An Exact Enumeration of Distance-Hereditary Graphs

Cédric Chauve* Éric Fusy† Jérémie Lumbroso‡

Abstract

Distance-hereditary graphs form an important class of graphs, from the theoretical point of view, due to the fact that they are the totally decomposable graphs for the split-decomposition. The previous best enumerative result for these graphs is from Nakano *et al.* (J. Comp. Sci. Tech., 2007), who have proven that the number of distance-hereditary graphs on n vertices is bounded by $2^{\lceil 3.59n \rceil}$.

In this paper, using classical tools of enumerative combinatorics, we improve on this result by providing an *exact* enumeration and *full asymptotic* of distance-hereditary graphs, which allows to show that the number of distance-hereditary graphs on n vertices is tightly bounded by $(7.24975\ldots)^n$—opening the perspective such graphs could be encoded on $3n$ bits. We also provide the exact enumeration and full asymptoticss of an important subclass, the 3-leaf power graphs.

Our work illustrates the power of revisiting graph decomposition results through the framework of analytic combinatorics.

Introduction

The decomposition of graphs into tree-structures is a fundamental paradigm in graph theory, with algorithmic and theoretical applications [5]. In the present work, we are interested in the *split-decomposition*, introduced by Cunningham and Edmonds [9, 10] and recently revisited by Gioan *et al.* [20, 21, 7]. For the classical modular and split-decomposition, the *decomposition tree* of a graph G is a tree (rooted for the modular decomposition and unrooted for the split decomposition) of which the leaves are in bijection with the vertices of G and whose internal nodes are labeled by indecomposable (for the chosen decomposition) graphs; such trees are called *graph-labeled trees* by Gioan and Paul [20]. Moreover, there is a one-to-one correspondence between such trees and graphs. The notion of a graph being *totally decomposable* for a decomposition scheme translates into restrictions on the labels that can appear on the internal nodes of its decomposition tree. For example, for the split-decomposition, totally decomposable graphs are the graphs whose decomposition tree's internal nodes are labeled only by cliques and stars; such graphs are called *distance-hereditary graphs*. They generalize the well-known *cographs*, the graphs that are totally decomposable for the modular decomposition, and whose enumeration has been well studied, in particular by Ravelomanana and Thimonier [26], also using techniques from analytic combinatorics

Efficiently encoding graph classes[1] is naturally linked to the enumeration of such graph classes. Indeed the number of graphs of a given class on n vertices implies a lower bound on the best possible encoding one can hope for. Until recently, few enumerative properties were known for distance-hereditary graphs, unlike their counterpart for the modular decomposition, the cographs. The best result so far, by Nakano *et al.* [24], relies on a relatively complex encoding on $4n$ bits, whose detailed analysis shows that there are at most $2^{\lfloor 3.59n \rfloor}$ unlabeled distance-hereditary graphs on n vertices. However, using the same techniques, their result also implies an upper-bound of 2^{3n} for the number of unlabeled cographs on n vertices, which is far from being optimal for these graphs, as it is known that, asymptotically, there are $Cd^n/n^{3/2}$ such graphs where $C = 0.4126\ldots$ and $d = 3.5608\ldots$ [26]. This suggests there is room for improving the best upper bound on the number of distance-hereditary graphs provided by Nakano *et al.* [24], which was the main purpose of our present work.

This paper. Following a now well established approach, which enumerates graph classes through a tree representation, when available (see for example the survey by Giménez and Noy [19] on tree-decompositions to count families of planar graphs), we provide *combinatorial specifications*, in the sense of Flajolet and Sedgewick [17], of the split-decomposition trees of distance-hereditary graphs and 3-leaf power graphs, both in the labeled and unlabeled cases. From these specifications, we can provide *exact enumerations*, *asymptotics*, and leave open the possibility of uniform random samplers allowing for further empirical studies of

*Dept. of Mathematics, Simon Fraser University, 8888 University Drive, V5A 1S6, Burnaby (BC), Canada, cedric.chauve@sfu.ca

†CNRS & LIX, École Polytechnique, 91120 Palaiseau, France, fusy@lix.polytechnique.fr

‡Dept. of Computer Science, Princeton University, 35 Olden Street, Princeton, NJ 08540, USA, lumbroso@cs.princeton.edu

[1] By which we mean, describing any graph from a class with as few bits as possible, as described for instance by Spinrad [28].

statistics on these graphs (see Iriza [23]).

In particular, we show that the number of distance-hereditary graphs on n vertices is bounded from above by 2^{3n}, which naturally opens the question of encoding such graphs on $3n$ bits, instead of $4n$ bits as done by Nakano *et al.* [24]. We also provide similar results for 3-leaf power graphs, an interesting class of distance hereditary graphs, showing that the number of 3-leaf power graphs on n vertices is bounded from above by 2^{2n}.

Main results. Our main contribution is to introduce the idea of symbolically specifying the trees arising from the split-decomposition, so as to provide the (previously unknown) exact enumeration of certain important classes of graphs.

Our grammars for distance-hereditary graphs are in Subsection 3, and our grammars for 3-leaf power graphs are in Subsection 2. We provide here the corollary that gives the beginning of the exact enumerations for the unlabeled and unrooted versions of both classes[2].

Corollary 1 (Enumeration of connected, unlabeled, unrooted distance-hereditary graphs)**.** *The first few terms of the enumeration,* EIS **A277862***, are*

$$1, 1, 2, 6, 18, 73, 308, 1484, 7492, 40010, 220676, 1253940, 7282316, 43096792, 259019070, 1577653196, 9720170360, 60492629435\dots$$

and the asymptotics is $c \cdot 7.249751250\dots^n \cdot n^{-5/2}$ *with* $c \approx 0.02337516194\dots$.

Corollary 2 (Enumeration of connected, unlabeled, unrooted 3-leaf power graphs)**.** *The first few terms of the enumeration,* EIS **A277863***, are*

$$1, 1, 2, 5, 12, 32, 82, 227, 629, 1840, 5456, 16701, 51939, 164688, 529070, 1722271, 5664786, 18813360, 62996841, 212533216\dots$$

and the asymptotics is $c \cdot 3.848442876\dots^n \cdot n^{-5/2}$ *with* $c \approx 0.70955825396\dots$.

1 Definitions and Preliminaries

For a graph G, we denote by $V(G)$ its vertex set and $E(G)$ its edge set. Moreover, for a vertex x of a graph G, we denote by $N(x)$ the neighbourhood of x, that is the set of vertices $y \in V(G)$ such that $\{x, y\} \in E(G)$; this notion extends naturally to vertex sets: if $V_1 \subseteq V(G)$, then $N(V_1)$ is the set of vertices in $V(G) \setminus V_1$ that is adjacent to at least one vertex in V_1. Finally, the subgraph of G induced by a subset V_1 of vertices is denoted by $G[V_1]$.

A graph on n vertices is *labeled* if its vertices are identified with the set $\{1, \dots, n\}$, with no two vertices having the same label. A graph is *unlabeled* if its vertices are indistinguishable.

A clique on k vertices, denoted K_k is the complete graph on k vertices (*i.e.*, there exists an edge between every pair of vertices). A star on k vertices, denoted S_k, is the graph with one vertex of degree $k-1$ (the *center* of the star) and $k-1$ vertices of degree 1 (the *extremities* of the star).

1.1 Split-decomposition trees.

We first introduce the notion of *graph-labeled tree*, due to Gioan and Paul [20], then define the split-decomposition and the corresponding tree, described as a graph-labeled tree.

Definition 1. A graph-labeled tree $(T, \mathcal{F})$ is a tree[3] T in which every internal node v of degree k is labeled by a graph $G_v \in \mathcal{F}$ on k vertices, such that there is a bijection ρ_v from the edges of T incident to v to the vertices of G_v.

Definition 2. A *split* [9] of a graph G with vertex set V is a bipartition (V_1, V_2) of V (*i.e.*, $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$) such that

(a) $|V_1| > 1$ and $|V_2| > 1$;
(b) every vertex of $N(V_1)$ is adjacent to every of $N(V_2)$.

A graph without any split is called a *prime* graph. A graph is *degenerate* if any partition of its vertices without a singleton part is a split: cliques and stars are the only such graphs.

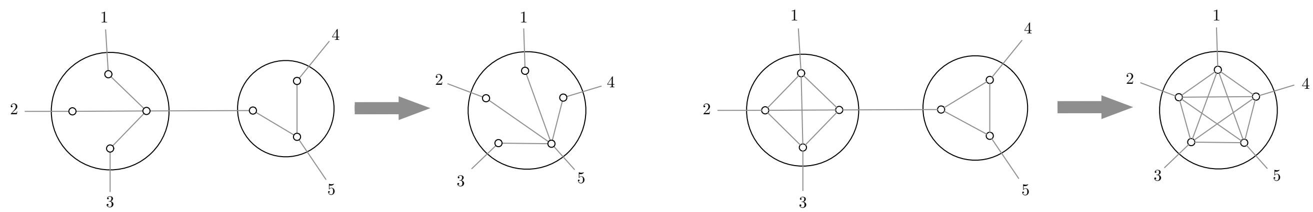
Informally, the split-decomposition of a graph G consists in finding a split (V_1, V_2) in G, followed by decomposing G into two graphs $G_1 = G[V_1 \cup \{x_1\}]$ where $x_1 \in N(V_1)$ and $G_2 = G[V_2 \cup \{x_2\}]$ where $x_2 \in N(V_2)$ and then recursively decomposing G_1 and G_2. This decomposition naturally defines an unrooted tree structure of which the internal vertices are labeled by degenerate or prime graphs and whose leaves are in bijection with the vertices of G, called a *split-decomposition tree*. A split-decomposition tree $(T, \mathcal{F})$ with $\mathcal{F}$ containing only cliques with at least three vertices and stars with at least three vertices is called a *clique-star tree*.

It can be shown that the split-decomposition tree of a graph might not be unique (*i.e.*, that several decompositions sequences of a given graph can lead to different split-decomposition trees), but following Cunningham [9], we obtain the following uniqueness result, reformulated in terms of graph-labeled trees by Gioan and Paul [20].

Theorem (Cunningham [9])**.** *For every connected graph G, there exists a unique split-decomposition tree such that:*

[2] With the symbolic grammars, it is then easy to establish recurrences [18, 29] to efficiently compute the enumeration–to the extent that we were trivially able to obtain the first 10 000 terms of the enumerations. See a survey by Flajolet and Salvy [16, §1.3] for more detail.

[3] This is a non-plane tree: the ordering of the children of an internal node does not matter—this is why in most of our grammars we describe the children as a SET instead of a SEQ, a sequence.



(a) Example of a star-join.

(b) Example of a clique-join.

Figure 1. The star-join and clique-join operations result in the merging of two internal nodes of a split-decomposition tree. A split-decomposition tree in which neither one of these operations may be applied (and in which all non-clique and non-star nodes are prime nodes) is said to be *reduced*.

(a) every non-leaf node has degree at least three;
(b) no tree edge links two vertices with clique labels;
(c) no tree edge links the center of a star to the extremity of another star.

Such a tree is called *reduced*, and this theorem establishes a one-to-one correspondence between graphs and their reduced split-decomposition trees. So enumerating the split-decomposition trees of a graph class provides an enumeration for the corresponding graph class, and we rely on this property in the following sections.

1.2 Decomposable structures. In order to enumerate classes of split-decomposition trees, we use the framework of decomposable structures, described by Flajolet and Sedgewick [17]. We refer the reader to this book for details and outline below the basics idea.

We denote by $\mathcal{Z}$ the combinatorial family composed of a single object of size 1, usually called *atom* (in our case, these refer to a leaf of a split-decomposition tree, *i.e.*, a vertex of the corresponding graph).

Given two disjoint families $\mathcal{A}$ and $\mathcal{B}$ of combinatorial objects, we denote by $\mathcal{A} + \mathcal{B}$ the *disjoint union* of the two families and by $\mathcal{A} \times \mathcal{B}$ the *Cartesian product* of the two families.

Finally, we denote by $\textsc{Set}(\mathcal{A})$ (resp. $\textsc{Set}_{\geqslant k}(\mathcal{A})$, $\textsc{Set}_k(\mathcal{A})$) the family defined as all sets (resp. sets of size at least k, sets of size exactly k) of objects from $\mathcal{A}$, and by $\textsc{Seq}_{\geqslant k}(\mathcal{A})$, the family defined as all sequences of at least k objects from $\mathcal{A}$.

Remark 1. Because this paper deals with classes both rooted (either at a vertex/leaf or an internal node) and unrooted, we use some notations to keep these distinct. But these notations are purely for clarity.

For instance, while we use $\mathcal{Z}_\bullet$ to denote a *rooted* vertex, and $\mathcal{Z}$ to denote an *unrooted* vertex, these are both translated in the same way in the associated generating functions and enumerations.

Remark 2. Decomposable structures specified by these grammars can either be:

- *labeled*: in a given object, each atom is labeled by a distinct number between 1 and n (the size of the object); this means that each "skeleton" of an object appears in $n!$ copies, for each of the possible way of labeling its individual atoms, and because each atom is distinguished, there are no symmetries;
- or *unlabeled*: in which case, an atom is indistinguishable from the next, and so certain symmetries must be taken into account (so that two objects which are not decomposed in the same way but have the same ultimate shape are not counted twice).

It is often the case that enumerations for labeled classes are easier to obtain than for unlabeled ones. Our grammars allow to derive generating functions, enumerations, and asymptotics for both.

2 3-Leaf Power Graphs

The first class that we discuss is that of 3-leaf power graphs: a chordal subset of distance-hereditary graphs[4].

Definition 3. A graph $G = (V, E)$ is a k-leaf power graph[5], if there is a tree T (called a k-leaf root of graph G) such that:
(a) the leaves of T are the vertices V;
(b) there is an edge $xy \in E$ if and only if the distance in T between leaves x and y is at most k, $d_T(x, y) \leqslant k$.

These families of graphs are relevant to phylogenetics [25]: from the the pairwise genetic distance between a collection of species (which is a graph), it is desirable to establish a tree which highlights the most likely ancestry (or more broadly, the evolutionary relationships) relations between species.

We begin with the enumeration of 3-leaf power graphs, the smaller combinatorial class, because the application of

[4] Not a maximal such subset, as it is known that *ptolemaic graphs* are the intersection of chordal graphs and distance-hereditary graphs.

[5] This is a specialization, introduced by Nishimura *et al.* [25, §1], of the concept of *graph powers*, in which the root is a tree—but the definition can be extended to the case where T is not a tree, but is a graph H (in which case, we consider the distance between any two vertices in graph H, not two leaves of a tree).

the dissymmetry theorem (used to obtain an enumeration of the unrooted class given the grammar for some rooted version of the class) in Subsection 2.3 is less involved for 3-leaf power graphs than it is for distance-hereditary graphs.

2.1 Grammar[6] from the split-decomposition.

The starting point is the characterization of the split-decomposition tree of 3-leaf power graphs, as introduced by Gioan and Paul [20].

Theorem (Characterization of 3-leaf power split-decomposition tree [20, § 3.3])**.** *A connected graph $G = (V, E)$ is a 3-leaf power graph if and only if:*

(a) its split-decomposition tree $ST(G)$ is a clique-star tree (implies that G is distance-hereditary);

(b) the set of star-nodes forms a connected subtree of T;

(c) the center of a star-node is incident either to a leaf or a clique node.

This is unsurprising given that an alternate (perhaps more pertinent) characterization is that a 3-leaf power graph can be obtained from a tree by replacing every vertex by a clique of arbitrary size.

Theorem 1. *The class $3\mathcal{LP}_\bullet$ of 3-leaf power graphs rooted at a vertex*[7] *is specified by*

$$3\mathcal{LP}_\bullet = \mathcal{L}_\bullet \times (\mathcal{S}_C + \mathcal{S}_X) + \mathcal{K}_\bullet \tag{2.1}$$

$$\mathcal{S}_C = \textsc{Set}_{\geqslant 2}(\mathcal{L} + \mathcal{S}_X) \tag{2.2}$$

$$\mathcal{S}_X = \mathcal{L} \times \textsc{Set}_{\geqslant 1}(\mathcal{L} + \mathcal{S}_X) \tag{2.3}$$

$$\mathcal{L} = \mathcal{Z} + \textsc{Set}_{\geqslant 2}(\mathcal{Z}) \tag{2.4}$$

$$\mathcal{L}_\bullet = \mathcal{Z}_\bullet + \mathcal{Z}_\bullet \times \textsc{Set}_{\geqslant 1}(\mathcal{Z}) \tag{2.5}$$

$$\mathcal{K}_\bullet = \mathcal{Z}_\bullet \times \textsc{Set}_{\geqslant 2}(\mathcal{Z}). \tag{2.6}$$

In this combinatorial specification, we define several classes of subtrees: we denote by $\mathcal{S}_X$ (resp. $\mathcal{S}_C$) the class of split-decomposition trees rooted at a star-node which are *linked to their parent* by an extremity of this star-node (resp. the center of this star-node).

Finally, because the structure of the split-decomposition tree of a 3-leaf power graph only allows for cliques that are incident to at most one star-node (and the rest of the edges must lead to leaves), we have three classes $\mathcal{L}$, $\mathcal{L}_\bullet$ and $\mathcal{K}_\bullet$ which express leaves and cliques[8].

Proof. In addition to the constraints specific to 3-leaf power split-decomposition trees given in the characterization above, because the split-decomposition trees we are enumerating are *reduced* (see Cunningham's Theorem in Section 1.1), there are two additional implicit constraints on their internal nodes:

- the center of a star cannot be incident to the extremity of another star (because then they would be merged with a *star-join* operation, as in Figure 1a, yielding a more concise split-decomposition tree);
- and two cliques may not be incident (or they would be merged with a *clique-join* operation, as in Figure 1b).

The star-nodes form a connected subtree, with each star-node connected to others through their extremities; the centers are necessarily connected to "leaves", and the extremities may be connected to "leaves"; "leaves" are either single nodes (actually leaves) or cliques (which are a set of more than two elements, because cliques have minimum size of 3 overall, including the parent node).

First, the following equation

$$\mathcal{S}_C = \textsc{Set}_{\geqslant 2}(\mathcal{L} + \mathcal{S}_X)$$

indicates that a subtree rooted at a star-node, linked to its parent (presumably a leaf) by its center, is a set of size at least 2 children, which are the extremities of the star-node: each extremity can either lead to a "leaf" or to another star-node entered through an extremity.

Next, the equation

$$\mathcal{S}_X = \mathcal{L} \times \textsc{Set}_{\geqslant 1}(\mathcal{L} + \mathcal{S}_X)$$

indicates that a subtree rooted at a star-node, linked to its parent by an extremity, is the Cartesian product of a "leaf" (connected through the center of the star-node) and a set of 1 or more children which are the extremities of the star-node: each leads either to a "leaf" or to another star-node entered through an extremity.

The "leaves" are then either an actual leaf of unit size, or a clique; the clique has to be of size at least 3 (including the incoming link) and the children can only be actual leaves. We are thus left with

$$\mathcal{L} = \mathcal{Z} + \textsc{Set}_{\geqslant 2}(\mathcal{Z}).$$

Finally, the rest of the grammar deals with the special cases that arise from when the split-decomposition tree does not contain any star-node at all. □

With the grammar for $3\mathcal{LP}_\bullet$, we are able to produce the exact enumeration for labeled rooted 3-leaf power graphs, and by a simple algebraic trick, for unlabeled rooted 3-leaf power graphs.

[6] All grammars that we produce in this article yield an incorrect enumeration for the first two terms (graphs of size 1 and 2), because Cunningham's Theorem, presented in Subsection 1.1 requires non-leaf nodes to have degree at least three: thus the special cases of graphs involving only 1 or 2 nodes must be treated non-recursively. While we could amend the grammars accordingly, we think it would be less elegant—especially since there is generally little confusion regarding those first few terms.

[7] Or, equivalently, rooted at a leaf of its split-decomposition tree.

[8] The class $\mathcal{L}$ is a class containing either a leaf; or a clique-node connected to all but one of its extremities to leaves. The class $\mathcal{L}_\bullet$ is that same class, in which one of the leaves has been distinguished (as the root of the tree).

Corollary 3 (Enumeration of labeled 3-leaf power graphs)**.** *Let $T(z)$ be the exponential generating function associated with the class $3\mathcal{LP}_\bullet$. Then, the enumeration of labeled,* unrooted 3*-leaf power graphs, for $n \geqslant 3$, is given by*

$$t_n = (n-1)![z^n]T(z), \tag{2.7}$$

to the effect that the first few terms of the enumeration, EIS **A277868***, are*

$$1, 1, 4, 35, 361, 4482, 68027, 1238841, 26416474, \\ 646139853, 17837851021, 548713086352, \ldots$$

2.2 Unrooting unlabeled objects. The trees described by the specification of $3\mathcal{LP}_\bullet$ have leaves which are labeled, one of which is the root. Thus because each label has equal opportunity of being the root, it is simple to obtain an enumeration of the labeled *unrooted* class by dividing by n.

When now considering *unlabeled* trees, however, proceeding in this way leads to an overcount of certain trees, because of new symmetries introduced by the disappearance of labels. Fortunately, we can use the *dissymmetry theorem for trees*, which expresses the enumeration of an unrooted class of trees in terms of the enumeration of the equivalent rooted class of trees.

This theorem was introduced by Bergeron *et al.* [2] in terms of ordered and unordered pairs of trees, and was eventually reformulated in a more elegant manner, such as in Flajolet and Sedgewick [17, VII.26 p. 481] or Chapuy *et al.* [6, §3]. It states

$$\mathcal{A} + \mathcal{A}_{\circ\to\circ} \simeq \mathcal{A}_\circ + \mathcal{A}_{\circ-\circ} \tag{2.8}$$

where $\mathcal{A}$ is the unrooted class of trees, and $\mathcal{A}_\circ, \mathcal{A}_{\circ-\circ}, \mathcal{A}_{\circ\to\circ}$ are the rooted class of trees respectively where only the root is distinguished, an edge from the root is distinguished, and a directed, outgoing edge from the root is distinguished[9].

The application of this theorem may initially be perplexing, and so we begin by making a couple of remarks.

Lemma 1. *In the dissymmetry theorem for trees, when rerooting at the nodes (or atoms) of a combinatorial tree-like class $\mathcal{A}$, leaves can be ignored.*

Proof. When we point a node of the class $\mathcal{A}$, we may distinguish whether it is an internal node or a leaf, which we respectively denote $\diamond$ and $\bullet$ in the *right hand side* of the following equations (please bear these notations; the purpose of this lemma is to show they can be safely omitted in what follows). Accordingly,

$$\begin{aligned} \mathcal{A}_{\circ\to\circ} &= \mathcal{A}_{\bullet\to\diamond} + \mathcal{A}_{\diamond\to\diamond} + \mathcal{A}_{\diamond\to\bullet} \\ \mathcal{A}_{\circ-\circ} &= \mathcal{A}_{\bullet-\diamond} + \mathcal{A}_{\diamond-\diamond} \\ \mathcal{A}_\circ &= \mathcal{A}_\bullet + \mathcal{A}_\diamond \end{aligned}$$

where the first equation should be understood as: if we mark a directed edge of the class $\mathcal{A}$, it can either go from an internal node to a leaf, from a leaf to an internal node, or from an internal node to another internal node[10].

These equations may be further simplified upon observing that any edge of which one of the endpoints is a leaf, is entirely determined by that leaf, to the effect that

$$\mathcal{A}_\bullet = \mathcal{A}_{\bullet-\diamond} = \mathcal{A}_{\bullet\to\diamond} = \mathcal{A}_{\bullet\to\diamond}.$$

Thus proving that one may disregard leaves entirely when applying the dissymmetry theorem for trees. □

Remark 3. While the dissymmetry theorem considers pointed internal nodes, our grammars $3\mathcal{LP}_\bullet$ and $\mathcal{DH}_\bullet$ (respectively derived from the split-decomposition of 3-leaf power graphs and distance-hereditary graphs) are pointed at the *leaves* of the split-decomposition tree (which correspond to the vertices of the original graph).

This is not, in fact, a discrepancy. When we apply the dissymmetry theorem, we implicitly *reroot* the trees from our grammars at internal nodes, which we express as subclasses $\mathcal{T}_x$ of trees rooted in some specific type of internal node x. Rerooting an already rooted tree is relatively easy (while unrooting a rooted tree is not!).

Remark 4. The dissymmetry theorem establishes a bijection between two disjoint unions; this allows us to recover an equation on the coefficients,

$$\begin{aligned} [z^n]A(z) = [z^n]&A_\circ(z) \\ &+ [z^n]A_{\circ-\circ}(z) \\ &- [z^n]A_{\circ\to\circ}(z). \end{aligned} \tag{2.9}$$

However the subtraction has no combinatorial meaning, which means that once the dissymetry theorem has been applied, we lose the symbolic meaning of the equation.

While this is enough to compute exact enumerations (by extracting the enumeration of each generating function and algebraically computing the equation), and is sufficient to deduce some asymptotics, there is not enough information, for instance, to yield a recursive sampler [18] or a Boltzmann sampler [14, 15]—and we are instead left with ad-hoc methods to generate unrooted objects, see Iriza [23, § 3.2].

[9] Drmota [12, §4.3.3, p. 293] presents an elegant proof of this result by appealing to the notion of *center* of the tree—which may be a single vertex or an edge; indeed, Drmota builds a bijection between the trees $\mathcal{A}_{\mathrm{nc}}$ rooted at a non-central vertex/edge and trees rooted at a directed edge, by orienting the root of the first class in the direction of the center.

[10] We are ignoring the very special case of a tree reduced to an edge, in which we may have an edge between two leaves; this explains why our unrooted grammars may, if uncorrected, be wrong for the first two terms. This is analogous to the initial term errors of our rooted grammars, as expressed in Footnote 6.

Unrooting the initial grammar while preserving the symbolic nature of the specification requires using a more complex combinatorial tool called *cycle-pointing*[11] introduced by Bodirsky *et al.* [4], and applied to these grammars by Iriza [23, §5.5], it has allowed us to generate the random graphs provided in figures to this article.

2.3 Applying the dissymmetry theorem.

Theorem 2. *The class $3\mathcal{LP}$ of unrooted 3-leaf power graphs is specified by*

$$3\mathcal{LP} = \mathcal{K} + \mathcal{T}_S + \mathcal{T}_{S-S} - \mathcal{T}_{S\to S} \tag{2.10}$$
$$\mathcal{T}_S = \mathcal{L} \times \mathcal{S}_C \tag{2.11}$$
$$\mathcal{T}_{S-S} = \mathrm{SET}_2\left(\mathcal{S}_X\right) \tag{2.12}$$
$$\mathcal{T}_{S\to S} = \mathcal{S}_X \times \mathcal{S}_X \tag{2.13}$$
$$\mathcal{S}_C = \mathrm{SET}_{\geq 2}\left(\mathcal{L} + \mathcal{S}_X\right) \tag{2.14}$$
$$\mathcal{S}_X = \mathcal{L} \times \mathrm{SET}_{\geqslant 1}\left(\mathcal{L} + \mathcal{S}_X\right) \tag{2.15}$$
$$\mathcal{L} = \mathcal{Z} + \mathrm{SET}_{\geqslant 2}\left(\mathcal{Z}\right) \tag{2.16}$$
$$\mathcal{K} = \mathrm{SET}_{\geq 3}\left(\mathcal{Z}\right). \tag{2.17}$$

Proof. From the dissymmetry theorem, we have the symbolic equation linking the rooted and unrooted decomposition tree of 3-leaf power graphs,

$$3\mathcal{LP} = 3\mathcal{LP}_\circ + 3\mathcal{LP}_{\circ-\circ} - 3\mathcal{LP}_{\circ\to\circ}.$$

As per Lemma 1, it suffices to consider only internal nodes, and the only type of internal node found in these split-decomposition trees is the star-node[12].

So we must reroot the grammar $3\mathcal{LP}_\bullet$, which is rooted at a leaf of the split-decomposition tree, to each of: a star-node, an undirected edge connecting two star-nodes, and a directed edge connecting two star-nodes.

Rerooting at a star-node, we must consider all the outgoing edges of the star. The center will lead either to a leaf, or to a clique—this is the rule $\mathcal{L}$; what remains are then the extremities, which can be expressed by the term $\mathcal{S}_C$. Since the center is distinguished, this is combined as a Cartesian product, hence

$$\mathcal{T}_S = \mathcal{L} \times \mathcal{S}_C.$$

Next, we reroot at an edge. Since these split-decomposition trees are reduced, two star-nodes can only be adjacent at their respective centers, or at two extremities; but because of the additional constraint for 3-leaf power graphs, two star-nodes can only be adjacent at their extremities.

Rerooting at an undirected edge will yield a set containing two elements; rerooting at a directed edge will yield a Cartesian product. Thus, we have

$$\mathcal{T}_{S-S} = \mathrm{SET}_2\left(\mathcal{S}_X\right)$$
$$\mathcal{T}_{S\to S} = \mathcal{S}_X \times \mathcal{S}_X.$$

Finally, as with the original vertex-rooted grammar $3\mathcal{LP}_\bullet$, we must deal with the special case of a graph reduced to a clique, as it does not involve any star-node. □

Corollary 2 (Enumeration of unlabeled, unrooted 3-leaf power graphs)**.** *The first few terms of the enumeration,* EIS **A277863***, are*

$$1, 1, 2, 5, 12, 32, 82, 227, 629, 1840, 5456, 16701, 51939, 164688, 529070, 1722271, 5664786, 18813360, 62996841, 212533216\ldots$$

3 Distance-Hereditary Graphs

A graph is *totally decomposable* by the split-decomposition if every induced subgraph with at least 4 vertices contains a split. And it is well-known [22] that the class of totally decomposable graphs is exactly distance-hereditary graphs.

Deriving the rooted grammar provided in Theorem 3 is easier than for 3-leaf power graphs, because there are few constraints on the split-decomposition tree of distance-hereditary graphs; as a result, applying the dissymmetry theorem will be a bit more involved because there are two types of internal nodes at which to reroot the tree.

Theorem 3. *The class $\mathcal{DH}_\bullet$ of distance-hereditary graphs rooted at a vertex is specified by*

$$\mathcal{DH}_\bullet = \mathcal{Z}_\bullet \times \left(\mathcal{K} + \mathcal{S}_C + \mathcal{S}_X\right) \tag{3.18}$$
$$\mathcal{K} = \mathrm{SET}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{S}_C + \mathcal{S}_X\right) \tag{3.19}$$
$$\mathcal{S}_C = \mathrm{SET}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X\right) \tag{3.20}$$
$$\mathcal{S}_X = \mathrm{SEQ}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C\right). \tag{3.21}$$

Proof. We describe a grammar for clique-star trees subject only to the irreducibility constraint: a star's center cannot be connected to the extremity of another star (see Figure 1a), and two cliques cannot be connected (see Figure 1b).

We start with the following rule

$$\mathcal{DH}_\bullet = \mathcal{Z}_\bullet \times \left(\mathcal{K} + \mathcal{S}_C + \mathcal{S}_X\right)$$

in which $\mathcal{Z}_\bullet$, the vertex at which the split-decomposition tree is rooted, can be connected either to a clique $\mathcal{K}$, or to a star's extremity $\mathcal{S}_X$, or to a star's center $\mathcal{S}_C$.

Next, we describe subtrees rooted at a clique

$$\mathcal{K} = \mathrm{SET}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{S}_C + \mathcal{S}_X\right),$$

[11] This operation, given a structure of size n, finds n ways to group its atoms/vertices in cycles which mirror the symmetries of the structure. This is analogous to atom/vertex-pointing in labeled objects, where each structure of size n can be pointed n different ways (each atom/vertex can be pointed because they are each distinguishable and there are no symmetries that would make two different pointings equivalent).

[12] In the split-decomposition of a 3-leaf power graph, clique-nodes cannot have any children other than leaves; as a result, they may be considered as leaves for the purpose of the dissymmetry theorem.

we are connected to our parent by one of the outgoing edges of the clique, and because clique-nodes have size at least 3 (see Cunningham's Theorem in Subsection 1.1 which requires non-leaf nodes to have degree at least 3), we are left with at least two subtrees to describe:

- these subtrees can either be a leaf $\mathcal{Z}$, or a star entered either by its center $\mathcal{S}_C$ or its extremity $\mathcal{S}_X$—they cannot be another clique because our tree could then be reduced with a clique-join operation;
- because of the symmetries within a clique (in particular there is no ordering of the vertices), the order of the subtrees does not matter, and so these are described by a SET operation.

By similar arguments, we describe subtrees rooted at a star which is connected to its parent by its center,

$$\mathcal{S}_C = \textsc{Set}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X\right).$$

Because the star's center is connected to its parent, we need only express what the extremities are connected to; each of these can be connected to a leaf, a clique, or another star by one of that star's extremity (to avoid a star-join). Again, as the extremities are indistinguishable from each other—the star is not planar—we describe the subtrees by a SET operation.

We are left with the subtrees rooted at star which is connected to its parent by an extremity; these may be described by

$$\mathcal{S}_X = \left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C\right) \times \textsc{Set}_{\geqslant 1}\left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X\right).$$

Indeed, the first term of the Cartesian product is the subtree to which the center is connected (either a leaf, a clique, or another star at its center); the SET expresses the remaining extremities—of which there is at least one. This equation can be simplified to obtain the one in the Theorem—but this simplification is proven in Appendix A. □

Remark 5. We notice the same symbolic rules for the clique-node $\mathcal{K}$ and the star-node $\mathcal{S}_C$ entered through the center, respectively in Equations (3.19) and (3.20). This suggest these nodes play a symmetrical role in the overall grammar, and that their associated generating function (and enumeration) are identical.

It would be mathematically correct to merge both rules, *e.g.*

$$\mathcal{DH}_\bullet = \mathcal{Z}_\bullet \times (\mathcal{K} + \mathcal{K} + \mathcal{S}) \tag{3.22}$$
$$\mathcal{K} = \textsc{Set}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{K} + \mathcal{S}\right) \tag{3.23}$$
$$\mathcal{S} = \textsc{Seq}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{K} + \mathcal{K}\right). \tag{3.24}$$

This may be convenient (and lead to additional simplifications) for some uses, such as the application of asymptotic theorems like those introduced by Drmota [11] (see Section 4) which requires classes be expressed as a single functional equation.

However the combinatorial meaning of the symbols is lost: in the above system, it can no longer be said that $\mathcal{K}$ represents a clique-node. This is problematic for parameter analysis (*e.g.*, if trying to extract the average number of clique-nodes in the split-tree of a uniformly drawn distance-hereditary graph).

Theorem 4. *The class $\mathcal{DH}$ of unrooted distance-hereditary graphs is specified by*

$$\mathcal{DH} = \mathcal{T}_K + \mathcal{T}_S + \mathcal{T}_{S-S} - \mathcal{T}_{K-S} - \mathcal{T}_{S\to S} \tag{3.25}$$
$$\mathcal{T}_K = \textsc{Set}_{\geqslant 3}\left(\mathcal{Z} + \mathcal{S}_C + \mathcal{S}_X\right) \tag{3.26}$$
$$\mathcal{T}_S = \left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C\right) \times \mathcal{S}_C \tag{3.27}$$
$$\mathcal{T}_{K-S} = \mathcal{K} \times \left(\mathcal{S}_C + \mathcal{S}_X\right) \tag{3.28}$$
$$\mathcal{T}_{S-S} = \textsc{Set}_2\left(\mathcal{S}_C\right) + \textsc{Set}_2\left(\mathcal{S}_X\right) \tag{3.29}$$
$$\mathcal{T}_{S\to S} = \mathcal{S}_C \times \mathcal{S}_C + \mathcal{S}_X \times \mathcal{S}_X \tag{3.30}$$
$$\mathcal{K} = \textsc{Set}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{S}_C + \mathcal{S}_X\right) \tag{3.31}$$
$$\mathcal{S}_C = \textsc{Set}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X\right) \tag{3.32}$$
$$\mathcal{S}_X = \textsc{Seq}_{\geqslant 2}\left(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C\right). \tag{3.33}$$

Proof. This is again an application of the dissymmetry theorem for trees, and as before, we may ignore the leaves, and mark only the internal nodes,

$$\mathcal{DH} = \mathcal{DH}_\circ + \mathcal{DH}_{\circ-\circ} - \mathcal{DH}_{\circ\to\circ}.$$

Unlike for 3-leaf power graphs in Subsection 2.3, the tree decomposition of distance-hereditary graphs clearly involves two types of internal nodes: cliques and stars. If we express all the rerooted trees we will have to express, we get the expression:

$$\begin{aligned}\mathcal{DH} = {}& \mathcal{T}_K + \mathcal{T}_S \\ & + \mathcal{T}_{S-S} + \mathcal{T}_{S-K} \\ & - \mathcal{T}_{S\to S} - \mathcal{T}_{S\to K} - \mathcal{T}_{K\to S}.\end{aligned} \tag{3.34}$$

Note that we do not have a tree rerooted at an edge involving two cliques, because as mentioned previously, the split-decomposition tree would not be reduced, since the two cliques could be merged with a clique-join.

A first simplification can be made, because a directed edge linking two internal nodes of different type is equivalent to a non-directed edge, because the nature of the two internal nodes already distinguishes them, thus in particular here

$$\mathcal{T}_{K\to S} \simeq \mathcal{T}_{K-S}.$$

In doing so, several terms cancel out, which leads us to:

$$\mathcal{DH} = \mathcal{T}_K + \mathcal{T}_S + \mathcal{T}_{S-S} - \mathcal{T}_{K-S} - \mathcal{T}_{S\to S}.$$

We then only have to express the rerooted classes:

$\mathcal{T}_K$ For a clique-node, we must account for at least three outgoing edges, which can be connected to anything besides another a clique-node.

$\mathcal{T}_S$ For a star-node, we reuse the same trick as previously: we express what the center can be connected to (either a leaf, a clique-node or the center of another star-node), and then we use $\mathcal{S}_C$ to express the remaining extremities, as explained in the unrooted grammar for the 3-leaf power graphs.

$\mathcal{T}_{K-S}$ The undirected edge already accounts for a connection between a clique-node and a star-node, so we must describe the remaining outgoing edges of these two combined nodes: for the clique, this can be expressed by reusing the subtree $\mathcal{K}$ (which is exactly a tree rooted at a clique which is missing one subtree—the one connected to the star-node); for the star, if it is connected to the clique-node by its extremity, we can use $\mathcal{S}_X$, otherwise $\mathcal{S}_C$.

$\mathcal{T}_{S-S}$ Two star-nodes can only be connected at two of their extremities, or their respective centers[13]; because the edge is undirected, we use a SET operation.

$\mathcal{T}_{S\to S}$ Same as above, except the edge now being directed, we use a Cartesian product to distinguish a source star-node and a destination star-node.

□

4 Asymptotics

Using singularity analysis of generating functions, we can now derive asymptotic estimates for the number of *unlabeled*[14] (rooted and unrooted) 3-leaf power graphs and distance-hereditary graphs, with respect to the number n of vertices.

The strategy in both cases is very similar, and uses a simplified version of the Drmota-Lalley-Woods theorem for tree-like objects, exposed both in the original article by Drmota [11] and in the Flajolet and Sedgewick book [17, §VII.6.3 p. 488].

The first step is to study the rooted case, starting from the decomposition grammar: Theorem 1 for 3-leaf power (3LP) graphs; Theorem 3, or rather the simplified version presented in Remark 5, for distance-hereditary (DH) graphs. This grammar can be translated into an equation system for the corresponding generating functions. Although this is not a prerequisite to analysis, the system is in fact sufficiently simple that, using suitable manipulations, it can be reduced to a single-line equation of the form $y = F(y, z)$, where y is one of the rooted generating functions and all the other functions have a simple expression in terms of y. The Drmota-Lalley-Wood theorem then ensures that the *rooted* generating functions classically have a square-root singularity, yielding asymptotic estimates of the form $c\cdot\rho^{-n}\cdot n^{-3/2}$.

The next step is to study the generating function $U(z)$ for *unrooted* 3LP (resp. DH) graphs. From the dissymmetry theorem (Theorem 2 for 3LP graphs, Theorem 4 for DH graphs), we obtain an expression of $U(z)$ in terms of y (the rooted generating function) and z, from which we can obtain a singular expansion of $U(z)$. As expected, the subtractions, from the dissymmetry theorem, involved in the expression of $U(z)$ yield a cancellation of the square-root terms, so that the leading singular terms are at the next order, yielding asymptotic estimates of the form $d \cdot \rho^{-n} \cdot n^{-5/2}$ (which are expected for unrooted "tree-like" structures).

Remark 6. A similar approach has been previously applied to another tree decomposition of graphs (decomposition into 2-connected blocks and a tree to describe the adjacencies between blocks). Based on this decomposition, the asymptotics for several families (both labeled and unlabeled) of graphs have been obtained (cacti graphs, outerplanar graphs [3], series-parallel graphs [13]), all of the form $c \cdot \rho^{-n} \cdot n^{-5/2}$.

4.1 Asymptotics of 3-leaf power graphs.

In this subsection, we will implement the method described above, and take the time to point out some of the finer points of the analysis. Most of this process is then recycled in the analysis of distance-hereditary graphs in the next subsection.

Rooted case. Our starting point is the grammar for rooted 3-leaf power graphs, as given by Theorem 1. We let $L(z)$ and $S_X(z)$ be the generating functions of $\mathcal{L}$ and $\mathcal{S}_X$. In particular, using the symbolic method, Eq. (2.3) in Theorem 1 translates into the functional equation:

$$S_X(z) = L(z) \cdot \left(\exp\left(\sum_{i\geqslant 1} \frac{1}{i}\left(L(z^i) + S_X(z^i)\right)\right) - 1\right).$$

Because our objects are unlabeled, we may identify a *set* of undistinguished elements with a *sequence* of undistinguished elements (this simplification is untrue in the labeled case),

$$\mathrm{SET}_{\geqslant 1}(\mathcal{Z}) \simeq \mathrm{SEQ}_{\geqslant}(\mathcal{Z}) \tag{4.35}$$

and thus let $L(z) = z/(1-z)$.

With this simplification, we find that $y := S_X(z)$ satisfies the functional equation

$$y = \frac{z}{1-z}\cdot\left(\exp\left(y + B(z)\right) - 1\right),$$

[13] For the 3-leaf power graphs, we only considered two star-nodes connected at two of their extremities, because part of the characterization of 3-leaf power graphs is that the center of stars are oriented away from other stars.

[14] As previously mentioned, labeled objects are a lot easier to enumerate because of the exact 1-to-n correspondance between labeled rooted and unrooted objects; therefore we do not study that case.

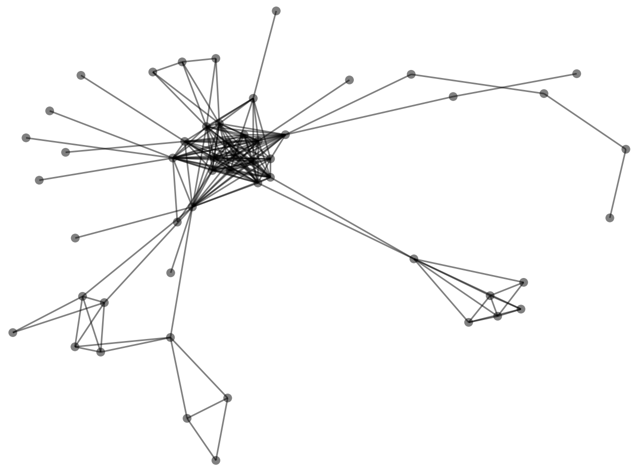


Figure 2. A randomly generated distance-hereditary graph with 52 vertices, produced using the Boltzmann samplers developed by Iriza [23].

where

$$B(z) := \sum_{i\geqslant 1} \frac{1}{i} \frac{z^i}{1-z^i} + \sum_{i\geqslant 2} \frac{1}{i} S_X(z^i).$$

This is a functional equation of the form $y = F(z, y)$, with $F(z, y)$ a bivariate formal power series with non-negative coefficients and that has nonlinear dependence on y.

We must show this power series exhibits certain properties, namely that it is a-positive, a-proper, and a-irreducible [17, p. 489]. To this end, let ρ be the radius of convergence of $S_X(z)$:

- The fact that $F(z, y)$ is superlinear in y, ensures that $S_X(z)$ converges to a finite positive value (which we denote by τ) when z tends to ρ from below.

- Furthermore, it is straightforward to combinatorially check that the coefficients of S_X, $[z^n]S_X(z)$, have exponential growth, *i.e.*, there is an $\alpha > 1$ such that $[z^n]S_X(z) \geqslant \alpha^n$, for large enough n. This in turn implies the radius of convergence ρ is such that $\rho < 1$, thus that $B(z)$ is analytic at ρ and hence $F(z, y)$ is analytic at (ρ, τ).

Under these conditions, the Drmota-Lalley-Woods theorem [17, Thm VII.6 p. 489] may be applied[15]. This results in a singular expansion of $S_X(z)$ around ρ of the form

$$S_X(z) = \tau - c \cdot Z + O\big(Z^2\big)$$

with

$$Z := \sqrt{1 - z/\rho} \quad \text{and} \quad c = \sqrt{2\rho \cdot \frac{F_z(\rho, \tau)}{F_{yy}(\rho, \tau)}}$$

where we use the notation $F_z := \partial F/\partial z$ for partial derivatives.

Moreover we can[16] apply classical transfer theorems [17, Thm VI.6 p. 404] to obtain

$$[z^n]S_X(z) \sim \frac{c}{\sqrt{2\pi}} \cdot \rho^{-n} \cdot n^{-3/2}.$$

Now, in order to evaluate the singularity ρ, we have to solve the system,

$$\begin{cases} y = F(z, y) \\ 1 = F_y(z, y) \end{cases} . \tag{4.36}$$

This presents a difficulty, due to the fact that $F(z, y)$ involves quantities of the form $S(z^i)$ for $i \geqslant 2$.

Following Flajolet and Sedgewick [17, §VII.5] however, we can accurately approximate these quantities by the truncated generating function ${S_X}^{[m]}(z^i)$, where ${S_X}^{[m]}(z)$ is the polynomial of degree m coinciding with the Taylor expansion of $S_X(z)$ to order m. Denoting by $F^{[m]}(z, y)$ the corresponding (now explicit) approximation of $F(z, y)$, we can solve for the system

$$\begin{cases} y = F^{[m]}(f, y) \\ 1 = {F_y}^{[m]}(f, y) \end{cases}$$

[15] It follows that the singularity of $S_X(z)$ is a square-root singularity, due to a branch point [17, p. 495]: this means in particular that $F_y(z, y) = 1$ and $(z, y) = (\rho, \tau)$.

[16] Because $S_X(z)$ has an analytic continuation to a Δ-domain of the form $\{|z| \leqslant \rho + \epsilon\} \cap \{z - \rho \notin \mathbb{R}_+\}$.

and the obtained solution $(\rho^{[m]}, \tau^{[m]})$ is found to converge exponentially fast as m increases; we find

$$\rho \approx 3.848442876\ldots$$

which is the exponential growth of 3-leaf power graphs.

Unrooted case. We now continue with the asymptotic enumeration of unrooted 3LP graphs. The grammar from Theorem 2 is a superset of the terms of Theorem 1, so our previous expressions, in particular for $S_X(z)$ and related, still hold. We now want to express the generating function $U(z)$ of unrooted 3LP graphs in terms of $S_X(z)$ and will show that the leading singular term is of order Z^3 due to a cancellation of the coefficients for terms of order Z.

To this end, will need to manipulate singular expansions up to order Z^3, and a first important fact is that, as an application of the Drmota-Lalley-Woods theorem, $S_X(z)$ admits such an expansion, of the form

$$S_X(z) = \tau - c \cdot Z + d \cdot Z^2 + e \cdot Z^3 + O\big(Z^3\big)\,.$$

Again, we define $U(z)$ as the generating function of unrooted 3LP graphs according to the number of vertices. It follows from the grammar given by Theorem 2 that

$$U(z) = K(z) + T_S(z) + T_{S-S}(z) - T_{S\to S}(z).$$

We now seek to express this as a single functional equation, for convenience. First, using the same combinatorial simplification of Eq. (4.35) that we used in the rooted case (which, again, is only possible because we are dealing with unlabeled objects), we have $K(z) = z^3/(1-z)$.

In order to express $T_S(z)$ in terms of $S_X(z)$, we make the following formal manipulations,

$$\begin{aligned}\mathcal{S}_X &= \mathcal{L} \times \textsc{Set}_{\geqslant 1}(\mathcal{L} + \mathcal{S}_X)\\ &= \mathcal{L} \times \textsc{Set}_{\geqslant 2}(\mathcal{L} + \mathcal{S}_X) + \mathcal{L} \times (\mathcal{L} + \mathcal{S}_X)\\ &= \mathcal{T}_S + \mathcal{L} \times (\mathcal{L} + \mathcal{S}_X)\end{aligned}$$

hence

$$T_S(z) = S_X(z) - L(z) \cdot (L(z) + S_X(z))$$

which, using our previous simplification of $L(z)$, yields

$$T_S(z) = S_X(z) - \frac{z}{1-z} \cdot \left(\frac{z}{1-z} + S_X(z)\right).$$

Finally, the rerooted subclasses, $\mathcal{T}_{S-S}$ and $\mathcal{T}_{S\to S}$, are fortunately already expressed in terms of the generating function of $\mathcal{S}_X$, as we have

$$T_{S-S}(z) = \frac{S_X(z^2) + S_X(z)^2}{2} \quad \text{and} \quad T_{S\to S}(z) = S_X(z)^2$$

so that by combining all these generating functions, we obtain

$$\begin{aligned}U(z) = \frac{z^3}{1-z} + \frac{S_X(z^2)}{2} + S_X(z)\\ - \frac{z}{1-z} \cdot \left(\frac{z}{1-z} + S_X(z)\right) - \frac{S_X(z)^2}{2}.\end{aligned}$$

This is of the form $U(z) = G(z, S_X(z))$, where we define

$$\begin{aligned}G(z, y) := \frac{z^3}{1-z} + \frac{S_X(z^2)}{2} + y\\ - \frac{z}{1-z} \cdot \left(\frac{z}{1-z} + y\right) - \frac{y^2}{2}.\end{aligned}$$

Here too, $G(z, y)$ is analytic at (ρ, τ) (as before, because $\rho < 1$ and $S_X(z^2)$ is analytic at ρ). As a consequence, $U(z)$ has a singular expansion at ρ of the form

$$U(z) = \tau' - c' \cdot Z + d' \cdot Z^2 + e' \cdot Z^3 + O\big(Z^4\big)\,,$$

again, with $Z := \sqrt{1 - z/\rho}$. Around ρ, we have

$$\begin{aligned}U(z) &= G(z, S_X(z))\\ &= G(\rho + (z-\rho), \tau - c \cdot Z + O(Z^2))\\ &= G(\rho, \tau) + G_y(\rho, \tau) \cdot (-cZ) + O(Z^2),\end{aligned}$$

in the last step, using the Taylor expansion of the bivariate function $G(z, y)$. Hence by coefficient identification, we have $\tau' = G(\rho, \tau)$ and more importantly $c' = c \cdot G_y(\rho, \tau)$. To show that $c' = 0$, we start with

$$G_y(z, y) = 1 - y - \frac{z}{1-z},$$

to show that this cancels out at (ρ, τ).

Recall that at $(z, y) = (\rho, \tau)$ the equations $y = F(z, y)$ and $1 = F_y(z, y)$ are satisfied—this is consequence of Eq. (4.36). These equations read

$$\begin{aligned}y &= \frac{z}{1-z} \cdot \Big(\exp\big(y + B(z)\big) - 1\Big),\\ 1 &= \frac{z}{1-z} \cdot \exp\big(y + B(z)\big).\end{aligned}$$

Subtracting the second equation from the first equation we get $y-1 = z/(1-z)$, which, due to our simplification of $\mathcal{L}$ at the beginning of this section, is equivalent to $y - 1 = -L(z)$, itself also equivalent to $0 = 1 - y - L(z) = G_y(z, y)$.

Since this equation holds at (ρ, τ), we conclude that $G_y(\rho, \tau) = 0$, to the effect that $c' = 0$.

We now need only verify that the leading singular term of $U(z)$ is indeed Z^3 (*i.e.*, we want to make sure that $e' \neq 0$). Were that not be the case, we would have $U(z) = \tau + d' \cdot Z^2 + O\big(Z^4\big)$, and using previously mentioned transfer theorems, this would imply that $[z^n]U(z) = o(\rho^{-n} \cdot n^{-5/2})$.

Note that an unrooted 3LP graph γ with n vertices gives rise to not more than n objects in $\mathcal{S}_X$ (precisely it gives rise to $n(\gamma)$ objects in $\mathcal{S}_X$, where $n(\gamma)$ is the number of dissimilar vertices of γ that are adjacent to a star-leaf in the split-decomposition tree), hence $n \cdot [z^n]U(z) \geq [z^{n-1}]S_X(z)$. Since $[z^{n-1}]S_X(z)$ is $\Theta(\rho^{-n}n^{-3/2})$ we conclude that $[z^n]U(z) = \Omega(\rho^{-n}n^{-5/2})$, and thus $e' \neq 0$. Using transfer theorems we conclude that $[z^n]U(z) \sim \frac{3\,e'}{4\sqrt{\pi}}\rho^{-n}n^{-5/2}$.

Leading constant. Let us now briefly explain how to compute the constant e', and more generally, how to accurately estimate the coefficients $u_0, u_1, \ldots, u_k$ in the singular expansion of $U(z)$ to any fixed order k:

$$U(z) = u_0 + u_1 Z + \cdots + u_k Z^k + o_{z\to\rho}(Z^k),$$

with $Z = \sqrt{1 - z/\rho}$. The first step is to estimate the coefficients in the singular expansion of $S_X(z)$, of the form

$$S_X(z) = c_0 + c_1 Z + \cdots + c_k Z^k + o_{z\to\rho}(Z^k),$$

For any fixed m (with the notations $F^{[m]}(z,y), \rho^{[m]}, \tau^{[m]}$ introduced above), we let $y^{[m]} := \tau^{[m]} + c_1^{[m]}Z + c_2^{[m]}Z^2 + c_3^{[m]}Z^3 + \cdots + c_k^{[m]}Z^k$, and consider the equation

$$-y^{[m]} + F^{[m]}(\rho^{[m]} \cdot (1 - Z^2), y^{[m]}) = 0,$$

which we expand order by order in Z, each coefficient $[Z^i]$ in $H := -y^{[m]} + F^{[m]}(\rho^{[m]} \cdot (1 - Z^2), y^{[m]})$ being a certain polynomial expression in $c_1^{[m]}, \ldots, c_k^{[m]}$. As it turns out, the coefficient $[Z^0]H$ and $[Z^1]H$ are 0, the coefficient $[Z^2]H$ is of the form $\frac{1}{2}(c_1^{[m]})^2 - a$ with $a \approx 1.46797$, which gives $c_1^{[m]} = -\sqrt{2a} \approx -1.71346$, and then for $3 \leq i \leq k+1$ the coefficient $[Z^i]H$ is of the form $c_1^{[m]}c_{i-1}^{[m]} - P_i(c_1^{[m]}, \ldots, c_{i-2}^{[m]})$ for a certain explicit polynomial P_i. This allows us to solve iteratively for the constants $c_2^{[m]}, c_3^{[m]}, \ldots$; we find $c_2^{[m]} \approx 1.45297$, $c_3^{[m]} \approx -0.33156$, etc, and we observe exponentially fast convergence as m increases.

Then, to obtain the coefficients u_i, we simply use the explicit expression $U(z) = G(z, S_X(z))$, which ensures that the singular expansion of $U(z)$ is the same as the singular expansion of $S_X(z) \cdot \big(1 - \frac{z}{1-z} - \frac{1}{2}S_X(z)\big)$. Expanding order by order in Z, we find that each u_i is a polynomial expression in $c_1, \ldots, c_i$, which allows us to compute the u_i's from the c_i's, giving $e' = u_3 \approx 1.67688$.

4.2 Distance-hereditary graphs. Since the following analysis is very similar in execution to that of the 3-leaf power graphs in the previous subsection, we will omit details.

We begin, as before, with the rooted case. Remark 5 notes that $\mathcal{K}$ and $\mathcal{S}_C$ play symmetric roles and hence have the same generating function, which we denote by $K(z)$. We denote by $S(z)$ the generating function of $\mathcal{S}_X$. Then from Eq. (3.24), we obtain

$$S(z) = \frac{(z + 2K(z))^2}{1 - z - 2K(z)}.$$

If we define $A(z) = z + K(z) + S(z) = z + K(z) + \frac{(z+2K(z))^2}{1-z-2K(z)}$ then Eq. (3.23) yields

$$K(z) = \exp\Big(\sum_{i\geq 1} \frac{1}{i} A(z^i)\Big) - 1 - A(z),$$

so that $y = K(z)$ satisfies the functional equation

$$y = \exp_{\geq 2}\Big(z + y + \frac{(z+2y)^2}{1 - z - 2y}\Big)\exp(B(z)),$$

with the notation $\exp_{\geq d}(t) = \sum_{i\geq d} \frac{t^i}{i!}$, and with $B(z) := \sum_{i\geq 2} \frac{1}{i} A(z^i)$.

As in Section 4.1, this is an equation of the form $y = F(z,y)$, with $F(z,y)$ a power series with non-negative coefficients and with nonlinear dependency on y. Thus, if we denote by ρ the radius of convergence of $K(z)$, then $K(z)$ converges to a finite positive constant (denoted by τ) when z tends to ρ from below; and since $y = K(z)$ does not diverge when z tends to ρ, then we must have $1 - z - 2y > 0$ at $(z,y) = (\rho,\tau)$ (no cancellation of the denominator appearing inside the exponential). Moreover, since the number of distance-hereditary graphs grows exponentially with n, we must have $\rho < 1$, from which we easily deduce that $B(z)$ is analytic at ρ, and that $F(z,y)$ is analytic at (ρ,τ). Hence, as in Section 4.1, the Drmota-Lalley-Woods theorem ensures that $K(z)$ has a singular expansion of the form

$$K(z) = \tau - c \cdot Z + O(Z^2), \ \text{ with } Z = \sqrt{1 - z/\rho},$$

and has an analytic continuation in a Δ-domain. Hence we can apply transfer theorems to obtain the asymptotic estimate $[z^n]K(z) \sim \frac{c}{\sqrt{\pi}}\rho^{-n}n^{-3/2}$. Again we can use an iterated scheme to evaluate the constants with increasing precision, we obtain $\rho \approx 7.249751250\ldots$.

For the unrooted case, similarly as for 3LP graphs, we express the generating function $U(z)$ of unrooted DH graphs in terms of $K(z)$, and verify that the leading singular term is of order Z^3. Again we have to use the fact that $K(z)$ admits a singular expansion up to terms of order Z^3, of the form

$$K(z) = \tau - cZ + dZ^2 + eZ^3 + O(Z^4).$$

Eq. (3.25) of Theorem 4 yields

$$U(z) = T_K(z) + T_S(z) + T_{S-S}(z) - T_{K-S}(z) - T_{S\to S}(z).$$

To express $T_K(z) + T_{S-S}(z)$ in terms of $K(z)$ we observe that

$$\begin{aligned}\mathcal{K} &= \textsc{Set}_{\geq 2}(\mathcal{Z}+\mathcal{K}+\mathcal{S})\\ &= \textsc{Set}_{\geq 3}(\mathcal{Z}+\mathcal{K}+\mathcal{S}) + \textsc{Set}_2(\mathcal{Z}+\mathcal{K}+\mathcal{S})\\ &= \mathcal{T}_K + \textsc{Set}_2(\mathcal{K}) + \textsc{Set}_2(\mathcal{S}) + \textsc{Set}_2(\mathcal{Z})\\ &\quad + \mathcal{Z}\times\mathcal{K} + \mathcal{Z}\times\mathcal{S} + \mathcal{K}\times\mathcal{S}\\ &= \mathcal{T}_K + \mathcal{T}_{S-S} + \textsc{Set}_2(\mathcal{Z}) + \mathcal{Z}\times\mathcal{K} + \mathcal{Z}\times\mathcal{S} + \mathcal{K}\times\mathcal{S}.\end{aligned}$$

Hence

$$T_K(z) + T_{S-S}(z) = K(z) - zK(z) - zS(z) - K(z)S(z).$$

Next, we have

$$T_{K-S}(z) = K(z)(K(z) + S(z)),$$

and

$$T_{S\to S}(z) = K(z)^2 + S(z)^2.$$

Finally, using $S(z) = \frac{(z+2K(z))^2}{1-z-2K(z)}$, we find $U(z) = G(z, K(z))$, where

$$G(z,y) := y - z^2 - \frac{(z+2y)^3}{(1-z-2y)^2}.$$

Remarkably, $U(z)$ admits here a rational expression in terms of z and $K(z)$, which was not the case for 3LP graphs (recall that the expression of $U(z)$ involved a term $S_X(z^2)$).

Similarly as for 3LP graphs, we note that $G(z,y)$ is analytic at (ρ,τ), so that $U(z)$ admits a singular expansion of the form

$$U(z) = \tau - c'Z + d'Z^2 + e'Z^3 + O(Z^4),$$

with the relation $c' = cG_y(\rho,\tau)$. We have

$$G_y(z,y) = \frac{(1+z+2y)(4y^2+4zy+z^2-8y-4z+1)}{(1-z-2y)^3}.$$

In order to verify that this cancels out at (ρ,τ), we again use the fact that at (ρ,τ), both equations $y = F(z,y)$ and $1 = F_y(z,y)$ are satisfied. Defining $R(z,y) = z + y + \frac{(z+2y)^2}{1-z-2y}$, these equations read

$$\begin{aligned}y &= \exp(R(z,y) + B(z)) - 1 - R(z,y),\\ 1 &= R_y(z,y)\exp(R(z,y)+B(z)) - R_y(z,y).\end{aligned}$$

Multiplying the first one by $R_y(z,y)$ and then subtracting the second one (so as to eliminate $\exp(R(z,y)+B(z))$), we obtain the following equation, which is satisfied at (ρ,τ):

$$0 = \frac{4y^2+4zy+z^2-8y-4z+1}{(1-z-2y)^3}.$$

We recognize the numerator as a factor in the numerator of $G_y(z,y)$, from which we conclude that $G_y(\rho,\tau) = 0$, and thus $c' = 0$. Similarly as for 3LP graphs, the fact that $[z^n]K(z) = \Theta(\rho^{-n}n^{-3/2})$ and $[z^n]U(z) \geq \frac{1}{n}[z^{n-1}]K(z)$ ensures that $e' \neq 0$, and $[z^n]U(z) \sim \frac{3\,e'}{4\sqrt{\pi}}\rho^{-n}n^{-5/2}$.

5 Exhaustive Enumeration

Since most of the classes enumerated in this paper, in their various flavors (labeled/unlabeled, rooted/unrooted, connected/disconnected), had no known enumeration, it became useful to have some reference enumerations to confirm the correctness of the grammars we deduced.

To this end, we have used the vertex incremental characterization of the studied classes of graphs. These are surprisingly readily available in the graph literature, and provide a convenient—and thankfully rather foolproof—way of finding reliable enumeration and exhaustive generation of these classes of graphs: by brute force, you grow a set of graphs using the vertex incremental operations, and eliminating isomorphic graphs as they are created. This is still exponential, but allows for the exhaustive generation of all graphs up to size $n = 11, 12\ldots$ vertices, see Figure 4, which is roughly enough to check whether our grammars are accurate at two depths of induction.

6 Conclusion

In this paper, we have taken well-known characterization results by established graph researchers [20], and have turned these characterizations into grammars, enumerations and asymptotics—for two classes of graphs for which these were previously unknown.

This illustrates that a tool long known by graph theorists is a very fruitful line of research in analytic combinatorics, of which this paper is likely only the beginning.

Future questions in this same line may focus, for instance, on the parameter analysis. For instance, Iriza [23, §7] has already empirically noted, that in the split-decomposition tree of an unrooted, unlabeled distance-hereditary graph, the number of clique-nodes grows approximately as $\sim 0.221n$ and the number of star nodes grows approximately as $\sim 0.593n$. This offers some intuition as to what is a typical "shape" for a distance-hereditary graph: many nodes concentrated in a small number of cliques and then long filaments in between as in Figure 2. But a more qualitative investigation is required.

Iriza also brings to light an issue with our methodology. While the dissymmetry theorem solves many issues that have frustrated many combinatoricians (the symmetries when enumerating unrooted trees), it does not provide a symbolic grammar for the unrooted graph classes. This prevents us from efficiently randomly generating graphs [23, §3.2]. An interesting line of inquiry would be to refine the application of cycle-pointing so that it is as straightforward as that of the dissymmetry theorem.

Another promising avenue is to investigate whether more complicated classes of graphs can easily be enumerated. Any superset of the distance-hereditary graphs (which are the totally decomposable graphs for the split-decomposition) will necessarily involve the presence of *prime nodes* (internal graph labels which are neither star graphs nor clique graphs). For instance, Shi [27] has done an experimental study of parity graphs (which have bipartite graphs as prime nodes).

Acknowledgments

We are grateful to the anonymous reviewers for their careful reading of our manuscript, and thoughtful corrections.

This work was begun in 2008, when the second author was visiting the first at Simon Fraser University; it was pursued in 2013-2014 during the third author's post-doc at the same university; the results in this paper were presented [8] at the Seventh International Conference on Graph Transformation (ICGT 2014).

Between then and now, it has improved from the careful remarks and the work of several of our students: Alex Iriza [23], who provided many of the figures, Jessica Shi, and Maryam Bahrani [1].

All of the figures (with the exception of Figure 2) in this article were created in OmniGraffle 6 Pro.

References

[1] Maryam Bahrani and Jérémie Lumbroso. Enumerations, Forbidden Subgraph Characterizations, and the Split-Decomposition. *arXiv preprint arXiv:1608.01465*, 2016.

[2] François Bergeron, Gilbert Labelle, and Pierre Leroux. *Combinatorial Species and Tree-like Structures*, volume 67 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, Cambridge, 1998.

[3] Manuel Bodirsky, Éric Fusy, Mihyun Kang, and Stefan Vigerske. Enumeration and asymptotic properties of unlabeled outerplanar graphs. *Electronic Journal of Combinatorics*, 14(R66):1–24, 2007.

[4] Manuel Bodirsky, Éric Fusy, Mihyun Kang, and Stefan Vigerske. Boltzmann samplers, Pólya theory, and cycle pointing. *SIAM Journal on Computing*, 40(3):721–769, 2011.

[5] Binh-Minh Bui-Xuan. *Tree-representation of set families in graph decompositions and efficient algorithms*. PhD thesis, Université d'Orléans, 2008.

[6] Guillaume Chapuy, Éric Fusy, Mihyun Kang, and Bilyana Shoilekova. A complete grammar for decomposing a family of graphs into 3-connected components. *The Electronic Journal of Combinatorics*, 15, 2008.

[7] Pierre Charbit, Fabien de Montgolfier, and Mathieu Raffinot. Linear Time Split Decomposition Revisited. *SIAM Journal on Discrete Mathematics*, 26(2):499–514, 2012.

[8] Cédric Chauve, Éric Fusy, and Jérémie Lumbroso. An Enumeration of Distance-Hereditary and 3-Leaf Power Graphs. Technical report, 2014.

[9] William H. Cunningham. Decomposition of Directed Graphs. *SIAM Journal on Algebraic Discrete Methods*, 3(2):214–228, 1982.

[10] William H. Cunningham and Jack Edmonds. A combinatorial decomposition theory. *Canadian Journal of Mathematics*, 32(3):734–765, 1980.

[11] Michael Drmota. Systems of Functional Equations. *Random Structures & Algorithms*, 10(1-2):103–124, 1997.

[12] Michael Drmota. Trees. In Miklos Bona, editor, *Handbook of Enumerative Combinatorics*, pages 281–334. CRC Press, 2015.

[13] Michael Drmota, Éric Fusy, Mihyun Kang, Veronika Kraus, and Juanjo Rué. Asymptotic Study of Subcritical Graph Classes. *SIAM Journal on Discrete Mathematics*, 25(4):1615–1651, 2011.

[14] Philippe Duchon, Philippe Flajolet, Guy Louchard, and Gilles Schaeffer. Boltzmann Samplers for the Random Generation of Combinatorial Structures. *Combinatorics, Probability and Computing*, 13(4–5):577–625, 2004.

[15] Philippe Flajolet, Éric Fusy, and Carine Pivoteau. Boltzmann Sampling of Unlabelled Structures. In David Applegate et al, editor, *Proceedings of the Ninth Workshop on Algorithm Engineering and Experiments and the Fourth Workshop on Analytic Algorithmics and Combinatorics*, pages 201–211. SIAM Press, 2007.

[16] Philippe Flajolet and Bruno Salvy. Computer Algebra Libraries for Combinatorial Structures. *Journal of Symbolic Computation*, 20(5):653–671, 1995.

[17] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.

[18] Philippe Flajolet, Paul Zimmermann, and Bernard van Cutsem. A Calculus for the Random Generation of Labelled Combinatorial Structures. *Theoretical Computer Science*, 132(1-2):1–35, 1994.

[19] Omer Giménez and Marc Noy. Counting planar graphs and related families of graphs. *Surveys in Combinatorics*, 365:169–210, 2009.

[20] Emeric Gioan and Christophe Paul. Split decomposition and graph-labelled trees: Characterizations and fully dynamic algorithms for totally decomposable graphs. *Discrete Applied Mathematics*, 160(6):708–733, 2012.

[21] Emeric Gioan, Christophe Paul, Marc Tedder, and Derek G. Corneil. Practical and Efficient Split Decomposition via Graph-Labelled Trees. *Algorithmica*, pages 1–55, 2013.

[22] Peter L. Hammer and Frédéric Maffray. Completely Separable Graphs. *Discrete Applied Mathematics*, 27(1):85–99, 1990.

[23] Alexander Iriza. Enumeration and random generation of unlabeled classes of graphs: A practical study of cycle pointing and the dissymmetry theorem. Master's thesis, Princeton University, arXiv preprint arXiv:1511.06037, 2015.

[24] Shin-ichi Nakano, Ryuhei Uehara, and Takeaki Uno. A New Approach to Graph Recognition and Applications to Distance-Hereditary Graphs. *Journal of Computer Science and Technology*, 24(3):517–533, 2009.

[25] Naomi Nishimura, Prabhakar Ragde, and Dimitrios M. Thilikos. On Graph Powers for Leaf-Labeled Trees. *Journal of Algorithms*, 42(1):69–108, 2002.

[26] Vlady Ravelomanana and Loÿs Thimonier. Asymptotic enumeration of cographs. *Electronic Notes in Discrete Mathematics*, 7(0):58–61, 2001.

[27] Jessica Shi. Enumeration of unlabeled graph classes: A study of tree decompositions and related approaches. 2015.

[28] Jeremy P. Spinrad. *Efficient Graph Representations*, vo-

lume 19 of *The Fields Institute for Research in Mathematical Sciences*. American Mathematical Society, 2003.

[29] Paul Zimmermann. *Séries génératrices et analyse automatique d'algorithmes*. PhD thesis, École Polytechnique, 1991.

A Distance-Hereditary Grammar Simplification

The class $\mathcal{DH}_\bullet$ of distance-hereditary graphs rooted at a vertex is originally specified by

$$\begin{aligned}\mathcal{DH}_\bullet &= \mathcal{Z}_\bullet \times (\mathcal{K} + \mathcal{S}_C + \mathcal{S}_X)\\ \mathcal{K} &= \textsc{Set}_{\geqslant 2}(\mathcal{Z} + \mathcal{S}_C + \mathcal{S}_X)\\ \mathcal{S}_C &= \textsc{Set}_{\geqslant 2}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X)\\ \mathcal{S}_X &= (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C) \times \textsc{Set}_{\geqslant 1}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X).\end{aligned}$$

The point of this appendix is to prove that the last equation can be simplified to

$$\mathcal{S}_X = \textsc{Seq}_{\geqslant 2}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C).$$

Although we first provide a straightforward formal derivation, we then follow it up with an intuitive explanation.

Proof. Indeed, while the elements of a SET have symmetries that are hard to take into account, this is not the case for sets of size 1, therefore

$$\textsc{Set}_{\geqslant 1}(\mathcal{U}) = \mathcal{U} + \textsc{Set}_{\geqslant 2}(\mathcal{U}).$$

By combining this fact with the definition of $\mathcal{S}_C$,

$$\mathcal{S}_C = \textsc{Set}_{\geqslant 2}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X),$$

we have that (parentheses in the right hand side purely for emphasis)

$$\textsc{Set}_{\geqslant 1}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X) \equiv (\mathcal{Z} + \mathcal{K} + \mathcal{S}_X) + \mathcal{S}_C$$

hence,

$$\begin{aligned}\mathcal{S}_X &= (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C) \times (\mathcal{Z} + \mathcal{K} + \mathcal{S}_X + \mathcal{S}_C)\\ &= \mathcal{S}_X \times (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C) + (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)^2\end{aligned}$$

we then proceed to manipulate this specification purely symbolically, implying

$$\mathcal{S}_X\left[1 - (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)\right] = (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)^2$$

and thus

$$\begin{aligned}\mathcal{S}_X &= \frac{(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)^2}{1 - (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)}\\ &= (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)^2 \times \textsc{Seq}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)\\ &= \textsc{Seq}_{\geqslant 2}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C).\end{aligned}$$

Finally

$$\mathcal{S}_X = \textsc{Seq}_{\geqslant 2}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C).$$

□

Remark 7. To understand this simplification from a combinatorial perspective, imagine that we have a connected subsequence of star-nodes connected by their extremities.

Without loss of generality, we can assume that all *but the last* of these internal star-nodes have only two extremities[17]—the one through which they are entered, and another one. We are then either in the situation illustrated by Figure 3a (in which the last star-node of the subsequence only has one additional extremity) or by Figure 3b (in which the last star-node has several extremities).

This subsequence of adjacent star-nodes connected by their extremities, translates to the grammar by a recursive expansion of the $\mathcal{S}_X$ rule: each of these has a $(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)$ child for the center of the star, and then one other children for the other extremity. This is repeated until we have reached the last adjacent star-node in the subsequence which can either have one or multiple extremities:

- If it has only one extremity, then this extremity connects to either a leaf or to a clique, thus $\mathcal{Z} + \mathcal{K}$ (Figure 3a).
- Otherwise, it has two or more undistinguished extremities, in which case we can *pretend* that this set of extremities is a $\mathcal{S}_C$ term (Figure 3b).

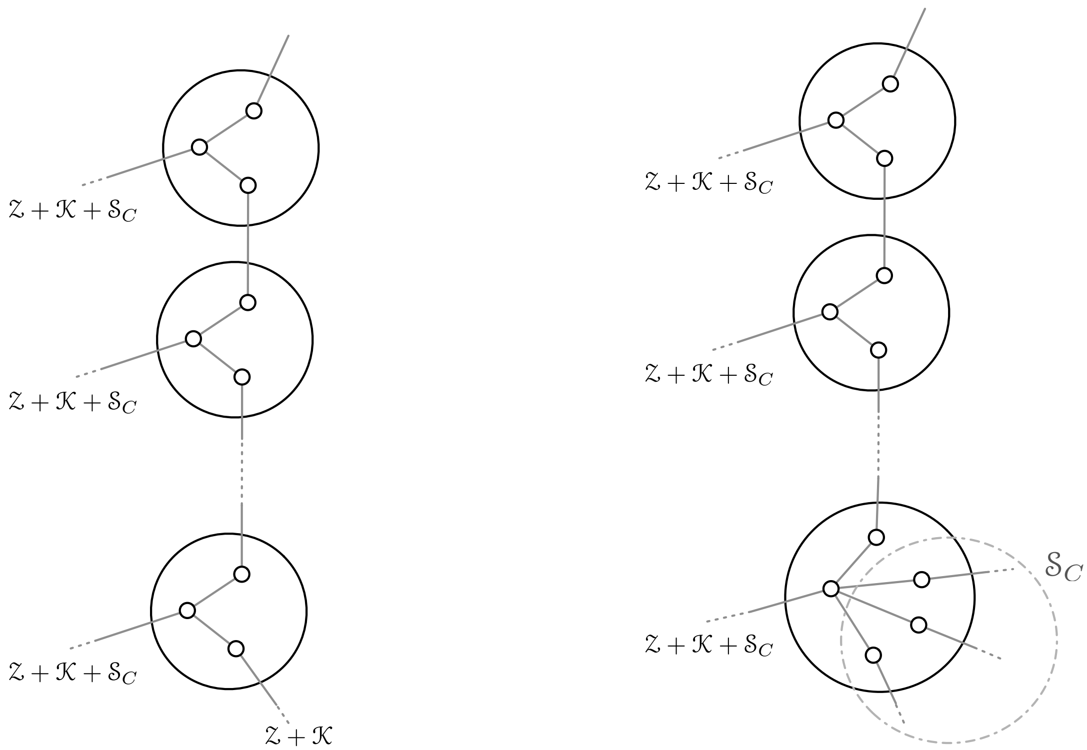
Recall that the original interpretation of $\mathcal{S}_X$,

$$\mathcal{S}_X = (\mathcal{Z} + \mathcal{K} + \mathcal{S}_C) \times \textsc{Set}_{\geqslant 1}(\mathcal{Z} + \mathcal{K} + \mathcal{S}_X),$$

is as follows: a distinguished center which can lead to either a leaf, a clique, or a star-node entered through its center; and a set of undistinguished extremities, each of which can lead to either a leaf, a clique, or another star-node entered through an extremity.

The new interpretation follows the figures: we have a sequence of $(\mathcal{Z} + \mathcal{K} + \mathcal{S}_C)$ terms for the center of each of the adjacent star-nodes (and we have at least one such star-node), and finally another such term to cover both possibilities, where the final star-node either has one extremity or several. This is equivalent to having a sequence of at least two of these terms, hence the simplified equation.

[17] The first star-node of the subsequence to have more than one extremity is the "last" star-node of that particular subsequence. In particular, it is possible for the subsequence to only have one single star-node.



(a) This case occurs when the last star in the subsequence of adjacent stars has *only one* extremity.

(b) This case occurs when the last star in the subsequence of adjacent stars has *at least two* extremities (here, it has three).

Figure 3. Combinatorial intuition behind the derivation of Appendix A.

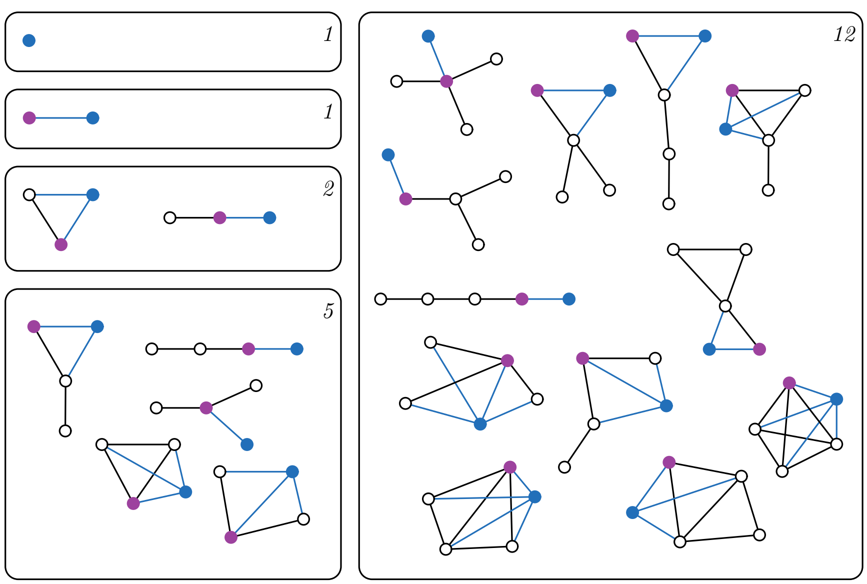


Figure 4. All unrooted, unlabeled 3-leaf power graphs of sizes 1 through 5, beginning the enumeration: 1, 1, 2, 5, 12, ... The coloring of the vertices illustrate one possible way to derive the graphs through vertex incremental operations sketched in Section 5: the newly added vertex is in blue, while the existing vertex it is added in reference to is in purple.