves incident to v_0. To compute $\lambda_{\mathrm{lead}}(\mathbf{X}_{G \odot e})$ and $\lambda_{\mathrm{lead}}(\mathbf{X}_{(G \odot e) \setminus \ell_e})$, we discuss the following two cases.

1. When $G \odot e$ is c-unicyclic with at least two non-trivial rooted trees, so is $(G \odot e) \setminus \ell_e$. In this case, we apply our inductive hypothesis and get

$$\lambda_{\mathrm{lead}}(\mathbf{X}_{G \odot e}) = \lambda_{\mathrm{LC}}(G \odot e) = \mathrm{sort}(\lambda_1 + \lambda_2, \lambda_3, \ldots, \lambda_{c+1}, \mu)\,, \text{ and}$$

$$\lambda_{\mathrm{lead}}(\mathbf{X}_{(G \odot e) \setminus \ell_e}) = \lambda_{\mathrm{LC}}((G \odot e) \setminus \ell_e) = \mathrm{sort}(\lambda_1 + \lambda_2 - 1, 1, \lambda_3, \ldots, \lambda_{c+1}, \mu).$$

2. When $G \odot e$ is c-unicyclic with exactly one non-trivial rooted tree, so is $(G \odot e) \setminus \ell_e$. This happens exactly when T_1, T_2 are non-trivial and T_k is trivial for all $3 \leqslant k \leqslant c+1$. In this case, we have $\lambda_3 = \cdots = \lambda_{c+1} = 1$. Let $e' = v_0 v_3$. We apply Theorem 32(1) and get

$$\begin{aligned}
\lambda_{\mathrm{lead}}(\mathbf{X}_{G \odot e}) &= \lambda_{\mathrm{LC}}((G \odot e) \setminus e') \\
&= \mathrm{sort}(\lambda_1 + \lambda_2, \lambda_3 + \lambda_4, \lambda_5, \ldots, \lambda_c, \lambda_{c+1}, \mu) \\
&= \mathrm{sort}(\lambda_1 + \lambda_2, 2, 1^{c-3}, \mu)\,, \text{ and} \\
\lambda_{\mathrm{lead}}(\mathbf{X}_{(G \odot e) \setminus \ell_e}) &= \lambda_{\mathrm{LC}}(((G \odot e) \setminus \ell_e) \setminus e') \\
&= \mathrm{sort}(\lambda_1 + \lambda_2 - 1, 1, \lambda_3 + \lambda_4, \lambda_5, \ldots, \lambda_c, \lambda_{c+1}, \mu) \\
&= \mathrm{sort}(\lambda_1 + \lambda_2 - 1, 2, 1^{c-2}, \mu).
\end{aligned}$$

Note that in both cases we have $\lambda_{\mathrm{lead}}(\mathbf{X}_{(G \odot e) \setminus \ell_e}) < \lambda_{\mathrm{lead}}(\mathbf{X}_{G \odot e})$.

Finally we compute $\lambda_{\mathrm{lead}}(\mathbf{X}_{G \setminus e})$ and compare it with $\lambda_{\mathrm{lead}}(\mathbf{X}_{(G \odot e) \setminus \ell_e})$. We discuss the following two cases.

1. When T_2 is a non-trivial rooted tree, then the edge $v_2 v_3$ is internal in $G \setminus e$ and we have
$$\lambda_{\mathrm{lead}}(\mathbf{X}_{G \setminus e}) = \lambda_{\mathrm{LC}}(G) = \mathrm{sort}(\lambda_1, \ldots, \lambda_{c+1}, \mu) \leqslant \lambda_{\mathrm{lead}}(\mathbf{X}_{(G \odot e) \setminus \ell_e}).$$

2. When T_2 is a trivial-rooted tree, then the edge $v_2 v_3$ is a leaf-edge in $G \setminus e$. In this case, $\lambda_2 = 1$ and $(G \odot e) \setminus \ell_e$ is c-unicyclic with at least two non-trivial rooted trees. We have

$$\begin{aligned}
\lambda_{\mathrm{lead}}(\mathbf{X}_{G \setminus e}) &= \mathrm{sort}(\lambda_1, \lambda_2 + \lambda_3, \lambda_4, \ldots, \lambda_{c+1}, \mu) = \mathrm{sort}(\lambda_1, 1 + \lambda_3, \lambda_4, \ldots, \lambda_{c+1}, \mu) \\
&> \mathrm{sort}(\lambda_1, 1, \lambda_3, \ldots, \lambda_{c+1}, \mu) = \lambda_{\mathrm{LC}}(G) = \mathrm{sort}(\lambda_1 + \lambda_2 - 1, 1, \lambda_3, \ldots, \lambda_{c+1}, \mu) \\
&= \lambda_{\mathrm{lead}}(\mathbf{X}_{(G \odot e) \setminus \ell_e}).
\end{aligned}$$

Therefore the leading partition of G must be $\lambda_{\mathrm{LC}}(G)$. $\square$

The following corollary is evident from the proof of Theorem 40, and will be used in the next section.

Corollary 41. *When $r \geqslant 2$ and DNC relation is applied to an edge e which is incident to one trivial tree root and one non-trivial tree root, then $\lambda_{\text{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) < \lambda_{\text{lead}}(\mathbf{X}_{G\odot e})$ and $\lambda_{\text{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) < \lambda_{\text{lead}}(\mathbf{X}_{G\setminus e})$.*

Example 42. Consider the graph G from Theorem 31. Let $e = v_1v_2$. Recall that $\lambda = (3,2,1,2)$, $\mu = (3,2,3,1,2)$. We compute that

$$\lambda_{\text{LC}}(G) = (3,3,3,2,2,2,2,1,1) = \lambda_{\text{lead}}(\mathbf{X}_{G\setminus e}) < \lambda_{\text{lead}}(\mathbf{X}_{(G\odot e)\setminus \ell_e}) = (4,3,3,2,2,2,1,1,1) \\ < \lambda_{\text{lead}}(\mathbf{X}_{G\odot e}) = (5,3,3,2,2,2,1,1).$$

Corollary 43. *Let λ_{lead} be the leading partition in the star-expansion of $\mathbf{X}_G$ where G is a connected unicyclic graph with a cycle of size c. If G has exactly 1 non-trivial rooted tree, then the number of leaves in G is equal to $|\lambda_{\text{lead}}| - \ell(\lambda_{\text{lead}}) - 1$; if G has at least 2 non-trivial rooted trees, then the number of leaves in G is equal to $|\lambda_{\text{lead}}| - \ell(\lambda_{\text{lead}})$.*

Proof. It follows from the proofs of Theorem 32(1) and Theorem 32(2) that

$$\lambda_{\text{lead}}(\mathbf{X}_G) = \begin{cases} \text{sort}(\lambda_1, \dots, \lambda_{c-2}, \lambda_{c-1} + \lambda_c, \mu) & \text{if } r = 1, \\ \text{sort}(\lambda_1, \dots, \lambda_c, \mu) & \text{if } r > 1, \end{cases}$$

where we recall $\lambda_i - 1$ is the number of leaves attached to v_i on the cycle and μ is the concatenation of leaf components not containing the cycle vertices. □

5.3 Application to cuttlefish graphs

Recall that a squid is a connected unicyclic graph with exactly one vertex of degree greater than 2. We define *cuttlefish* to be squids such that all tentacles are of length 1. Let $\mathcal{U}_n$ denote the set of connected unicyclic graphs on n vertices, and let $\mathcal{U}_{c,n}$ denote the set of such graphs whose cycles are of order c.

Proposition 44. *The chromatic symmetric function can distinguish cuttlefish from among all connected unicyclic graphs on n vertices.*

Proof. Let $C_{c,t}$ denote the cuttlefish on $c+t$ vertices which is constructed by attaching t leaves to a single vertex of a c-cycle. Since by Corollary 11 we can obtain the size of the cycle of a connected unicyclic graph from the coefficient of $c_{(n)}$, it is sufficient to show that $\mathbf{X}_G$ distinguishes $C_{c,t}$ from other graphs in $\mathcal{U}_{c,c+t}$. Note that C_n is the only graph in $\mathcal{U}_{c,c}$, so we henceforth suppose $t > 0$.

By Theorem 32(1), the leading partition of $\mathbf{X}_{C_{c,t}}$ in the star-basis is $\lambda_{\text{lead}} = (t+1, 2, 1^{c-3})$. We will show that $C_{c,t}$ is the only graph in $\mathcal{U}_{c,n}$ that has this leading partition.

Let $G \in \mathcal{U}_{c,c+t}$ with $G \not\cong C_{c,t}$. If $r = 1$, then by Theorem 32(1), $\lambda_{\text{lead}}(\mathbf{X}_G) = (\lambda_1, 2, \mu, 1^{c-3})$ where $\lambda_1 - 1$ is the number of leaves attached to the root and μ is the leaf components in the remaining part of the tree. If $\lambda_{\text{lead}}(\mathbf{X}_G) = (t+1, 2, 1^{c-3})$, then μ is empty and $\lambda_1 - 1 = t$, making $G = C_{c,t}$. If $r \geqslant 2$, then each non-trivial tree has at least one leaf component of size at least two. To match the first two entries of $\lambda_{\text{lead}}(\mathbf{X}_{C_{c,t}})$, the only possibility is $r = 2$ where one of the trees is St_{t+1}, another is St_2, and the rest are trivial. But now there are $c-2$ cycle vertices remaining as their own leaf components, so $|G| = c + t + 1$, which is a contradiction. This completes the proof. □

Remark 45. Martin, Morin, and Wagner show in [11, Theorem 12] that no two squids have the same CSF. Note that while the class of squids is larger than that of cuttlefish, their result differs from the proposition above in that they do not distinguish squids from among all unicyclic graphs. On the other hand, both results implicitly give a procedure for reconstructing the graph from the CSF when the ambient class of graphs is known.

6 Leading coefficients for unicylic graphs

In this section, we provide formulas for the leading coefficient of any unicyclic graph of order n. The formulas depend on the number of rooted trees in the unicyclic graph, denoted by r. In the case where $r = 0$, the leading coefficient is $(-1)^{n-2}$, as discussed following Proposition 22. We now recall a result from [6] on the leading coefficient for trees, which will be useful in our extensions to the unicyclic case. Recall that a vertex is called deep if it is internal and has no leaf as a neighbor.

Theorem 46 ([6, Theorem 4.28]). *Suppose that T is a tree of order t with deep vertices $\{u_1, \ldots, u_p\}$ such that $d_i = \deg(u_i)$ for all i. Then, the leading coefficient of $\mathbf{X}_T$ is*

$$(-1)^p \prod_{i=1}^{p} (d_i - 1).$$

Theorem 46 can be reformulated using the elementary symmetric functions, see Section 2.2.

Remark 47. *Elementary symmetric polynomials* in a finite number of variables are defined analogously to the elementary symmetric functions. Observe that we have the identity

$$\prod_{i=1}^{p} (d_i - 1) = \sum_{j=0}^{p} (-1)^j e_{p-j}(d_1, \ldots, d_p).$$

That is, the leading coefficient is an alternating sum of elementary symmetric polynomials in the degrees of the deep vertices.

6.1 One non-trivial rooted tree ($r = 1$)

We first give a formula for the leading coefficient of any unicyclic graph when $r = 1$ depending on whether or not the root is a sprout, that is, a deep vertex on the cycle of degree at least 3. In this subsection, we will use ε_i's to denote edges since we will use e_i's for the elementary symmetric functions.

Theorem 48. *Let G be a connected unicyclic graph with a cycle of size c and exactly one non-trivial rooted tree T. Let v be the root of T. Label the edges on the cycle within G with edges ε_1 and ε_c incident to v, so that the labels increase $\varepsilon_2, \varepsilon_3, \ldots, \varepsilon_c$ in a clockwise direction. Consider the tree T' obtained by deleting edge ε_1. Suppose that T' has p deep vertices $\{u_1, \ldots, u_p\}$ such that $d_i = \deg(u_i)$ for all i.*

1. *If the root $v = u_1$ is a sprout in G, then*

$$c_{\lambda_{\text{lead}}} = (-1)^p \left[(c-2) \prod_{i=1}^{p} (d_i - 1) + \prod_{i=2}^{p} (d_i - 1) \right].$$

2. If the root $v = u_1$ is not a sprout in G, then

$$c_{\lambda_{\text{lead}}} = (-1)^p(c-2)\prod_{i=1}^{p}(d_i - 1).$$

Proof. First, we consider case (1) by identifying graphs from the leaves in the DNC-tree that have the same λ_{lead} as $\mathbf{X}_G$. By the fact that the star-expansion algorithm provides a cancellation-free formula, it suffices to determine the contribution to $c_{\lambda_{\text{lead}}}$ from each of these graphs. We can ignore the terms obtained through a leaf-contraction of ε_1 since, by Theorem 36, it will not have the same leading partition as $\mathbf{X}_G$. Upon deleting ε_1, we obtain the tree T' where $\lambda_{\text{lead}}(\mathbf{X}_{T'}) = \lambda_{\text{lead}}(\mathbf{X}_G)$ by Theorem 32(1). By Theorem 46, the contribution from the deletion of ε_1 to $c_{\lambda_{\text{lead}}}$ is

$$(-1)^p\prod_{i=1}^{p}(d_i - 1).$$

Now, consider $G' = (G \odot \varepsilon_1) \setminus \ell_{\varepsilon_1}$. If $c = 3$, then G' is a tree together with an isolated vertex and v is no longer a deep vertex; it has one leaf incident to it. Hence, the cycle's vertices contribute one part of size 2 (leaf component containing v) and one part of size 1 to λ_{lead} according to Theorem 32(1), and the other leaf components of the graph are unaffected. Thus, the resulting forest has the same leading partition as $\mathbf{X}_G$, and its contribution to $c_{\lambda_{\text{lead}}}$ is

$$-(-1)^{p-1}\prod_{i=2}^{p}(d_i - 1),$$

by Theorem 46, with the additional minus sign coming from the dot-contraction in the DNC relation applied to ε_1. Combining with the above gives the claimed formula when $c = 3$.

If $c > 3$, G' is a unicyclic graph with one non-trivial rooted tree and an isolated vertex, and by Theorem 32(1) $\lambda_{\text{lead}}(G') = \lambda_{\text{LC}}(G' \setminus \varepsilon_c)$. Letting v' be the other vertex on the edge ε_1, then G' is obtained from G by deleting ε_1, identifying v and v' into a single vertex v_0 (which is connected to the tree), and adding an isolated vertex v^*. We show that $\lambda_{\text{lead}}(G') = \lambda_{\text{LC}}(G' \setminus \varepsilon_c) = \lambda_{\text{LC}}(G \setminus \varepsilon_c) = \lambda_{\text{lead}}(G)$. First we note the leaf components for the non-trivial rooted tree (including v in $G \setminus \varepsilon_c$ and v_0 in $G' \setminus \varepsilon_c$) are equal in $G' \setminus \varepsilon_c$ and $G \setminus \varepsilon_c$. Second, for the vertices on the cycle (excluding v in $G \setminus \varepsilon_c$ and v_0 in $G' \setminus \varepsilon_c$), $G \setminus \varepsilon_c$ has leaf components of orders $(2, 1^{c-3})$ and $G' \setminus \varepsilon_c$ has leaf components of orders $(2, 1^{c-4})$ from the cycle vertices. Collecting these together with the isolated vertex in G', we see that $\lambda_{\text{lead}}(G') = \lambda_{\text{lead}}(G)$.

To get $c_{\lambda_{\text{lead}}}$, we must continue down the DNC-tree by next applying the relation to ε_2 in G'. Again, graphs arising from leaf-contraction will have CSFs with lexicographically larger leading partitions, so we can ignore them. Deletion of the edge leads to a forest with the same λ_{lead} and $c_{\lambda_{\text{lead}}}$ as $\mathbf{X}_{T'}$ but with one fewer deep vertices of degree two. Together with the sign contributed by the previous dot-contraction, this yields another term contributing

$$(-1)^p\prod_{i=1}^{p}(d_i - 1).$$

To G', we continue to apply the DNC relation to the edges $\varepsilon_3, \varepsilon_4, \ldots, \varepsilon_{c-2}$ in consecutive order. At each step, we gain $(-1)^p \prod_{i=1}^{p}(d_i - 1)$ through deletion. In the last dot-contraction, we gain

$$(-1)^{c-2}(-1)^{p-(c-2)} \prod_{i=2}^{p}(d_i - 1).$$

This term is analogous to the $c = 3$ case, where $(-1)^{c-2}$ represents the number of dot-contractions and $(-1)^{p-(c-2)}$ corresponds to the remaining deep vertices at the final step. Combining all these DNC-tree contributions gives the claimed formula for case (1).

The proof for case (2) follows a similar approach to case (1). Suppose v is not deep in G. Then, the deletion of ε_1 produces a tree T' with leading coefficient exactly

$$(-1)^p \prod_{i=1}^{p}(d_i - 1).$$

If $c = 3$, both dot-contraction and leaf-contraction of ε_1 increase the size of the leaf component containing v by one, so their leading partitions are different from $\mathbf{X}_{T'}$ and thus do not contribute to λ_{lead}. On the other hand, if $c > 3$, dot-contraction does not increase the size of the leaf component containing v, and we may continue applying the DNC relations on $\varepsilon_2, \varepsilon_3, \ldots, \varepsilon_{c-2}$. Each deletion contributes the same term as above to $c_{\lambda_{\text{lead}}}$, occurring a total of $(c-2)$ times. The final dot-contraction on edge ε_{c-2} does not contribute to $c_{\lambda_{\text{lead}}}$, as it will change the leading partition of the CSF of the resulting forest, just as in the $c = 3$ case. □

Example 49. Consider G as in Theorem 34 and Figure 10, which has one non-trivial rooted tree and v_1 is deep in $G \setminus e$. We get $d_1 = 2$, $p = 1$, and $c = 3$. We compute

$$c_{\lambda_{\text{lead}}} = (-1)^1 \left[(3-2) \cdot 1 + 1\right] = -2\,.$$

This is consistent with

$$\mathbf{X}_G = 2\mathfrak{st}_{(6)} - 4\mathfrak{st}_{(5,1)} + \mathfrak{st}_{(4,2)} + 2\mathfrak{st}_{(4,1,1)} + 2\mathfrak{st}_{(3,3)} - 2\mathfrak{st}_{(3,2,1)}.$$

6.2 More than one non-trivial rooted tree ($r \geqslant 2$)

Here, the formulas again depend on whether or not the root vertices of non-trivial trees on the cycle are deep or not.

Definition 50. Let G be a unicyclic graph with more than one non-trivial tree at distinct cycle vertices. Let the set of sprouts be denoted $v_1, \ldots, v_s$, and their degrees $b_i = \deg(v_i)$. Let the non-sprout deep vertices be denoted $u_1, \ldots, u_p$ with $d_i = \deg(u_i)$ as before.

Proposition 51. *Let G be a connected unicyclic graph with no sprouts and $r \geqslant 2$. Then*

$$c_{\lambda_{\text{lead}}} = (-1)^p \prod_{i=1}^{p}(d_i - 1).$$

Proof. If there is a cycle edge incident to two non-trivial tree roots, then applying DNC to that edge shows that the only contribution to the leading coefficient will come from the tree obtained by deletion. If there is no such edge, then the only contribution to the leading coefficient will come from dot-contraction, so we dot-contract until we have two adjacent non-trivial tree roots and then same argument applies. The sign is the same in both cases. □

Note that unlike the $r = 1$ case, deletion of edges that are not incident to two roots of non-trivial trees no longer produces a tree with the same leading partition, which explains the absence of the factor c in the formula of Theorem 51. If we have at least one sprout, then our formula will depend on both the number of sprouts, which we denote by s, as well as r, the number of non-trivial trees in total.

Theorem 52. *Let b_i $(1 \leqslant i \leqslant s)$ be the degrees of sprout vertices, and d_i $(1 \leqslant i \leqslant p)$ be the degrees of non-sprout deep vertices of a connected unicyclic G with $r \geqslant 2$ and $c \geqslant 3$. Then*

$$c_{\lambda_{\text{lead}}} = (-1)^{p+s} \begin{cases} \displaystyle\prod_{i=1}^{p}(d_i - 1) \left[\left(\prod_{i=1}^{s}(b_i - 1)\right) - \sum_{i=1}^{s} b_i + 2s - 1\right] & \textit{if } r = s, \\ \displaystyle\prod_{i=1}^{p}(d_i - 1) \left[\left(\prod_{i=1}^{s}(b_i - 1)\right) - 1\right] & \textit{if } r = s + 1, \\ \displaystyle\prod_{i=1}^{p}(d_i - 1) \prod_{i=1}^{s}(b_i - 1) & \textit{if } r \geqslant s + 2. \end{cases}$$

Proof. We proceed by induction on r. *Base cases ($r = 2, 3$):* First suppose that $r = 2$. If $c > 3$, this means that there is a cycle edge e incident to the root of one trivial tree and one non-trivial tree. Since $r = 2$, by Theorem 32 (2) $\lambda_{\text{lead}}(\mathbf{X}_G) = \lambda_{\text{LC}}(G)$. Then, by Theorem 41, both deletion and leaf-contraction of e will produce graphs whose CSFs will have larger leading partitions, meaning we can ignore these terms for the purpose of computing $c_{\lambda_{\text{lead}}}(\mathbf{X}_G)$ and consider only the CSF of the graph obtained by the dot-contraction of e. By Theorem 4, the leading partition of $\mathbf{X}_G$ and $\mathbf{X}_{(G \odot e) \setminus \ell_e}$ are the same, and we may repeat this argument on any cycle edge of $(G \odot e) \setminus \ell_e$, the dot-contracted graph, which is incident to a cycle vertex of degree two and the root of a non-trivial tree. Thus we continue to dot-contract edges of G until we obtain a graph which consists of the union of a unicyclic component G' which contains a triangle ($c' = 3$) and two non-trivial trees, together with some isolated vertices. The only impact of these dot-contractions on $c_{\lambda_{\text{lead}}}$ is the sign, which is accounted for in the factor $(-1)^p$ in the formula above. This is because we will have dot-contracted $c - 3$ times to obtain a triangle, and removed the same number of non-sprout deep vertices from G in the process.

Let e' be the edge of G' between the roots of the non-trivial trees. The case $s = 0$ is handled by Theorem 51. If $s = 1$, then both dot-contraction and leaf-contraction of e' give different leading partitions than $\lambda_{\text{lead}}(\mathbf{X}_G)$, so we can just consider $c_{\lambda_{\text{lead}}}(\mathbf{X}_{G' \setminus e'})$. This graph is now a tree for which all of the non-sprout deep vertices have the same degree as in G, and the unique vertex that was a sprout with degree b_1 in G now has degree $b_1 - 1$. Applying the tree formula of Theorem 46 to $G' \setminus e'$ recovers the second case of the claimed formula above. Finally if $s = 2$, we can again ignore dot-contraction and leaf-contraction, but now deletion of e' gives that the sprout vertices enter into the tree formula as a factor of

$$((b_1-1)-1)((b_2-1)-1) = (b_1-1)(b_2-1)-(b_1+b_2)+3 = (b_1-1)(b_2-1)-(b_1+b_2)+2s-1$$

as claimed, completing the case $r = 2$. Now suppose $r = 3$. As in the $r = 2$ case, we may dot-contract until we obtain G' whose connected unicyclic component consists of a triangle with the same three rooted trees attached and $|c_{\lambda_{\text{lead}}}(\mathbf{X}_{G'})| = |c_{\lambda_{\text{lead}}}(\mathbf{X}_G)|$. If $s \leqslant 1$, then we let e' the edge between two non-sprout roots, and the deletion term of DNC applied to e' gives the only contribution to $c_{\lambda_{\text{lead}}}$. As $G' \setminus e'$ is a tree, the claimed formula holds.

Next let $s = 2$, and let e' be the edge between sprouts v_1 and v_2 in G'. Applying DNC to e', we see that both deletion and dot-contraction terms contribute to $c_{\lambda_{\text{lead}}}$. For brevity, we only discuss the "sprout factors" in our formula, as the degrees of the other deep vertices of the graph are unchanged by DNC operations on cycle edges. Deletion yields a tree with sprout factor $(b_1-2)(b_2-2)$, while dot-contraction merges the sprouts into a single deep vertex of degree b_1+b_2-3 since the triangle collapses, yielding another tree. Adding the sprout factor contributions of these trees gives

$$b_1b_2 - 2(b_1+b_2) + 4 + (b_1+b_2-4) = (b_1-1)(b_2-1) - 1$$

as claimed.

If $s = 3$, suppose we apply DNC to the edge between sprouts v_1 and v_2. Again we must consider the deletion and dot-contraction terms. Now, the sprout factor contributions of the corresponding trees combine to give

$$(b_1-2)(b_2-2)(b_3-1) + (b_1+b_2-4)(b_3-2) = b_1b_2b_3 - (b_1b_2 + b_1b_3 + b_2b_3) + 4.$$

We see that this sprout factor is the same as

$$(b_1-1)(b_2-1)(b_3-1) - (b_1+b_2+b_3) + 5$$

as claimed. This completes the base cases.

Inductive step: Now suppose G has $r > 3$, and that our formula for $c_{\lambda_{\text{lead}}}$ holds for unicyclic graphs with $r-1$ non-trivial trees. As in the $r = 3$ case, by Theorem 4 when $c > r$ we may apply dot-contraction to edges of the cycle until we obtain a unicyclic graph G' with cycle size $c' = r$ and such that $|c_{\lambda_{\text{lead}}}(\mathbf{X}_{G'})| = |c_{\lambda_{\text{lead}}}(\mathbf{X}_G)|$.

We consider three cases for the graph G'.

(a) There are two adjacent non-sprout cycle vertices.

(b) There are two adjacent sprouts.

(c) Every pair of neighbors on the cycle consists of sprout and a non-sprout.

Case (a): This case is easiest. First notice that this assumption implies $r \geqslant s+2$. Let e' be an edge of G' between two non-sprouts, and apply DNC to this edge. As above, we only need to consider the contribution of $G' \setminus e'$ which is now a tree in which all deep vertices have the same degree as they do in G. Thus, we recover the third case of our claimed formula for $c_{\lambda_{\text{lead}}}$.

Case (b): Let e' be an edge between sprouts v_1 and v_2. The deletion of e' produces a tree with the same leading partition and sprout factor contribution

$$(b_1 - 2)(b_2 - 2) \prod_{i=3}^{s} (b_i - 1), \tag{4}$$

while dot-contraction produces a unicyclic graph G'' with number of non-trivial trees $r'' = r - 1$ and number of sprouts $s'' = s - 1$.

If $r = s$, then $r'' = s''$ and the merging of v_1 and v_2 produces a sprout of degree $b_1 + b_2 - 2$. Hence, by the inductive hypothesis, the sprout factor contribution of G'' is

$$(b_1 + b_2 - 3) \left(\prod_{i=3}^{s} (b_i - 1) \right) - (b_1 + b_2 - 2 + b_3 + \cdots + b_s) + 2s - 3.$$

Adding with the contribution of $G' \setminus e'$ gives the desired formula.

If $r = s + 1$, the sprout factor contribution of $G' \setminus e'$ is unchanged. On the other hand, dot-contraction gives a sprout factor of

$$(b_1 + b_2 - 3) \left(\prod_{i=3}^{s} (b_i - 1) \right) - 1,$$

by the inductive hypothesis. Adding the quantity we obtained in (4) gives the desired result.

Finally, if $r \geqslant s + 2$ then G'' contributes sprout factor

$$(b_1 + b_2 - 3) \prod_{i=3}^{s} (b_i - 1),$$

and again adding with (4) gives the result.

Case (c): Note that the condition of this case together with $r \geqslant 4$ implies that $r \geqslant s+2$. Let e' be an edge between sprout v_1 and another cycle vertex which is adjacent to v_1. We again consider sprout factor contributions of the deletion and dot-contraction terms of DNC applied to e'. $G' \setminus e'$ has a sprout factor of

$$(b_1 - 2) \prod_{i=2}^{s} (b_i - 1)$$

while the dot-contraction gives a graph G'' with $r'' = r - 1$, $s'' = s - 1$ to which we apply the inductive hypothesis. As $r'' \geqslant s'' + 2$, the sprout factor is just

$$\prod_{i=2}^{s} (b_i - 1),$$

and adding these terms gives the result. This concludes this case, the inductive step, and the proof of the theorem. □

Example 53. Consider G in Figure 2, which has $r = 3$ non-trivial rooted trees. The only sprout is v_3 with degree 3 and the non-sprout deep vertices are v_4 and v_{11}, both with degree 2. So $s = 1, p = 2$ and we get

$$c_{\lambda_{\text{lead}}} = (-1)^{2+1} \cdot 1 \cdot 1 \cdot 2 = -2\,.$$

7 Bicyclic graphs

A simple connected graph G is said to be *bicyclic* if the number of edges is equal to the number of vertices plus one. Let $\mathcal{B}_n$ denote the set of connected bicyclic graphs of order n. For any $G \in \mathcal{B}_n$, the unique minimal bicyclic subgraph $\tilde{G}$ falls into one of two cases:

1. $\tilde{G}$ is constructed from two disjoint cycles C_s and C_t by adding a path P_h (with $h \geqslant 1$) from a vertex in C_s to any vertex in C_t. In the case where $h = 1$, we have two cycles that share a single vertex.

2. $\tilde{G}$ is constructed from three paths whose edge sets are disjoint, $P_{k+1}, P_{\ell+1}, P_{m+1}$, where $k, \ell, m \geqslant 1$ and at most one of them is equal to 1. The initial vertices of all three paths are identified as a single vertex u, and the terminal vertices are identified as a single vertex v. In other words, $\tilde{G}$ can be regarded as consisting of two cycles $C_{\ell+k}$ and $C_{\ell+m}$ which share a common path of $P_{\ell+1}$.

For any $G \in \mathcal{B}_n$, we say that G is a *type-I* (resp. *type-II*) bicyclic graph if its unique minimal bicyclic subgraph is of the form of case (1) (resp. case (2)) above. We illustrate these two types in Figure 12.

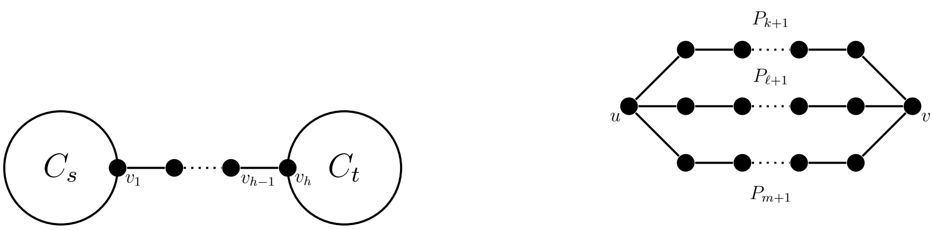

Figure 12: Left: a type-I bicyclic graph; right: a type-II bicyclic graph.

Example 54. In Figure 13 we illustrate a type-I bicyclic graph in $\mathcal{B}_{16}$. Its unique minimal bicyclic subgraph has cycles of size 4 and 5 connected by a path with $h = 3$.

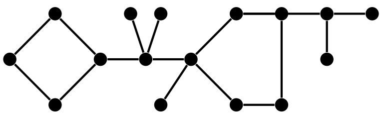

Figure 13: A type-I bicyclic graph in $\mathcal{B}_{16}$

Proposition 55. *Let G be a bicyclic graph on n vertices.*

1. *If G is a type-I bicyclic graph formed by cycles of length $s, t \geqslant 3$, then*
$$c_{(n)} = (s-1)(t-1).$$

2. *If G is a type-II bicyclic graph formed by two cycles of length $s, t \geqslant 3$ which meet along a path $P_{\ell+1}$ for $\ell \geqslant 1$, then*
$$c_{(n)} = (s-1)(t-1) - 2\binom{\ell}{2}.$$
Note that when $\ell = 1$, the formula reduces to $c_{(n)} = (s-1)(t-1)$.

Proof. First we consider case (1). We apply the DNC relation to an edge of the cycle containing s vertices. The term corresponding to deletion of the edge contributes $t-1$ to the coefficient $c_{(n)}$, by Theorem 11. Next, the dot-contraction term can be ignored since it produces a disconnected graph, for which $c_{(n)} = 0$ since all partitions from disconnected graphs have length at least 2. We now proceed by induction on s for the leaf-contraction term. When $s = 3$, the leaf-contraction term also contributes $(t-1)$ yielding $2(t-1)$ total for $c_{(n)}$, as claimed. Now if $s > 3$, the leaf-contraction term reduces us to the case of cycles of size $s-1$ and t, which contributes $(s-2)(t-1)$ by the inductive hypothesis. Finally, adding this to the contribution of deletion produces the result.

Next we consider case (2) where the cycles of size s, t have $\ell \geqslant 1$ edges in common. We proceed with induction on ℓ. For the base case, suppose $\ell = 1$ and e is the single shared edge. We apply the DNC relation to e. If $s = 3$, deleting the edge e yields a unicyclic graph with a cycle of size $2 + (t-1) = t+1$, which by Theorem 11 contributes t to $c_{(n)}$. If $t = 3$, leaf-contracting e would result in a tree which contributes $1 = t-2$ to $c_{(n)}$ by Theorem 8. If $t \geqslant 4$, leaf-contracting e would produce a unicyclic graph with a cycle of size $t-1$, which contributes $t-2$ to $c_{(n)}$. We may again ignore the dot-contraction term. Thus by the DNC relation, $c_{(n)} = t + t - 2 = 2(t-1)$. For $s > 3$, applying DNC to e now gives $s+t-3$ from deletion, and $(s-2)(t-2)$ from leaf-contraction by applying the formula for case (1). Adding these contributions yields the result. Supposing the result holds whenever the cycles have $\ell - 1$ edges in common, we apply DNC to an edge of the common path. Deletion contributes $s+t-2\ell-1$, since we are left with one cycle of length $s+t-2\ell$. Leaf-contraction reduces to the $\ell - 1$ case where our inductive hypothesis has contribution

$$(s-2)(t-2) - 2\binom{\ell-1}{2} = st - 2s - 2t + 4 - (\ell-1)(\ell-2).$$

Adding these contributions, we have

$$st - t - s + 1 - 2(\ell-1) - (\ell-1)(\ell-2) = (s-1)(t-1) - \ell(\ell-1)$$

as desired. □

In what follows, we will show how to recover the sizes of the two cycles in a bicyclic graph. In order to do this we need another equation in addition to $(s-1)(t-1) = c_{(n)}$ in type-I, and we need two additional equations in type-II as we have three unknowns. We begin with a proposition.

Proposition 56. *[13, Lemma 4.19] Let $j \geqslant 1$, G be a unicyclic graph with a cycle of length $c \geqslant 3$ and chromatic symmetric function $\mathbf{X}_G = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$. Then,*

$$\sum_{\substack{\lambda \vdash n \\ \ell(\lambda) = j}} c_\lambda = (-1)^{j-1}\binom{c-1}{j}.$$

In particular, these sums are zero whenever $j \geqslant c$.

We have some similar results for bicyclic graphs.

Proposition 57. *Let G be a connected bicyclic graph on n vertices of type-I containing cycles of length $t \geqslant s \geqslant 3$. Let $\boldsymbol{X}_G = \sum_{\lambda \vdash n} c_\lambda \mathfrak{st}_\lambda$. Then for $j \geqslant 1$,*

$$\sum_{\lambda \vdash n,\ \ell(\lambda)=j} c_\lambda = (-1)^{j-1} \sum_{i=1}^{j} \binom{s-1}{i}\binom{t-1}{j-i+1} = (-1)^{j-1}\left[\binom{s+t-2}{j+1} - \binom{t-1}{j+1} - \binom{s-1}{j+1}\right].$$

In particular, the sum vanishes when $j \geqslant s+t-2$ and is non-zero for $1 \leqslant j \leqslant s+t-3$.

Proof. We induct on the length of the cycle s. For the base case, suppose $s = 3$. We apply the DNC relation to an edge e on the s-cycle, we have that $G \setminus e$ and $G \odot e$ are both connected unicyclic graphs on n vertices containing a t-cycle, while $(G \odot e) \setminus \ell_e$ is the disjoint union of a connected unicyclic graph on $n-1$ vertices and an isolated vertex. Hence, by Theorem 56, $\mathbf{X}_{G\setminus e}$ and $\mathbf{X}_{G\odot e}$ each contribute $(-1)^{j-1}\binom{t-1}{j}$ to the sum, and $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ contributes 0 to the sum when $j = 1$, and contributes $(-1)^{j-2}\binom{t-1}{j-1}$ when $j > 1$. Therefore, by the DNC relation, $\sum_{\lambda\vdash n,\ \ell(\lambda)=j} c_\lambda = (-1)^{j-1}2\binom{t-1}{j} = (-1)^{j-1}\binom{3-1}{1}\binom{t-1}{j}$ when $j = 1$ and when $j > 1$,

$$\begin{aligned}\sum_{\lambda \vdash n,\ \ell(\lambda)=j} c_\lambda &= (-1)^{j-1}\binom{t-1}{j} - (-1)^{j-2}\binom{t-1}{j-1} + (-1)^{j-1}\binom{t-1}{j} = (-1)^{j-1}\left[2\binom{t-1}{j} + \binom{t-1}{j-1}\right] \\ &= (-1)^{j-1}\left[\binom{3-1}{1}\binom{t-1}{j} + \binom{3-1}{2}\binom{t-1}{j-1}\right].\end{aligned}$$

This completes the base case. For the inductive step, assume that the statement holds for all such bicyclic graphs of type-I with cycles of size $t, z-1 \geqslant 3$. Let G be a bicyclic graph on n vertices with cycles of size $t, z \geqslant 3$ sharing at most one vertex. By applying the DNC relation to an edge e on the z-cycle, we have that $G \setminus e$ is a connected unicyclic graph on n vertices containing a t-cycle; $(G \odot e) \setminus \ell_e$ is the disjoint union of a bicyclic graph (of type-I) on $n-1$ vertices containing cycles of sizes t and $z-1$ and an isolated vertex; and $G \odot e$ is a bicyclic graph on n vertices (of type-I) containing cycles of size t and $z-1$. Hence, by Theorem 56, $\mathbf{X}_{G\setminus e}$ contributes $(-1)^{j-1}\binom{t-1}{j}$ to the sum; $\mathbf{X}_{(G\odot e)\setminus\ell_e}$ contributes $(-1)^{j-2}\sum_{i=1}^{j-1}\binom{z-2}{i}\binom{t-1}{j-i}$ by the inductive hypothesis; and, again by the inductive hypothesis, $\mathbf{X}_{G\odot e}$ contributes $(-1)^{j-1}\sum_{i=1}^{j}\binom{z-2}{i}\binom{t-1}{j-i+1}$. Therefore, we have that

$$\begin{aligned}\sum_{\lambda \vdash n,\ \ell(\lambda)=j} c_\lambda &= (-1)^{j-1}\binom{t-1}{j} - (-1)^{j-2}\sum_{i=1}^{j-1}\binom{z-2}{i}\binom{t-1}{j-i} + (-1)^{j-1}\sum_{i=1}^{j}\binom{z-2}{i}\binom{t-1}{j-i+1} \\ &= (-1)^{j-1}\left[\binom{t-1}{j} + \sum_{i=1}^{j-1}\binom{z-2}{i}\binom{t-1}{j-i} + \sum_{i=1}^{j}\binom{z-2}{i}\binom{t-1}{j-i+1}\right] \\ &= (-1)^{j-1}\Bigg[\binom{t-1}{j} + \sum_{i=1}^{j-2}\binom{z-2}{i}\binom{t-1}{j-i} + \binom{z-2}{j-1}\binom{t-1}{1} + \binom{z-2}{1}\binom{t-1}{j} \\ &\qquad + \sum_{i=2}^{j-1}\binom{z-2}{i}\binom{t-1}{j-i+1} + \binom{z-2}{j}\binom{t-1}{1}\Bigg].\end{aligned}$$

Reindex the last sum, combine with terms of the second, and apply Pascal's rule to get

$$\begin{aligned}\sum_{\lambda\vdash n,\ \ell(\lambda)=j} c_\lambda &= (-1)^{j-1}\left[\tbinom{z-1}{1}\tbinom{t-1}{j} + \sum_{i=1}^{j-2}\left(\tbinom{z-2}{i}+\tbinom{z-2}{i+1}\right)\tbinom{t-1}{j-i} + \tbinom{z-2}{j-1}\tbinom{t-1}{1} + \tbinom{z-2}{j}\tbinom{t-1}{1}\right]\\ &= (-1)^{j-1}\left[\tbinom{z-1}{1}\tbinom{t-1}{j} + \sum_{i=1}^{j-2}\tbinom{z-1}{i+1}\tbinom{t-1}{j-i} + \tbinom{z-1}{j}\tbinom{t-1}{1}\right]\\ &= (-1)^{j-1}\left[\sum_{i=1}^{j}\tbinom{z-1}{i}\tbinom{t-1}{j-i+1}\right].\end{aligned}$$

The result follows from applying Vandermonde's identity.

It is easy to see that if $j \geqslant s+t-2$, then all binomial coefficients $\binom{s+t-2}{j+1}$, $\binom{s-1}{j+1}$, and $\binom{t-1}{j+1}$ are zero. Thus, the sums vanish for these values of j. Now, we show that the summation is strictly positive for $1 \leqslant j \leqslant s+t-3$. Since

$$\tbinom{s+t-2}{j+1} = \sum_{i=0}^{j+1}\tbinom{s-1}{i}\tbinom{t-1}{j-i+1} = \sum_{i=1}^{j}\tbinom{s-1}{i}\tbinom{t-1}{j-i+1} + \tbinom{t-1}{j+1} + \tbinom{s-1}{j+1},$$

then $\binom{s+t-2}{j+1} > \binom{t-1}{j+1} + \binom{s-1}{j+1}$ if there is at least one i such that $\binom{s-1}{i}\binom{t-1}{j-i+1} > 0$. This happens when there is an i in the summation such that $1 \leqslant i \leqslant s-1$ and $1 \leqslant j-i+1 \leqslant t-1$, which is equivalent to $1 \leqslant i \leqslant s-1$ and $j-t+2 \leqslant i \leqslant j$. Thus, there is an i in both intervals if $1 \leqslant j$ and $j-t+2 \leqslant s-1$, in other words $1 \leqslant j \leqslant s+t-3$. $\square$

Proposition 58. *Let G be a connected bicyclic graph of type-II, formed by two cycles of length $s, t \geqslant 3$ which meet along a path $P_{\ell+1}$ for $\ell \geqslant 1$, and with chromatic symmetric function $\mathbf{X}_G = \sum_{\lambda\vdash n} c_\lambda \mathfrak{st}_\lambda$. Then*

$$\sum_{\substack{\lambda\vdash n\\ \ell(\lambda)=j}} c_\lambda$$
$$= (-1)^{j-1}\left[\sum_{i=1}^{\ell}\left[\binom{s+t-2(\ell+1)+i}{j}\right] + \binom{s+t-\ell-2}{j+1} - \binom{s-1}{j+1} - \binom{t-1}{j+1}\right].$$

In particular, the sum vanishes when $j \geqslant s+t-\ell-1$ and is non-zero for $1 \leqslant j < s+t-\ell-1$. We also note that the formula reduces to that of a type-I bicyclic graph when $\ell = 1$, and, although type-II bicyclic graphs assume $\ell \geqslant 1$, setting $\ell = 0$ also agrees with the type-I formula.

Proof. We proceed by induction on the length of the common path $P_{\ell+1}$ and assume without loss of generality that $s \leqslant t$. For the base case, suppose $\ell = 1$. If $s = 3$, we let e be an edge of the cycle containing s vertices that is not the single edge in $P_{\ell+1}$. We apply the DNC relation to e to obtain $\mathbf{X}_G = \mathbf{X}_{G\setminus e} - \mathbf{X}_{(G\odot e)\setminus \ell_e} + \mathbf{X}_{G\odot e}$ and consider $\sum_{\substack{\lambda\vdash n\\ \ell(\lambda)=j}} c_\lambda$ for each of the three graphs coming from the DNC relation. Notice that the deletion and leaf-contraction of e each result in a connected unicyclic graph with a cycle of length t and the dot-contraction of e results in a unicyclic graph with a cycle of length t together with an isolated vertex. For $j = 1$, $\mathbf{X}_{G\setminus e}$ and $\mathbf{X}_{G\odot e}$ each contribute $t-1$ to the sum by

Theorem 11 while $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ contributes 0 to the sum since it is disconnected. The DNC relation and Pascal's rule show

$$\begin{aligned} c_{(n)} &= 2t-2 = \binom{t}{1} + \binom{t-1}{1} - \binom{2}{2} \\ &= \binom{t}{1} + \binom{t}{2} - \binom{2}{2} - \binom{t-1}{2} \\ &= (-1)^0\binom{3+t-2(2)+1}{1} + \binom{3+t-1-2}{2} - \binom{3-1}{2} - \binom{t-1}{2}. \end{aligned}$$

For $j \geqslant 2$, Theorem 56 states that $\mathbf{X}_{G\setminus e}$ and $\mathbf{X}_{G\odot e}$ each contribute $(-1)^{j-1}\binom{t-1}{j}$ to the sum. With $\mathbf{X}_{(G\odot e)\setminus \ell_e}$, all partitions of length j result from multiplying those of length $j-1$ by $\mathfrak{st}_{(1)}$, and therefore by Theorem 56, $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ contributes $(-1)^{j-2}\binom{t-1}{j-1}$ to the sum. Since $j \geqslant 2$, it follows that $\binom{2}{j+1} = 0$. The DNC relation and Pascal's rule show

$$\begin{aligned} \sum_{\substack{\lambda \vdash n \\ \ell(\lambda)=j}} c_\lambda &= (-1)^{j-1}\left[\binom{t-1}{j} + \binom{t-1}{j-1} + \binom{t-1}{j} - \binom{2}{j+1}\right] \\ &= (-1)^{j-1}\left[\binom{t}{j} + \binom{t-1}{j} - \binom{2}{j+1}\right] \\ &= (-1)^{j-1}\left[\binom{t}{j} + \binom{t}{j+1} - \binom{2}{j+1} - \binom{t-1}{j+1}\right] \\ &= (-1)^{j-1}\left[\binom{3+t-2(2)+1}{j} + \binom{3+t-1-2}{j+1} - \binom{3-1}{j+1} - \binom{t-1}{j+1}\right]. \end{aligned}$$

If $s \geqslant 4$, we let e be the single shared edge. Notice that $G \setminus e$ is a connected unicyclic graph with cycle length $s+t-2$, $(G \odot e) \setminus \ell_e$ is a type-I graph with cycles of length $s-1$ and $t-1$ together with an isolated vertex, and $(G \odot e)$ is a connected type-I graph with cycles of size $s-1$ and $t-1$. By Theorem 56, $\mathbf{X}_{G\setminus e}$ contributes $(-1)^{j-1}\binom{s+t-3}{j}$ to the sum. By Theorem 57, $\mathbf{X}_{(G\odot e)\setminus \ell_e}$ contributes $(-1)^{j-2}\left[\binom{s+t-4}{j} - \binom{s-2}{j} - \binom{t-2}{j}\right]$ and $\mathbf{X}_{G\odot e}$ contributes $(-1)^{j-1}\left[\binom{s+t-4}{j+1} - \binom{s-2}{j+1} - \binom{t-2}{j+1}\right]$. Using these contributions in the DNC relation and Pascal's rule, we have

$$\begin{aligned} \sum_{\substack{\lambda \vdash n \\ \ell(\lambda)=j}} c_\lambda &= (-1)^{j-1}\left[\binom{s+t-3}{j} + \binom{s+t-4}{j} - \binom{s-2}{j} - \binom{t-2}{j} + \binom{s+t-4}{j+1} - \binom{s-2}{j+1} - \binom{t-2}{j+1}\right] \\ &= (-1)^{j-1}\left[\binom{s+t-3}{j} + \binom{s+t-3}{j+1} - \binom{s-1}{j+1} - \binom{t-1}{j+1}\right]. \end{aligned}$$

For the inductive step, suppose the statement holds for $\ell = w \geqslant 1$ and let G be as in the statement with $\ell = w+1$. Observe that we may assume that $s, t \geqslant 4$ since if there were a 3-cycle, we could rearrange the paths so that the common path has length 1, which is covered by the base case. Let e be an edge of the common path, P_{w+2}. As above, we apply the DNC relation on e and analyze the contributions to the sum from each of the three graphs. The deletion of e results in a unicyclic graph with cycle length $s+t-2(w+1)$ and thus $\mathbf{X}_{G\setminus e}$ contributes $(-1)^{j-1}\binom{s+t-2(w+1)-1}{j}$ to the sum. The dot-contraction of e results in a type-II bicyclic graph with cycles of size $s-1$ and $t-1$ that meets along a path of length w together with an isolated vertex and thus by assumption contributes

$$(-1)^{j-2}\left(\sum_{i=1}^{w}\left[\binom{(s-1)+(t-1)-2(w+1)+i}{j-1}\right]\right.$$

$$+\binom{(s-1)+(t-1)-w-2}{j}-\binom{s-2}{j}-\binom{t-2}{j}\Bigg).$$

Similarly, the leaf-contraction of e contributes

$$(-1)^{j-1}\Bigg(\sum_{i=1}^{w}\Bigg[\binom{(s-1)+(t-1)-2(w+1)+i}{j}\Bigg]$$
$$+\binom{(s-1)+(t-1)-w-2}{j+1}-\binom{s-2}{j+1}-\binom{t-2}{j+1}\Bigg).$$

Thus by the DNC relation and Pascal's rule,

$$\begin{aligned}\sum_{\substack{\lambda\vdash n\\ \ell(\lambda)=j}} c_\lambda = (-1)^{j-1}\Bigg[&\tbinom{s+t-2(w+1)-1}{j}\\ &+\sum_{i=1}^{w}\Big[\tbinom{s+t-2(w+2)+i}{j-1}\Big]+\tbinom{s+t-(w+1)-3}{j}-\tbinom{s-2}{j}-\tbinom{t-2}{j}\\ &+\sum_{i=1}^{w}\Big[\tbinom{s+t-2(w+2)+i}{j}\Big]+\tbinom{s+t-(w+1)-3}{j+1}-\tbinom{s-2}{j+1}-\tbinom{t-2}{j+1}\Bigg]\\ =(-1)^{j-1}\Bigg[&\tbinom{s+t-2(w+2)+1}{j}+\sum_{i=2}^{w+1}\Big[\tbinom{s+t-2(w+2)+i}{j}\Big]+\tbinom{s+t-(w+1)-2}{j+1}-\tbinom{s-1}{j+1}-\tbinom{t-1}{j+1}\Bigg],\end{aligned}$$

which after combining the first two terms as a single sum, we get the desired result.

It is easy to see that all binomials in the statement of the proposition vanish when $j \geqslant s+t-\ell-1$. If $\ell = 1$, then the formula reduces to type-I, hence the non-vanishing is proven in Proposition 57. In fact, if any of k, m, or ℓ are equal to 1, the formula reduces to that of type-I. Hence, we assume that $k, m, \ell \geqslant 2$ (where k, m, ℓ are as in our definition of bicyclic graphs, see Figure 12). Next we show that when $1 \leqslant j \leqslant s+t-\ell-2$, $\sum_{\substack{\lambda\vdash n\\ \ell(\lambda)=j}} c_\lambda \neq 0$. We claim that

$$\sum_{i=1}^{\ell}\binom{s+t-2(\ell+1)+i}{j}+\binom{s+t-\ell-2}{j+1}-\binom{s-1}{j+1}-\binom{t-1}{j+1} \tag{5}$$

is greater than zero if $1 \leqslant j \leqslant s+t-\ell-2$. Notice that it is greater than zero whenever $j \geqslant t-1$ since we assume that $s \leqslant t$, and for $j \geqslant t-1$ the last two subtracted binomials are zero. It remains to prove that expression (5) is positive for $1 \leqslant j \leqslant t-2$. To show this, we use the following identity that is a direct consequence of Pascal's relation for any two positive integers $a > b$ we have

$$\binom{a}{j+1}-\binom{b}{j+1}=\sum_{i=b}^{a-1}\binom{i}{j}. \tag{6}$$

We also use the fact that $\binom{a}{i} > \binom{b}{i}$ whenever $a > b \geqslant 0$ and $1 \leqslant i \leqslant a$. By Equation (6) with $a = s+t-\ell-2$ and $b = t-1$, we get

$$\binom{s+t-\ell-2}{j+1}-\binom{t-1}{j+1}=\sum_{i=t-1}^{s+t-\ell-3}\binom{i}{j} \geqslant 0$$

Notice that this means that since $\binom{s-1}{j+1} = 0$ for $j \geqslant s-1$, we actually only need to show that expression (5) is positive for $1 \leqslant j \leqslant s-2$. We use Eq. (6) with $a = s-1$ and $b = j$, to show that

$$\binom{s-1}{j+1} = \sum_{i=j}^{s-2} \binom{i}{j}. \tag{7}$$

Substituting into expression (5) we get

$$\begin{aligned} |\sum_{\substack{\lambda \vdash n \\ \ell(\lambda)=j}} c_\lambda| &= \sum_{i=1}^{\ell} \binom{s+t-2(\ell+1)+i}{j} + \sum_{i=t-1}^{s+t-\ell-3} \binom{i}{j} - \sum_{i=j}^{s-2} \binom{i}{j} \\ &= \sum_{i=1}^{\ell} \binom{s+t-2(\ell+1)+i}{j} + \left(\sum_{i=t-1}^{s+t-\ell-3} \binom{i}{j} - \sum_{i=l}^{s-2} \binom{i}{j} \right) - \sum_{i=j}^{\ell-1} \binom{i}{j}. \end{aligned} \tag{8}$$

Examining the middle expression of (8) and reindexing, we have

$$\sum_{i=t-1}^{s+t-\ell-3} \binom{i}{j} - \sum_{i=\ell}^{s-2} \binom{i}{j} = \sum_{i=0}^{s-\ell-2} \left[\binom{t-1+i}{j} - \binom{\ell+i}{j} \right] > 0\,, \tag{9}$$

as $t-1 = (\ell+m)-1 > \ell$ and $s-\ell-2 = (\ell+k)-\ell+2 \geqslant 0$.

This implies that it suffices to show

$$\sum_{i=1}^{\ell} \binom{s+t-2(\ell+1)+i}{j} - \sum_{i=j}^{\ell-1} \binom{i}{j} \geqslant 0. \tag{10}$$

If $\ell \leqslant j$ then the subtracted summation of inequality (10) is empty and we are done. Otherwise, if $\ell > j$, we can reindex both summations and get

$$\begin{aligned} \sum_{i=1}^{\ell} \binom{s+t-\ell-1-i}{j} - \sum_{i=1}^{\ell-j} \binom{\ell-i}{j} &\geqslant \sum_{i=1}^{\ell} \binom{t-1-i}{j} - \sum_{i=1}^{\ell-j} \binom{\ell-i}{j} \\ &= \sum_{i=\ell-j+1}^{\ell} \binom{t-1-i}{j} + \left(\sum_{i=1}^{\ell-j} \left[\binom{t-1-i}{j} - \binom{\ell-i}{j} \right] \right) \geqslant 0, \end{aligned}$$

where the first inequality follows from $\ell \leqslant s-1$, and last inequality follows from $\ell \leqslant t-1$. So, the summation in parenthesis is greater or equal to zero. By inequality (9) this finishes the proof. □

Recall that the *girth* of a graph G is defined as the length of the smallest cycle. To show that we can recover the cycle sizes of bicyclic graphs, we will make use of the following result.

Proposition 59 ([11, Prop. 3])**.** *The girth g of a graph G is recoverable from $\mathbf{X}_G$.*

Notice that in type-II, the formulas of Propositions 55(2) and 58 are independent of the choice of s and t. On the other hand, we can always assume that $\ell \leqslant k \leqslant m$ and that $s = \ell + k$ is the girth of the graph so that $t = \ell + m$ (see Figure 12). That is, the path common to the two cycles has length $\ell \leqslant s/2 < s$. We will make this assumption to facilitate the proof of our main result for this section.

Theorem 60. *For any bicyclic graph G with cycle sizes $s \leqslant t$ meeting along $\ell < s/2$ edges, we can recover s and t from $\mathbf{X}_G$. Furthermore, if $\ell \geqslant 2$ we can detect that G is of type-II and recover ℓ.*

Proof. First note that the girth of a bicyclic graph is equal to the size of the smaller cycle, $s = \ell + k$, so this is immediately recovered from Theorem 59. Next we show how we can distinguish between the cases $\ell \geqslant 2$ and $\ell \leqslant 1$. Let

$$j^* := \min\{j \in [n] \mid \sum_{\substack{\lambda \vdash n \\ \ell(\lambda)=j}} c_\lambda = 0\},$$

and

$$\hat{t} := \frac{c_{(n)}}{s-1} + 1.$$

Observe that if G is of type-I, i.e., $\ell = 0$, or is of type-II with $\ell = 1$, then $t = \hat{t}$, while if G is of type-II with $\ell \geqslant 2$ then the size $t = \ell + m$ of the other cycle satisfies

$$t = \frac{c_{(n)} + 2\binom{\ell}{2}}{s-1} + 1 = \hat{t} + \frac{\ell(\ell-1)}{s-1} > \hat{t}.$$

Observe that if $\hat{t}$ is not integral, then we may immediately conclude G is of type-II and that $\ell \geqslant 2$. Otherwise, it must be the case that $s-1 \mid \ell(\ell-1)$, and we proceed as follows.

By Theorem 57 and Theorem 58, we have that if G is type-I or $\ell = 1$ then $j^* = s + \hat{t} - 2$ while if G is of type-II then $j^* = s + t - \ell - 1$. We claim that $\ell \geqslant 2$ exactly when $s + \hat{t} - 2 \neq j^*$. This will follow from the inequality

$$\hat{t} - 1 \geqslant t - \ell$$

which we claim is strict for $\ell \geqslant 2$. Observe that if $\ell \geqslant 2$ we have

$$\begin{aligned} t - \ell &= \frac{c_{(n)} + \ell(\ell-1)}{s-1} + 1 - \ell \\ &= \hat{t} - 1 + \left(1 - \frac{\ell(s-\ell)}{s-1}\right). \end{aligned}$$

The claim now follows from the basic inequality $\ell(s-\ell) > s-1$ for $1 < \ell < s-1$, with equality at the endpoints. Note that with our assumptions that $\ell \leqslant k$ and $s = \ell + k$ (see the paragraph preceding the theorem statement), ℓ is always at most $s/2$ and when $\ell \geqslant 2$, we have $s > 3$.

Having shown that we can distinguish the case $\ell \leqslant 1$ (including type-I where $\ell = 0$) from type-II with $\ell \geqslant 2$, we can conclude that if G is type-I or $\ell = 1$, we have $t = \hat{t} = j^* + 2 - s$. On the other hand, if G is type-II we can obtain ℓ by identifying the equations $t = \frac{c_{(n)} + 2\binom{\ell}{2}}{s-1} + 1 = j^* - s + \ell + 1$ and solving for ℓ using the quadratic equation. This produces solutions

$$\ell = \frac{1}{2}(s \pm \sqrt{s^2 - 4((s-j^*)(s-1) + c_{(n)})}).$$

Then substituting for j^* and $c_{(n)}$, we see the discriminant gives

$$s^2 - 4((s-1)(\ell+1-t) + (s-1)(t-1) - \ell(\ell-1)) = s^2 - 4((s-1)(\ell) - \ell(\ell-1))$$

$$
\begin{aligned}
&= s^2 - 4\ell(s-\ell) \\
&= (s-2\ell)^2.
\end{aligned}
$$

This shows that the solutions to the quadratic are exactly the natural number values ℓ and $k = s - \ell$, as desired. The smaller of these solutions in our convention is ℓ, which can then be used with j^* to recover $t = j^* - s + \ell + 1$. $\square$

We remark that it is impossible to distinguish between type-I bicyclic graphs and the $\ell = 1$ case of type-II bicyclic graphs. Indeed, the first example provided by Stanley of non-isomorphic graphs with the same chromatic symmetric function exhibits this difficulty [14] (see Figure 14).

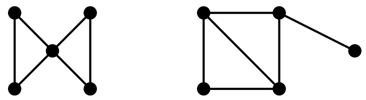

Figure 14: Two bicyclic graphs with identical chromatic symmetric functions

Example 61. Consider the two graphs in Figure 14. They both have chromatic symmetric function:

$$\mathbf{X}_G = 4\mathfrak{st}_{(5)} - 8\mathfrak{st}_{(4,1)} + 4\mathfrak{st}_{(3,2)} + 4\mathfrak{st}_{(3,1,1)} - 3\mathfrak{st}_{(2,2,1)}.$$

Now suppose we only know that $\mathbf{X}_G$ comes from a certain bicyclic graph G with cycle lengths s, t with $s \leqslant t$. We can recover the number of vertices $n = 5$ from the degree. Next we use [11, Prop. 3] to find $s = 3$. Then we may follow the instructions from the proof of Theorem 60:

1. $j^* := \min\{j \in [n] \mid \sum_{\substack{\lambda \vdash n \\ \ell(\lambda)=j}} c_\lambda = 0\} = 4\,,$
2. $\hat{t} := \frac{c_{(n)}}{s-1} + 1 = \frac{4}{3-1} + 1 = 3\,,$
3. $s + \hat{t} - 2 = 3 + 3 - 2 = 4 = j^*\,.$

Therefore we must have $\ell < 2$ and $t = \hat{t} = 3$.

On the other hand, we have the following corollary.

Corollary 62. *No bicyclic graph of type-I or of type-II with* $\min\{k, \ell, m\} = 1$ *can have the same CSF as a type-II bicyclic graph with* $\min\{k, \ell, m\} \geqslant 2$ *where* k, ℓ, m *are as in our definition of bicyclic graphs.*

8 Examples and data

It is known that a graph is bipartite if and only if the graph contains no cycles of odd length. This implies that every tree is a bipartite graph. In [12] the authors construct an infinite number of unicyclic graphs with a 3-cycle with the same CSF. A natural question to then ask is whether the bipartite property of trees plays a role in the tree isomorphism problem. That is, one might ask: If G_1 and G_2 are non-isomorphic bipartite graphs, is

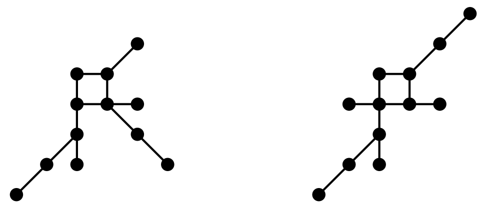

Figure 15: Two bipartite graphs with the same CSF on 12 vertices.

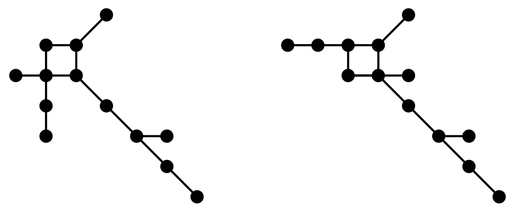

Figure 16: Two graphs with the same CSF on 13 vertices.

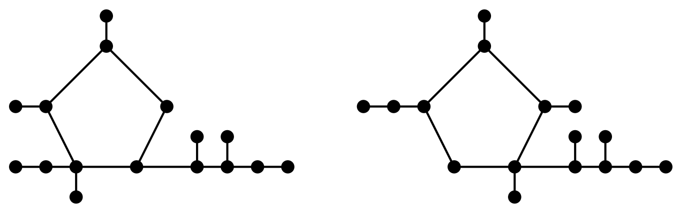

Figure 17: Two graphs with the same CSF on 16 vertices.

it always true that $\mathbf{X}_{G_1} \neq \mathbf{X}_{G_2}$? We exhibit in Figures 15 and 16 examples showing that this is not the case. We note that these are the only pairs of non-isomorphic connected 4-unicyclic graphs up to $n = 17$ vertices.

There is a large family of pairs of non-isomorphic connected unicyclic graphs with a 3-cycle and an even number of vertices that have identical CSFs but distinct numbers of non-trivial rooted trees, internal edges, and degree sequences. On the other hand, the following conjecture has been verified for $n \leqslant 16$ vertices.

Conjecture 63. Let G_1 and G_2 be connected unicyclic graphs, either with a 3-cycle and an odd number of vertices or a c-cycle for $c \geqslant 4$. Then if $\mathbf{X}_{G_1} = \mathbf{X}_{G_2}$ then G_1 and G_2 have the same degree sequence, number of non-trivial rooted trees, and number of internal edges.

In Figures 15 to 17, each pair of graphs shares the same degree sequence, number of non-trivial rooted trees, and number of internal edges. By Theorem 12 two unicyclic graphs with the same CSF must have the same cycle size c. However, it is not known whether there exist two unicyclic graphs with different degree sequences that have the same CSF with $c > 3$. We have verified the following using SageMath [15].

1. CSF distinguishes 3-unicyclic graphs on n vertices if $n = 3, 4, 5, 7, 9, 11, 13, 15$.

2. There are pairs of 3-unicyclic graphs with the same CSF when the number of vertices is $6, 8, 10, 12, 14, 16$.

3. CSF distinguishes 4-unicyclic graphs on n vertices if $n = 14, 15, 16, 17$.

4. CSF distinguishes 5-unicyclic graphs on n vertices if $n \leqslant 15$.
5. CSF distinguishes c-unicyclic graphs on n vertices if $n \leqslant 16$ and $6 \leqslant c \leqslant 16$.

References

[1] José Aliste-Prieto, Anna de Mier, Rosa Orellana, and José Zamora. Marked graphs and the chromatic symmetric function. *SIAM J. Discrete Math.*, 37(3):1881–1919, 2023.

[2] José Aliste-Prieto and José Zamora. Proper caterpillars are distinguished by their chromatic symmetric function. *Discrete Math.*, 315-316:158–164, 2014.

[3] Soojin Cho and Stephanie van Willigenburg. Chromatic classical symmetric functions. *J. Comb.*, 9(2):401–409, 2018.

[4] Samantha Dahlberg and Stephanie van Willigenburg. Lollipop and lariat symmetric functions. *SIAM J. Discrete Math.*, 32(2):1029–1039, 2018.

[5] Arunkumar Ganesan, Narayanan Narayanan, B. V. Raghavendra Rao, and Sagar S. Sawant. Proper q-caterpillars are distinguished by their chromatic symmetric functions. *Discrete Math.*, 347(11):8, 2024.

[6] Michael Gonzalez, Rosa Orellana, and Mario Tomba. The chromatic symmetric function in the star-basis. *Discrete Math.*, 349(2):28, 2026. Id/No 114691.

[7] Michael Gonzalez, Rosa Orellana, and Mario Tomba. On the reconstruction of trees from their chromatic symmetric functions. *Sém. Lothar. Combin.*, 93B (2025), Art. 78, 12 pp.

[8] Jake Huryn and Sergei Chmutov. A few more trees the chromatic symmetric function can distinguish. *Involve*, 13(1):109–116, 2020.

[9] S. Heil and Caleb Ji. On an algorithm for comparing the chromatic symmetric functions of trees. *Australas. J. Comb.*, 75:210–222, 2019.

[10] Martin Loebl and Jean-Sébastien Sereni. Isomorphism of weighted trees and Stanley's isomorphism conjecture for caterpillars. *Ann. Inst. Henri Poincaré D, Comb. Phys. Interact.*, 6(3):357–384, 2019.

[11] Jeremy L. Martin, Matthew Morin, and Jennifer D. Wagner. On distinguishing trees by their chromatic symmetric functions. *J. Comb. Theory, Ser. A*, 115(2):237–253, 2008.

[12] Rosa Orellana and Geoffrey Scott. Graphs with equal chromatic symmetric functions. *Discrete Math.*, 320:1–14, 2014.

[13] Rosa Orellana and Foster Tom. *Linear relations on star coefficients of the chromatic symmetric function.* arXiv:2601.13390, 2026

[14] Richard P. Stanley. A symmetric function generalization of the chromatic polynomial of a graph. *Adv. Math.*, 111(1):166–194, 1995.

[15] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 10.5)*, 2025.

[16] Douglas B. West. *Introduction to graph theory.* Upper Saddle River, NJ: Prentice Hall, 1996.

[17] David G. L. Wang and Monica M. Y. Wang. The e-positivity of two classes of cycle-chord graphs. *J. Algebr. Comb.*, 57(2):495–514, 2023.

[18] Yuzhenni Wang, Xingxing Yu, and Xiao-Dong Zhang. A class of trees determined by their chromatic symmetric functions. *Discrete Math.*, 347(9):11, 2024. Id/No 114096.