Proper q-caterpillars are distinguished by their Chromatic Symmetric Functions

G. Arunkumar Narayanan Narayanan Raghavendra Rao B. V.
Sagar S. Sawant

Abstract

Stanley's Tree Isomorphism Conjecture posits that the chromatic symmetric function can distinguish non-isomorphic trees. This conjecture is already established for caterpillars and other subclasses of trees. We prove the conjecture's validity for a new class of trees that generalize proper caterpillars, thus confirming the conjecture for a broader class of trees.

Mathematics Subject Classification(2010): 05C15, 05C25, 05C31, 05C60.

Keywords: chromatic symmetric function, U-polynomial, integer compositions, caterpillars.

1 Introduction

The chromatic symmetric function, introduced by R. Stanley [13], generalizes the chromatic polynomial of a graph to a symmetric function and is widely studied as it determines generating functions of various graph statistics. This raises the question: Does the chromatic symmetric function distinguish all graphs up to isomorphism? Unfortunately, the answer is negative. Stanley presented two non-isomorphic graphs (both containing cycles) that share the same chromatic symmetric function. However, the question remains open for trees and is conjectured to be true, famously known as Stanley's Tree Isomorphism Conjecture. Substantial progress has been made in confirming the conjecture for various subclasses of trees. Martin et al. [10] proved its validity for specific classes of caterpillars, spiders, and certain unicyclic graphs. Additionally, Aliste-Prieto and Zamora [2] showed that the conjecture holds for proper caterpillars, while Loebl and Sereni [9] extended this result to all caterpillars. To explore additional examples of graphs that can be distinguished based on their chromatic symmetric function, refer to [1, 6, 7, 8, 14, 15].

Let $\mathbb{Q}$ and $\mathbb{P}$ denote the rational numbers and positive integers, respectively. In this paper, we consider the following generalization of proper caterpillars.

Definition 1 (proper q-caterpillars). *Let $q \geq 1$ be fixed. A proper q-caterpillar T is constructed as follows: We begin with a path $S = \langle v_1, \ldots, v_\ell \rangle$ (with endpoints v_1 and v_ℓ) called the spine, with $\ell > 0$. For every $1 \leq i \leq \ell$, we glue (endpoint of the path identified with a vertex on the spine) p_i additional paths of length exactly q to the vertices v_i, respectively, where $p_i \in \mathbb{P}$.*

In this context, proper 1-caterpillars have been distinguished by their chromatic symmetric functions up to isomorphism [2]. We show that for all $q \geq 1$, the chromatic symmetric function of a tree determines whether the tree is a proper q-caterpillar or not. (See Proposition 2.1.) Further, we prove that for all $q \geq 2$, proper q-caterpillars are distinguished by their chromatic symmetric functions.

Theorem 1.1. *For $q \geq 2$, the chromatic symmetric function distinguishes isomorphism classes of proper q-caterpillars.*

The proof uses ideas involved in [2], that is, associating proper q-caterpillars with the integer compositions, and the interrelations of the chromatic symmetric function, U-polynomial and $\mathcal{L}$-polynomial. The U-polynomial, introduced by Noble and Welsh [11], is a Tutte-Grothendieck invariant equivalent to the chromatic symmetric function when restricted to trees; that is, one can be obtained from the other by certain change of variables. Consequently, Stanley's Tree Isomorphism Conjecture is equivalent to distinguishing trees by their U-polynomial.

We obtain a characterization of proper q-caterpillars that only involves invariants determined by the chromatic symmetric function in Proposition 2.1. Lemma 2.5 relates the U-polynomial of the proper q-caterpillars with the $\mathcal{L}$-polynomial of the corresponding integer composition. Further, we exhibit the irreducible factorization of the integer composition corresponding to a proper q-caterpillar. (See Lemma 2.6.) Finally, we combine these ideas to prove Theorem 1.1. Note that for $q \geq 2$, every integer composition $(p_1, p_2, \ldots, p_\ell)$ with each component being positive corresponds to a unique proper q-caterpillar with p_i many paths of length q incident on the vertex v_i of the spine $\langle v_1, v_2, \ldots, v_\ell \rangle$. Therefore, Theorem 1.1 asserts that for each such integer composition, there are infinitely many trees (one for each $q \geq 2$) that can be distinguished by their chromatic symmetric function, thereby attaining a significant improvement in the pool of trees that are known to satisfy Stanley's Tree Isomorphism Conjecture.

The rest of the paper is organized as follows. In Section 1.1, we include fundamental graph notions including the chromatic symmetric function and U-polynomial. In Section 2.1, we present the proof of Proposition 2.1. The factorization of integer compositions is discussed in Section 2.2. We prove Lemma 2.5 in Section 2.3. This, combined with the factorization in Section 2.2, leads to the proof of Theorem 1.1.

1.1 Preliminaries

Let $G = (V, E)$ be a simple graph with the set of vertices V and the set of edges E. A $\mathbb{P}$-*coloring* of a graph G is a function $f \colon V \to \mathbb{P}$, and such a coloring is said to be *proper* if, for every edge $uv \in E$, the colors $f(u)$ and $f(v)$ are distinct. The *content of a coloring* f is the $\mathbb{P}$-tuple $(|f^{-1}(1)|, |f^{-1}(2)|, |f^{-1}(3)|, \ldots)$, denoted by $c(f)$, that encodes the cardinality of the color classes of f. Throughout this paper, whenever we consider a $\mathbb{P}$-tuple, we assume that all but finitely many of its components are zero. In what follows, we adopt the graph notions and terminology in accordance with [4].

Let $\mathbf{x} = (x_1, x_2, \ldots)$ be a collection of commutative indeterminates. For a $\mathbb{P}$-tuple α, let $\mathbf{x}^\alpha$ be the monomial having the i^{th} component of α as the exponent of x_i. A *symmetric function* is a formal power series in indeterminates $\mathbf{x}$ that is invariant under all permutations of $\mathbf{x}$. Let $\mathrm{Sym}_R(\mathbf{x})$ denote the collection of symmetric functions with coefficients over ring R. We refer the reader to [12] for detailed exposition to the theory of symmetric functions. We now define the chromatic symmetric function introduced by Stanley [13].

The *chromatic symmetric function* of a graph $G = (V, E)$ is defined as

$$\mathbf{X}_G := \sum_{\substack{f:V\to\mathbb{P} \\ proper}} \mathbf{x}^{c(f)}. \tag{1.1}$$

The above function is indeed symmetric in $\mathbf{x}$ since a permutation of the colors does not affect the properness of colorings. Moreover, the chromatic symmetric function X_G is homogeneous in $\mathbf{x}$ with degree $|V|$.

The coefficients arising in the expansion of the chromatic symmetric function in various bases of the $\mathrm{Sym}_{\mathbb{Q}}(\mathbf{x})$ encodes numerous combinatorics of the graph. We are particularly interested in the expansion with respect to the power sum symmetric function basis.

For $k \in \mathbb{P}$, the *power sum symmetric function of degree* k is defined as

$$p_k(\mathbf{x}) = \sum_{i\in\mathbb{P}} x_i^k.$$

An integer *partition* of a positive integer n is a weakly decreasing sequence of positive integers whose sum is n. The notation $\lambda \vdash n$ indicates that λ is an integer partition of n. For any partition $\lambda = \lambda_1\,\lambda_2\,\cdots\,\lambda_k \vdash n$, the power sum symmetric function

$$p_\lambda(\mathbf{x}) := \prod_{i=1}^{k} p_{\lambda_i}(\mathbf{x}),$$

and the collection $\{p_\lambda\}_{n,\lambda\vdash n}$ forms a $\mathbb{Q}$-basis of $\mathrm{Sym}_{\mathbb{Q}}(\mathbf{x})$.

Given a graph $G = (V, E)$ and a subset $F \subseteq E$, let $G[F]$ denote the subgraph of G with vertex set V and edge set F. We call the subgraph $G[F] = (V, F)$ as the *spanning subgraph* of G. Let $\lambda[F]$ be the partition of $|V|$ formed by the orders of the connected components of the spanning subgraph $G[F]$.

Theorem 1.2 ([13, Theorem 2.5])**.** *For a graph G, the expansion of the chromatic symmetric function in the power sum symmetric function basis is*

$$\mathbf{X}_G = \sum_{F\subseteq E} (-1)^{|F|} p_{\lambda[F]}(\mathbf{x}). \tag{1.2}$$

The U-polynomial defined by Noble and Welsh [11] establishes a strong connection with the chromatic symmetric function due to the expansion (1.2). For a partition $\lambda = \lambda_1\,\lambda_2\,\cdots\,\lambda_k \vdash n$, let $\mathbf{x}_\lambda$ denote the monomial $x_{\lambda_1}x_{\lambda_2}\cdots x_{\lambda_k}$,

Definition 2. *Given a graph $G = (V, E)$, the U-polynomial of the graph is defined as*

$$U_G(\mathbf{x}; y) = \sum_{F\subseteq E} \mathbf{x}_{\lambda[F]}(y-1)^{|F|-|V|+\mathcal{K}(F)},$$

where $\kappa(F)$ is the number of connected components in the spanning subgraph $G[F]$, or equivalently the length of partition $\lambda[F]$.

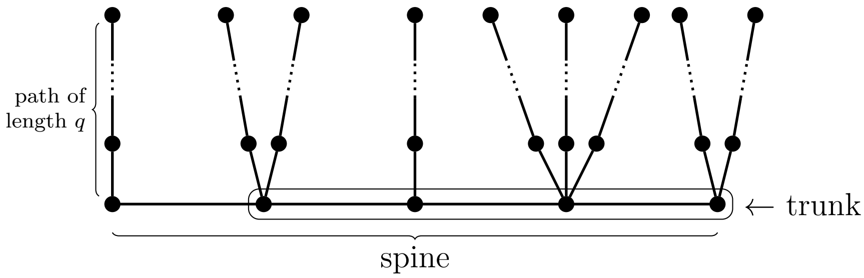


Figure 1: Example of a proper q-caterpillar with spine of order 5, trunk of order 4 and the multiset of twigs is $\{\underbrace{q, q, \ldots, q}_{8 \text{ times}}, q+1\}$.

Any spanning subgraph $T[F]$ of a tree $T = (V, E)$ must have $|V| - |F|$ connected components, leading to $|F| - |V| + \kappa(F) = 0$ for all $F \subseteq E$. Therefore, for any tree T, we have

$$(-1)^{|V|} U_T\left(-p_1(\mathbf{x}), -p_2(\mathbf{x}), -p_3(\mathbf{x}), \ldots; y\right) = (-1)^{|V|} \sum_{F \subseteq E} (-1)^{\mathcal{K}(F)} p_{\lambda[F]}(\mathbf{x}) = \mathbf{X}_T\,. \quad (1.3)$$

This implies that the two graph invariants are equivalent when restricted to trees, that is, two trees have the same chromatic symmetric function if and only if they have the same U-polynomial.

2 Proper q-Caterpillars

We prove Theorem 1.1 in this section. We begin by characterizing proper q-caterpillars in terms of certain statistics that can be identified through the chromatic symmetric function. This allows us to distinguish proper q-caterpillars from other types of trees.

2.1 Characterization of proper q-caterpillars

Given a tree $T = (V, E)$, the *trunk* T° of T is the smallest subtree containing all vertices of degree at least three. For each pendant vertex u of T, there exists a unique path starting at u and ending at some vertex in the trunk such that all internal vertices of the path have degree two. Each such path is called a *twig*, and let $\textsc{Twig}(T)$ be the multiset representing the lengths of twigs in T. Evidently, every tree containing a vertex of degree at least three can be decomposed into the trunk T° and some twigs. Crew proved that the order of T°, and the multiset $\textsc{Twig}(T)$ can be determined by the chromatic symmetric function [6].

It is clear that a path T is a proper q-caterpillar if and only if its order is either $q+1, 2q+1$ or $2q+2$. For proper q-caterpillars that are not paths, we characterize the proper q-caterpillars using tree-invariants such as, order of the tree, degree sequence, multiset of twigs and the diameter of the tree. The recovery of these invariants from the chromatic symmetric function of trees is detailed in [5, 6, 10].

Proposition 2.1. *Let $q \geq 1$ be fixed and $T = (V, E)$ be a tree that is not a path. Then T is a proper q-caterpillar if and only if it satisfies the following:*

(i) $|T^\circ| = |V| - \delta_1 - \delta_2$ *where* δ_i *is the number of vertices of degree* i *in* T*.*

(ii) $\textsc{Twig}(T)$ *only contains integers* q *and* $q{+}1$*, with* $m_{q+1} \leq 2$ *where* m_{q+1} *is the multiplicity of* $q{+}1$ *in* $\textsc{Twig}(T)$*.*

(iii) $\operatorname{diam}(T) = (|T^\circ| - 1) + 2q + m_{q+1}$.

Proof. ($\Rightarrow$) It is clear that every proper q-caterpillar that is not a path satisfies the above three conditions.
($\Leftarrow$) A tree satisfying $|T^\circ| = 1$ and (ii) is indeed a proper q-caterpillar. Thus we may assume that $|T^\circ| \geq 2$. Note that $\operatorname{diam}(T) \leq 2q + \operatorname{diam}(T^\circ) + m_{q+1}$ along with (iii) implies that $(|T^\circ| - 1) \leq \operatorname{diam}(T^\circ)$, and hence T° is a path, say $\langle w_1, w_2, \ldots, w_k\rangle$ (with endpoints w_1 and w_k). From (i), it follows that T° consists only of vertices of degree at least 3, owing to which every vertex of the trunk must be incident to at least one twig. To prove that T is a proper q-caterpillar, it suffices to prove that twigs of length $q+1$ (if they exist) are incident to the distinct endpoints of the trunk. For $1 \leq i \leq k$, let w_i be incident to n_i many twigs P_i^t ($1 \leq t \leq n_i$), and fix $0 \leq r_i \leq n_i$ where the length of the path P_i^t is equal to q if $r_i < t \leq n_i$ and $q+1$ otherwise. Let u_i^t be the pendant vertex of the twig $P_i^t(1 \leq t \leq n_i)$. In the resulting tree T, we have the following

$$d(u_i^t, u_j^s) = \begin{cases} q + |i-j| + q & \text{if } r_i < t \leq n_i \text{ and } r_j < s \leq n_j, \\ (q+1) + |i-j| + q & \text{if } 1 \leq t \leq r_i \text{ and } r_j < s \leq n_j, \\ q + |i-j| + (q+1) & \text{if } r_i < t \leq n_i \text{ and } 1 \leq s \leq r_j, \\ (q+1) + |i-j| + (q+1) & \text{if } 1 \leq t \leq r_i \text{ and } 1 \leq s \leq r_j. \end{cases}$$

From the above computation, the endpoints of the path in T of length $\operatorname{diam}(T)$ must be w_1^t and w_k^s for some $1 \leq t \leq n_1$ and $1 \leq s \leq n_k$. This together with (ii) and (iii) dictates the position of $q{+}1$-twigs as follows:

$$m_{q+1} = \begin{cases} 0 & \text{if } r_1 = r_k = 0, \\ 1 & \text{if exactly one of } r_1 \text{ or } r_k \text{ is non-zero}, \\ 2 & \text{if both } r_1 \text{ and } r_k \text{ are non-zero}. \end{cases}$$

Therefore, the tree T is a proper q-caterpillar, and this completes the proof. □

Note. The trunk of the proper q-caterpillar may not coincide with the spine (see Figure 1). However, it is always a subpath of the spine.

Before proceeding to the proof of Theorem 1.1, we revisit the factorization of integer compositions introduced in [3]. This factorization is instrumental for determining the isomorphism classes of proper q-caterpillars.

2.2 Monoid of Integer Compositions

Let $n \in \mathbb{P}$ be a positive integer. An *integer composition* α of n, denoted by $\alpha \vDash n$, is an ordered sequence of positive integers $\alpha_1\,\alpha_2\,\cdots\,\alpha_r$ whose sum is n. The integer α_i is called the i^{th} component of α. The length of the integer composition α, denoted by $\ell(\alpha)$, is the number of components r in α. Let $\mathscr{C}$ denote the set of all integer compositions.

For any two compositions $\alpha = \alpha_1\,\alpha_2\,\cdots\,\alpha_r$ and $\beta = \beta_1\,\beta_2\,\cdots\,\beta_s$, their *concatenation* is given by

$$\alpha \cdot \beta := \alpha_1\,\alpha_2\,\cdots\,\alpha_r\,\beta_1\,\beta_2\,\cdots\,\beta_s,$$

whereas the near-concatenation operation is defined as

$$\alpha \odot \beta := \alpha_1\,\alpha_2\,\cdots\,\alpha_{r-1}\,(\alpha_r + \beta_1)\,\beta_2\,\cdots\,\beta_s.$$

Let $\alpha^{\odot q}$ denote the q-fold near-concatenation $\underbrace{\alpha \odot \alpha \odot \cdots \odot \alpha}_{q \text{ times}}$, for any positive integer q. The *composition* of two integer compositions is given by

$$\alpha \circ \beta := \beta^{\odot \alpha_1} \cdot \beta^{\odot \alpha_2} \cdot \cdots \cdot \beta^{\odot \alpha_r}.$$

For example, $2\,1 \circ 2\,3 = (2\,3\, \odot 2\,3) \cdot 2\,3 = 2\,5\,3\,2\,3$.

Proposition 2.2 ([3, Proposition 3.3])**.** *$(\mathscr{C}, \circ)$ is a non-commutative monoid with the integer composition* 1 *as the identity element.*

Let α and β be two integer compositions of n. Then α is said to be a *refinement* of β if α is obtained by subdividing some (or no) parts of β, denoted by $\alpha \preceq \beta$. In this case, we also say that β is a *coarsening* of α. For example, $2\,3\,1\,3\,2 \preceq 2\,4\,5$. Let $(\mathscr{C}, \preceq)$ be the poset with the refinement order. In [3], Billera, Thomas and Willigenburg defined an equivalence relation on $\mathscr{C}$ based on the refinement of the integer compositions. We consider the polynomial interpretation of that equivalence relation called the $\mathcal{L}$-polynomial [2, 3].

The $\mathcal{L}$-polynomial of an integer composition α is defined as

$$\mathcal{L}(\mathbf{x}; \alpha) = \sum_{\beta \succeq \alpha} x_{\beta_1} x_{\beta_2} \ldots x_{\beta_r}.$$

For instance, the $\mathcal{L}$-polynomial of the composition $2\,2\,1\,2$ is $x_1x_2^3 + x_1x_2x_4 + 2x_2^2x_3 + 2x_2x_5 + x_3x_4 + x_7$. Note that the equality of the $\mathcal{L}$-polynomial induces an equivalence relation on integer compositions. Let $[\alpha]_{\mathcal{L}}$ denote the equivalence class of α under this equivalence relation. We recall its description using the unique factorization in $(\mathscr{C}, \circ)$ [3].

A factorization $\alpha = \varepsilon \circ \eta$ is said to be *trivial* if one of the following is satisfied:

a) either ε or η is the identity composition 1,

b) both ε and η are of length 1,

c) both ε and η have all parts equal to 1.

An integer composition is said to be *irreducible* if it admits only trivial factorizations. A factorization $\alpha = \eta_1 \circ \eta_2 \circ \cdots \circ \eta_k$ is said to be an *irreducible factorization* if each integer composition η_i is irreducible and no $\eta_i \circ \eta_{i+1}$ is a trivial factorization.

Theorem 2.3 ([3, Theorem 3.6])**.** *Every integer composition admits a unique irreducible factorization.*

For an integer composition α, let α^* be the integer composition obtained by reversing α, that is, the i^{th} component of α^* is $\alpha_{\ell(\alpha)-i+1}$ for $1 \leq i \leq \ell(\alpha)$.

Theorem 2.4 ([3, Theorem 4.1])**.** *Let $\alpha = \eta_1 \circ \eta_2 \circ \cdots \circ \eta_k$ be the irreducible factorization of α. Then*

$$[\alpha]_{\mathcal{L}} = \{\varepsilon_1 \circ \varepsilon_2 \circ \cdots \circ \varepsilon_k \ \mid \ \varepsilon_i = \eta_i \text{ or } \varepsilon_i = {\eta_i}^*, \text{ for all } i = 1, 2, \dots k\},$$

Example. Consider the integer composition 4 10 4 10 with its irreducible factorization given by $1\ 1 \circ 2\ 5 \circ 2$. Then the equivalence class

$$[4\ 10\ 4\ 10]_{\mathcal{L}} = \{1\ 1 \circ 2\ 5 \circ 2, 1\ 1 \circ 5\ 2 \circ 2\} = \{4\ 10\ 4\ 10, 10\ 4\ 10\ 4\}.$$

2.3 Proof of Theorem 1.1

In Section 2.1, we have established that the chromatic symmetric function of a tree identifies whether the tree is a proper q-caterpillar or not. Our objective now is to demonstrate that the chromatic symmetric function distinguishes non-isomorphic proper q-caterpillars. To accomplish this, we associate every proper q-caterpillar with a unique integer composition such that any two proper q-caterpillars are isomorphic if and only if their corresponding compositions are either the same or reverses of one another. The proof technique is similar to the proof of distinguishing proper 1-caterpillars [2].

Let $q \geq 2$, and T be a proper q-caterpillar. Let $\langle v_1, v_2, \dots v_\ell \rangle$ denote the spine of T. Let p_i represent the number of paths in T of length q, starting from a leaf and ending at v_i. We define a composition $\varphi(T)$ of length ℓ whose i^{th} component is $q \cdot p_i + 1$. Conversely, for any integer composition α with all components greater than one and congruent to 1 modulo q, we construct a proper q-caterpillar $\tau(\alpha)$ as follows: consider a path with $\ell(\alpha)$ vertices, which serves as the spine, and glue $\frac{\alpha_i - 1}{q}$ new paths of length q to the i^{th} vertex of the spine. The mapping φ and τ are inverses of each other. For instance, the proper q-caterpillar in Figure 1 corresponds to the integer composition $q{+}1\ \ 2q{+}1\ \ q{+}1\ \ 3q{+}1\ \ 2q{+}1$.

Remark 1. *Any two proper q-caterpillars S and T are isomorphic if and only if $\varphi(S) = \varphi(T)$ or $\varphi(S) = \varphi(T)^*$.*

The following lemma, which is a generalization of [2, Proposition 2.5], states that the $\mathcal{L}$-polynomial of the compositions associated to proper q-caterpillars can be obtained as an evaluation of the U-polynomial. This, along with (1.3) asserts that the chromatic symmetric function of a proper q-caterpillar T determines the $\mathcal{L}$-polynomial of $\varphi(T)$.

Lemma 2.5. *Let $q \geq 1$. For any proper q-caterpillar $T = (V, E)$ and the composition $\varphi(T)$ associated to T, we have*

$$U_T(\underbrace{0, 0, \dots, 0}_{q \text{ times}}, x_{q+1}, x_{q+2}, \dots) = \mathcal{L}(\varphi(T); \mathbf{x}). \tag{2.4}$$

Proof. The U-polynomial with $x_1 = x_2 = \cdots = x_q = 0$ can be interpreted as the subset-sum over $F \subseteq E$ such that each connected component of the induced subgraph $T[F]$ has order at least $q + 1$. This implies that such an F must contain all non-spine edges (otherwise, the induced subgraph $T[F]$ would contain a connected component of order at most q). Thus every monomial $\mathbf{x}_{\lambda[F]}$ in $U_T(0, 0, \dots, 0, x_{q+1}, x_{q+2}, \dots)$ corresponds to the subset $F' := F \cap S$, where S is the set of spine edges. Any such subset F' of

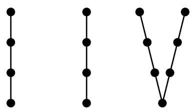
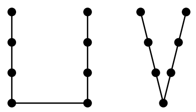
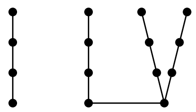
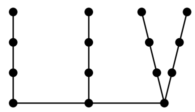
Subgraphs				
Integer Compositions	4 4 7	8 7	4 11	15
Monomials	$x_4^2x_7$	x_7x_8	x_4x_{11}	x_{15}

Table 1: Subgraphs containing all non-spine edges of a proper 3-caterpillar $\tau(4\ \ 4\ \ 7)$, their corresponding compositions, and the monomials contributed by them.

spine-edges determines a unique coarsening $\varphi(T)^{F'}$ of the composition $\varphi(T)$ in the poset $(\mathscr{C}, \preceq)$ obtained as follows: for every maximal path $\langle v_{i_1}, v_{i_2}, \ldots, v_{i_k}\rangle$ induced by the spine vertices in $T[F']$, add the corresponding components $\varphi(T)_{i_1}, \varphi(T)_{i_2}, \ldots, \varphi(T)_{i_k}$ to get a single component of $\varphi(T)^{F'}$. Furthermore, the monomial contributed by the subset F in $U_T(0, 0, \ldots, 0, x_{q+1}, x_{q+2}, \ldots)$ is same as the monomial contributed by the coarsening $\varphi(T)^{F'}$ in $\mathcal{L}(\varphi(T); \mathbf{x})$. (See Table 1 for an example.) Therefore, we have

$$U_T(0, 0, \ldots, 0, x_{q+1}, x_{q+2}, \ldots) = \sum_{\substack{F \subseteq E \\ F \text{ contains all} \\ \text{non-spine edges}}} \mathbf{x}_{\lambda[F]} = \mathcal{L}(\varphi(T); \mathbf{x}).$$

□

The following lemma helps in determining the irreducible factorization of the integer compositions associated with the proper q-caterpillars.

Lemma 2.6. *Let $q \geq 2$ and h be positive integers such that q does not divide h. Let γ be an integer composition in which each component is h modulo q, and the greatest common divisor (gcd) of all components is 1. Then either γ is irreducible, or its irreducible factorization is $\gamma = (1^m) \circ \omega$, where (1^m) denotes the integer composition of length m with all components equal to 1, for some $m \geq 1$.*

Proof. We may assume that γ is not irreducible. We prove using induction on length of γ. Let $\gamma = \zeta \circ \eta$ be a non-trivial factorization of γ. We claim that each component of ζ must be equal to 1. Assume to the contrary that ζ contains at least one component greater than 1, and let i be the smallest index with the i^{th} component $\zeta_i > 1$. The gcd of all components of γ being 1 implies that the length of η must be at least 2. Since $\gamma_1 = \eta_1$ and $\gamma_{\ell(\gamma)} = \eta_{\ell(\eta)}$, both η_1 and $\eta_{\ell(\eta)}$ are congruent to h modulo q. For $k = \ell(\eta) \cdot i$, consider the k^{th} component of $\gamma = \zeta \circ \eta$. By the given hypotheses, we get γ_k to be congruent to h modulo q, but the factorization implies

$$\varepsilon_k = (\zeta \circ \eta)_k = \eta_1 + \eta_{\ell(\eta)} \equiv 2h \pmod{q}.$$

This is not possible because h is non-zero modulo q. Therefore ζ must have all the components equal to 1, that is, $\gamma = (1^r) \circ \eta$ for some $r \geq 2$. Note that η satisfies the given hypothesis and its length $\ell(\eta) < \ell(\gamma)$. Using induction, either η is irreducible or its irreducible factorization is $(1^s) \circ \omega$, and consequently, the irreducible factorization of γ is $(1^r) \circ \eta$ or $(1^{rs}) \circ \omega$, respectively. Thus γ admits the required irreducible factorization. □

Using Lemma 2.6, we can conclude that the proper q-caterpillars are distinguished by the chromatic symmetric functions up to isomorphism.

Proof of Theorem 1.1. Let $q \geq 2$. Let S and T be two proper q-caterpillars with the same chromatic symmetric function. Lemma 2.5 implies that the $\mathcal{L}$-polynomial of $\varphi(S)$ and $\varphi(T)$ are equal as well, that is, $\varphi(S) \in [\varphi(T)]_{\mathcal{L}}$. Note that it suffices to prove the equivalence class $[\varphi(T)]_{\mathcal{L}} = \{\varphi(T), \varphi(T)^*\}$, as it would imply $\varphi(S) = \varphi(T)$ or $\varphi(S) = \varphi(T)^*$. This, along with Remark 1 would imply that S is isomorphic to T. If the gcd of all components of $\varphi(T)$ is 1, then by Lemma 2.6 either $\varphi(T)$ is irreducible or its irreducible factorization is $(1^r) \circ \omega$. On the other hand, if the gcd of all components is d which is greater than 1, then factorize $\varphi(T) = \varepsilon \circ d$. Note that the gcd of all components of ε is 1, and each component is congruent to h modulo q, where h is the least positive integer satisfying $d{\cdot}h \equiv 1 \pmod{q}$. By Lemma 2.6, either ε is irreducible or its irreducible factorization must be $(1^r) \circ \omega$ for some $r \geq 2$. This implies that the irreducible factorization of $\varphi(T)$ is $\varepsilon \circ d$ or $(1^r) \circ \omega \circ d$. In either case, the irreducible factorization of $\varphi(T)$ contains at most one non-palindrome composition. This, along with Theorem 2.4 concludes that $[\varphi(T)]_{\mathcal{L}} = \{\varphi(T), \varphi(T)^*\}$. This completes the proof. □

3 Concluding remarks and future directions

While Stanley's Tree Isomorphism Conjecture remains open, our result demonstrates that the ideas presented in [2] can be extended to a more general class of trees that resemble proper caterpillars.

Along this line, we call a tree $T(V, E)$ as a *generalized caterpillar* if the trunk of the tree forms a path. Further, a *generalized proper caterpillar* is a generalized caterpillar in which every vertex of the trunk has degree at least 3. Equivalently, a tree is a generalized proper caterpillar if and only if it satisfies $|T^\circ| = |V| - \delta_1 - \delta_2$, where T° represents the trunk of the tree and δ_i denotes the number of vertices of degree i in T. We believe that further generalizations of Lemma 2.5 might hold for generalized proper caterpillars. In particular, we propose the following question:

Question 3.1. *Do the U-polynomials of generalized proper caterpillars relate to the $\mathcal{L}$-polynomials of the associated integer compositions?*

For instance, consider a tree T obtained from a proper q-caterpillar S by gluing an additional twig of length $q+1$ at i^{th} vertex of the spine. Let $\varphi'(T)$ be the integer composition obtained from $\varphi(S)$ by replacing the i^{th} component $\varphi(S)_i$ with $\varphi(S)_i + (q+1)$. Then, it can be seen that

$$U_T(\underbrace{0, 0 \ldots, 0}_{q}, x_{q+1}, \ldots) = \mathcal{L}(\varphi'(T); \mathbf{x}) + x_{q+1}\mathcal{L}(\varphi(S); \mathbf{x}).$$

Observe that the U-polynomial of T is expressed as the sum of the $\mathcal{L}$-polynomials of $\varphi(S)$ and $\varphi'(T)$. It turns out that in such cases, Theorem 2.4 cannot be applied directly. Nevertheless, we do believe that distinguishing such trees by U-polynomial is feasible.

Acknowledgements

We thank José Aliste-Prieto for valuable comments and suggestions that helped improve the presentation significantly. We also thank Anna de Mier for providing valuable feedback. We extend our gratitude to the anonymous referees for their comments that greatly enhanced the quality of the manuscript.

The first author was supported by SERB SRG GRANT- SRG/2022/001281 and Indian Institute of Technology, Madras NFIG GRANT - RF/22-23/0985/MA/NFIG/009003. The fourth author thanks the Indian Institute of Technology, Madras, India, for the financial support under the HTRA fellowship.

References

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

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

[3] Louis J. Billera, Hugh Thomas, and Stephanie van Willigenburg. Decomposable compositions, symmetric quasisymmetric functions and equality of ribbon schur functions. *Advances in Mathematics*, 204(1):204–240, 2006.

[4] A. Bondy and U.S.R. Murty. *Graph Theory*. Graduate Texts in Mathematics. Springer London, 2011.

[5] Sharad Chaudhary and Gary Gordon. Tutte polynomials for trees. *Journal of Graph Theory*, 15(3):317–331, 1991.

[6] Logan Crew. A note on distinguishing trees with the chromatic symmetric function. *Discrete Mathematics*, 345(2):112682, 2022.

[7] Angèle M. Foley, Joshua Kazdan, Larissa Kröll, Sofía Martínez Alberga, Oleksii Melnyk, and Alexander Tenenbaum. Spiders and their kin: an investigation of Stanley's chromatic symmetric function for spiders and related graphs. *Graphs Combin.*, 37(1):87–110, 2021.

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

[9] Martin Loebl and Jean-Sébastien Sereni. Isomorphism of Weighted Trees and Stanley's Isomorphism Conjecture for Caterpillars. *Annales de l'Institut Henri Poincaré (D) Combinatorics, Physics and their Interactions*, 2018.

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

[11] Steven D. Noble and Dominic J. A. Welsh. A weighted graph polynomial from chromatic invariants of knots. *Annales de l'Institut Fourier*, 49(3):1057–1087, 1999.

[12] Richard P. Stanley and Sergey Fomin. *Enumerative Combinatorics.* Cambridge University Press, jan 1999.

[13] R.P. Stanley. A symmetric function generalization of the chromatic polynomial of a graph. *Advances in Mathematics*, 111(1):166–194, 1995.

[14] Shuhei Tsujie. The chromatic symmetric functions of trivially perfect graphs and cographs. *Graphs Combin.*, 34(5):1037–1048, 2018.

[15] Wang, Yuzhenni, Xingxing Yub, and Xiao-Dong Zhang. A class of trees determined by their chromatic symmetric functions. arXiv:2308.03980, 2023.