Random planar trees and the Jacobian conjecture

Elia Bisi, Piotr Dyszewski, Nina Gantert, Samuel G. G. Johnston, Joscha Prochno, Dominik Schmid

Angaben zur Veröffentlichung / Publication details:

Bisi, Elia, Piotr Dyszewski, Nina Gantert, Samuel G. G. Johnston, Joscha Prochno, and Dominik Schmid. 2026. "Random planar trees and the Jacobian conjecture." *Journal of the London Mathematical Society* 113 (1): e70416. https://doi.org/10.1112/jlms.70416.

Received: 29 March 2024 | Revised: 1 August 2025 | Accepted: 1 December 2025

DOI: 10.1112/jlms.70416

RESEARCH ARTICLE

Journal of the London Mathematical Society

Random planar trees and the Jacobian conjecture

Elia Bisi[1] | **Piotr Dyszewski**[2] | **Nina Gantert**[3] | **Samuel G. G. Johnston**[4] | **Joscha Prochno**[5] | **Dominik Schmid**[6]

[1]Department of Mathematics and Computer Science, University of Florence, Florence, Italy

[2]Mathematical Institute, Wrocław University, Wroclaw, Poland

[3]Department of Mathematics, Technical University of Munich, Garching, Germany

[4]Department of Mathematics, King's College London, London, UK

[5]Faculty of Computer Science and Mathematics, University of Passau, Passau, Germany

[6]Department of Mathematics, University of Augsburg, Augsburg, Germany

Correspondence

Nina Gantert, Department of Mathematics, Technical University of Munich, Boltzmannstraße 3, D-85748 Garching, Germany.
Email: nina.gantert@tum.de

Funding information

National Science Centre, Poland, Grant/Award Number: 2020/39/D/ST1/00258; EPSRC, Grant/Award Number: EP/S036202/1; Austrian Science Fund, Grant/Award Number: P32405

Abstract

We develop a probabilistic approach to the celebrated Jacobian conjecture, which states that any Keller map (i.e. any polynomial mapping $F: \mathbb{C}^n \to \mathbb{C}^n$ whose Jacobian determinant is a non-zero constant) has a compositional inverse which is also a polynomial. The Jacobian conjecture may be formulated in terms of a problem involving labellings of rooted trees; we give a new probabilistic derivation of this formulation using multi-type branching processes. Thereafter, we develop a simple and novel approach to the Jacobian conjecture in terms of a problem involving shuffling subtrees of d-Catalan trees, that is, planar d-ary trees. We also show that, if one can construct a certain Markov chain on large d-Catalan trees which updates its value by randomly shuffling certain nearby subtrees, and in such a way that the stationary distribution of this chain is uniform, then the Jacobian conjecture is true. Finally, we use the local limit theory of large random trees to show that the subtree shuffling conjecture is true in a certain asymptotic sense, and thereafter use our machinery to prove an approximate version of the Jacobian conjecture, stating that inverses of Keller maps have small power series coefficients for their high-degree terms.

MSC 2020
05C05, 14R15, 60J80 (primary), 13F25, 60J10 (secondary)

1 | INTRODUCTION AND OVERVIEW

1.1 | The Jacobian conjecture

The Jacobian conjecture, a problem from algebraic geometry which is concerned with the invertibility of polynomial mappings in n variables over a field of characteristic zero, is one of the outstanding open problems in all of mathematics. It was first conjectured by Keller in 1939 [18] and has remained unsolved for over 80 years; even the case $n = 2$ is still open. It features as Problem 16 on Smale's list of *18 mathematical problems for the 21st century* [25] and a resolution seems to require completely novel ideas and methods.

We now give a brief account of the Jacobian conjecture, assuming no specialist knowledge of commutative algebra or algebraic geometry. For the sake of simplicity we work over $\mathbb{C}$, though this formulation is equivalent to more general formulations over fields of characteristic zero [29]. Consider the ring $\mathbb{C}[X_1, \dots, X_n]$ of polynomials in n variables with coefficients in $\mathbb{C}$. A **polynomial mapping** $F = (F_1, \dots, F_n)$ is a function $F : \mathbb{C}^n \to \mathbb{C}^n$ such that each component F_i is an element of $\mathbb{C}[X_1, \dots, X_n]$. A polynomial mapping F is said to be a **polynomial automorphism** if there exists a second polynomial mapping G such that $F \circ G = G \circ F = I$, where I is the identity mapping on $\mathbb{C}^n$. We define the **Jacobian matrix** of F, denoted JF, to be the $n \times n$ matrix whose (i, j)th entry is $(JF)_{i,j} := \frac{\partial}{\partial X_j} F_i$. We denote by $\det JF$ the **Jacobian determinant** of F, which is itself an element of $\mathbb{C}[X_1, \dots, X_n]$.

We say a polynomial mapping is a **Keller mapping** if its Jacobian determinant is equal to a non-zero constant in $\mathbb{C}$. It is easily verified that every polynomial automorphism is a Keller mapping. The celebrated Jacobian conjecture is the converse of this statement.

Conjecture 1.1 (The Jacobian conjecture). *Every Keller mapping is a polynomial automorphism.*

There is a vast literature on the conjecture and its history. We refer the reader to van den Essen's monograph [29] for more information.

1.2 | The combinatorial approach to the Jacobian conjecture

While the Jacobian problem most commonly lies in the remit of algebraic geometry, over the years several authors have noted that the problem may be attacked using combinatorial arguments [1, 17, 23, 34, 37–39]. The combinatorial formulation of the Jacobian conjecture was crystallised in recent work by the fourth and fifth authors [17], and we now give a brief outline of this very concrete approach, which rests on the following three observations:

- We begin by appealing to the celebrated Bass–Connell–Wright reduction of the Jacobian conjecture [8]: to prove the Jacobian conjecture, it is sufficient to fix $d \geqslant 3$ and establish, for all n, the polynomial invertibility of every map of the form $F = I - H : \mathbb{C}^n \to \mathbb{C}^n$, where H is **homogeneous** of degree d (i.e. each of its components is homogeneous of degree d) and the Jacobian matrix JH is **nilpotent** (i.e. $(JH)^n = 0$). In fact, for such mappings, the condition that JH is nilpotent is equivalent to Keller's condition that $\det(JF) = 1$ (see [17, Section 2.5]).

  For any such map, we write

$$F_i(X_1, \dots, X_n) := X_i - \sum_{|\alpha|=d} \frac{H_{i,\alpha}}{\alpha!} X^\alpha \tag{1.1}$$

  for the ith component of F, where the sum is taken over multi-indices $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n_{\geqslant 0}$ with $|\alpha| := \alpha_1 + \cdots + \alpha_n = d$, and we set $\alpha! := \alpha_1! \cdots \alpha_n!$ and $X^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. Any such F (or H) is thus determined by the set of complex coefficients $H = (H_{i,\alpha} : i \in [n], |\alpha| = d)$.
- Whenever $F = I - H$ is a polynomial mapping, where H is homogeneous degree d (and JH may or may not be nilpotent), F has a power series inverse F^{-1} whose ith component takes the form

$$F_i^{-1}(X_1, \dots, X_n) = X_i + \sum_{k \geqslant 1} \sum_{|\alpha|=(d-1)k+1} g_{i,\alpha} X^\alpha.$$

  Moreover, for $|\alpha| = (d-1)k + 1$, the complex coefficients $g_{i,\alpha}$ may be written in terms of the formula [17, Theorem 2.3]

$$g_{i,\alpha} = \Phi_i^\alpha(H) := \frac{1}{(d!)^k} \sum_{T \in C_k^{(d)}} \sum_{(i,\alpha) \text{ labellings } \mathcal{T} \text{ of } T} \mathcal{E}_H(\mathcal{T}). \tag{1.2}$$

  Here, $C_k^{(d)}$ is the set of d-Catalan trees (i.e. planar, rooted, d-ary trees) with $(d-1)k + 1$ leaves (i.e. with k internal vertices). If $T = (V, E) \in C_k^{(d)}$, then a labelling $\mathcal{T}$ of T is a pair $\mathcal{T} = (T, \tau)$ where $\tau : V \to [n]$ is a function giving vertex v **type** $\tau(v)$. For $i \in \mathbb{N}$ and $\alpha \in \mathbb{Z}^n_{\geqslant 0}$, an (i, α) **labelling** of T is any labelling of T giving the root type i and giving exactly α_ℓ of the leaves type ℓ, for all $\ell \in [n]$. Finally, if for any internal vertex v, we let $\mu(v)$ be the multi-index whose jth component counts the number of type j children of v (so that $|\mu(v)| = d$), then the H-**weight** $\mathcal{E}_H(\mathcal{T})$ of the labelling $\mathcal{T}$ of T is given by

$$\mathcal{E}_H(\mathcal{T}) := \prod_{\substack{v \text{ internal} \\ \text{vertex of } \mathcal{T}}} H_{\tau(v),\mu(v)},$$

  where the product is taken over all internal vertices in V. See Figure 1.
- The nilpotency of JH imposes a condition on the coefficients of H, which we now describe. We define a d-Catalan **fern** to be an element of $C_n^{(d)}$ of height n, with a designated 'sink' vertex in generation n, and the property that at most one vertex in each generation is not a leaf (see Figure 2). Fix an arbitrary d-Catalan fern U. It then turns out that the nilpotency of JH is equivalent to the statement

$$\Psi_{i,j}^\alpha(H) = 0 \qquad \text{for every } i, j \in [n] \text{ and } \alpha \in \mathbb{Z}^n_{\geqslant 0} \text{ with } |\alpha| = (d-1)n, \tag{1.3}$$

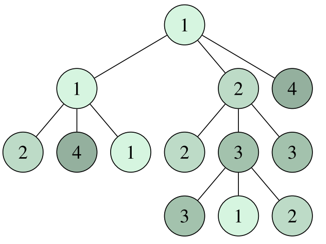

FIGURE 1 An (i,α) labelling $\mathcal{T}$ of a tree $T \in C_k^{(d)}$. In the figure, the degree is $d = 3$, the number of internal vertices is $k = 4$, the dimension is $n = 4$, the root type is $i = 1$ and the leaf type is $\alpha = (2,3,2,2)$. The H-weight of the depicted labelled tree $\mathcal{T}$ is $\mathcal{E}_H(\mathcal{T}) = H^2_{1,(1,1,0,1)}H_{2,(0,1,2,0)}H_{3,(1,1,1,0)}$. The coefficient $g_{i,\alpha}$ of X^α for the inverse of $F = I - H$ with H homogeneous degree d may be computed as a sum over (i,α) labellings of d-Catalan trees; see (1.2).

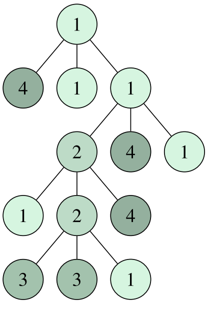

FIGURE 2 An (i,j,α) labelling of a 3-ary fern of height $n = 4$, with root type $i = 1$, sink type $j = 3$ and non-sink leave type $\alpha = (4,0,1,3)$. If JH is nilpotent with $(JH)^n = 0$, then weighted sums over (i,j,α) labellings of a fixed d-ary fern of length n are zero; see (1.3)–(1.4).

where

$$\Psi^\alpha_{i,j}(H) := \sum_{(i,j,\alpha)\text{ labellings } \mathcal{U} \text{ of } U} \mathcal{E}_H(\mathcal{U}). \tag{1.4}$$

Here, an (i,j,α) **labelling** of U is any labelling of U giving the root type i, the sink type j, and giving exactly α_ℓ of the non-sink leaves type ℓ, for all $\ell \in [n]$ (see [17, Lemma 2.8]). Clearly, definition (1.4) does not depend on the choice of the d-ary fern U.

We prove (1.2) and the equivalence of (1.3) with the nilpotency of JH in the sequel.

To recapitulate, if H is degree d homogeneous, then the coefficients of the inverse mapping of $F = I - H$ are given in (1.2) as a sum over weighted trees (see Figure 1). If we assume that JH is nilpotent, then the coefficients of H must satisfy the nilpotency condition (1.3)-(1.4) involving sums over labellings of a d-ary fern (see Figure 2).

Now, as mentioned above, the Jacobian conjecture is equivalent to the statement that every F of the form $F = I - H$ with H d-homogeneous and JH nilpotent has the property that F^{-1} is a polynomial mapping. In other words, for such F, it must be the case that the combinatorial quantity $\Phi^\alpha_i(H)$ defined in (1.2) for the coefficients of F^{-1} is zero for all but finitely many choices of α. Combining all the above observations, we arrive at the following combinatorial formulation of the Jacobian conjecture:

Conjecture 1.2 (Jacobian conjecture – combinatorial version). *There exists $d \geqslant 3$ such that the following is true for all $n \geqslant 1$. Suppose $H = (H_{i,\alpha} : i \in [n], |\alpha| = d)$ is a collection of complex numbers satisfying $\Psi^{\alpha}_{i,j}(H) = 0$ for all $i, j \in [n]$ and $\alpha \in \mathbb{Z}^n_{\geqslant 0}$ with $|\alpha| = (d-1)n$. Then, for each $i \in [n]$, $\Phi^{\alpha}_i(H) = 0$ for all but finitely many α.*

We stress that Conjecture 1.2 is equivalent to the Jacobian conjecture, Conjecture 1.1.

In this article, we give a rapid new derivation of the combinatorial formulation of the Jacobian conjecture [17] using a probabilistic approach centred around multi-type Galton–Watson trees. We then substantially reduce the complicated conjecture in [17] involving labelled trees to a new and far simpler conjecture involving unlabelled trees. We show it is also possible to reformulate the latter in terms of stationary distributions of Markov chains on trees that update by shuffling certain subtrees. Finally, we apply the local limit theory of large random trees to attack this simpler conjecture, showing it is true in a certain asymptotic sense, and relating this approximate form back to the original Jacobian conjecture. In short, probabilistic machinery has the potential to tackle the Jacobian conjecture, and thus the *raison d'etre* of the present article is to attract probabilists to the problem.

In Section 1.3, we give an outline of our main results, with the understanding that full details surrounding the definitions will be given in the sequel.

1.3 | Main results

The basic object we study are d-**Catalan trees**: these are rooted, planar, d-ary trees. Here and throughout, by d-**ary tree** we mean a tree in which each vertex has either 0 or d children. We write

$$C^{(d)}_k := \{d\text{-Catalan trees with } k \text{ internal vertices}\}.$$

Take an ancestral path $(v_0, \ldots, v_p)$ of length p in a tree $T \in C^{(d)}_k$, and consider the subtrees subtended by the $(d-1)p$ siblings of $v_1, \ldots, v_p$ (some of which may be leaves). The **shuffle class** associated with this ancestral path is the subset of $C^{(d)}_k$ consisting of trees $T' \in C^{(d)}_k$ that can be obtained from T by rearranging these $(d-1)p$ trees. We call any such subset a length-p shuffle class. See Figure 3 for an illustration and Section 3.1 for more precise definitions.

We make the following conjecture about the vector space of functions on $C^{(d)}_k$.

Conjecture 1.3 (Subtree shuffling conjecture). *Fix $d \geqslant 2$ and $p \geqslant 1$. Then, there exists an integer $k_0(d,p)$ such that, for all $k \geqslant k_0(d,p)$, the constant function $\mathbf{1}$ on $C^{(d)}_k$ lies in*

$$\operatorname{span}\{\mathbf{1}_S : S \text{ is a length-}p\text{ shuffle class in } C^{(d)}_k\},$$

that is, in the span of the indicator functions of the length-p shuffle classes.

Let us interpret this conjecture. Spelling out the statement, we are saying that to each length-p shuffle class $S \subseteq C^{(d)}_k$ we may associate a real number λ_S in such a way that for every tree $T \in C^{(d)}_k$ we have

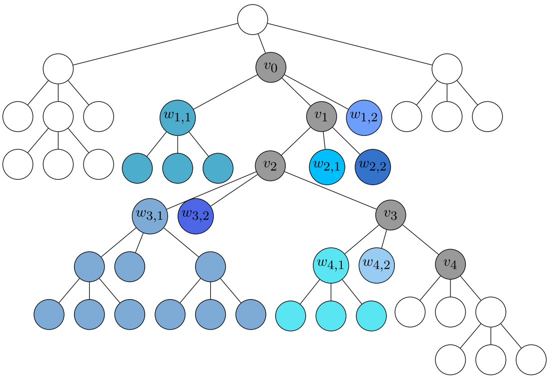

FIGURE 3 A d-Catalan tree with $d = 3$ and $k = 15$ internal vertices. An ancestral path $(v_0, \dots, v_4)$ of length $p = 4$ is highlighted in grey. Each vertex v_i, $i = 1, \dots, p$, has siblings $w_{i,j}$, $j = 1, \dots, d-1$. Thus, there are $(d-1)p = 8$ subtrees coming off the path, each of which is subtended by one of the $w_{i,j}$; these subtrees are highlighted in various shades of blue. Five of these trees are leaves; two of them (those subtended by $w_{1,1}$ and $w_{4,1}$) are the unique 3-Catalan tree with one internal vertex; finally, the one subtended by $w_{3,1}$ is a 3-Catalan tree with three internal vertices. The associated shuffle class consists of the set of $8!/(5!2!1!)$ trees that may be obtained by rearranging the positions of the subtrees subtended by the $w_{i,j}$ while keeping the positions of the white and grey vertices fixed. The notation of the v_i and $w_{i,j}$ corresponds to that of Section 3.1.

$$\sum_{\mathcal{S} \text{ length-}p \text{ shuffle class}} \lambda_{\mathcal{S}} \mathbf{1}_{\mathcal{S}}(T) = 1. \tag{1.5}$$

We can think of $\lambda_{\mathcal{S}}$ as a 'score' given to each tree T in $C_k^{(d)}$ for membership in $\mathcal{S}$. Conjecture 1.3 then states that, when the number of vertices is large, the shuffle classes form a sufficiently rich collection of subsets of $C_k^{(d)}$ so that there exists an allocation of scores to shuffle classes for which each tree gets the same total score.

The reason for our interest in Conjecture 1.3 is the following theorem, which we will prove in Section 4.1:

Theorem 1.4 (Subtree shuffling conjecture implies Jacobian conjecture). *If, for some $d \geqslant 3$, Conjecture 1.3 is true for all $p \in \mathbb{N}$, then so is the Jacobian conjecture.*

It is possible to adapt Conjecture 1.3 and the subsequent Theorem 1.4 to a problem involving tree-valued Markov chains, as we now outline. A ***p*-shuffle chain** is a Markov chain $(T_j)_{j \geqslant 0}$ taking values in $C_k^{(d)}$ that updates its value from T_j to T_{j+1} by choosing a length-p ancestral path within T_j according to some random T_j-dependent rule, and then creating T_{j+1} from T_j by choosing uniformly a rearrangement of the subtrees coming off this ancestral path. See Section 4.2 for the precise definition.

Conjecture 1.5 (Shuffle chain conjecture). *Fix $d \geqslant 2$ and $p \geqslant 1$. Then, there exists an integer $k_0(d, p)$ such that, for all $k \geqslant k_0(d, p)$, there exists a p-shuffle chain whose stationary distribution is the uniform distribution on $C_k^{(d)}$.*

We will show the following.

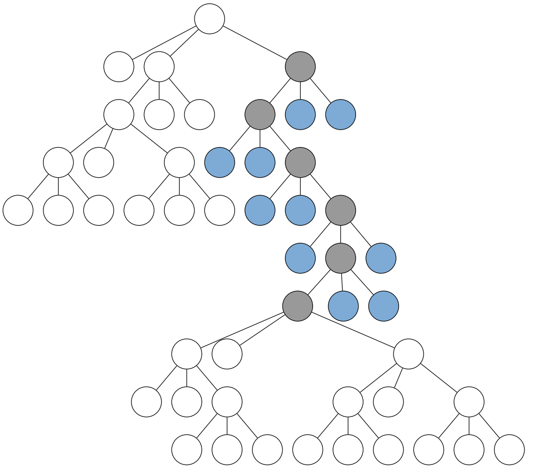

FIGURE 4 A p-perfect path of length $p = 5$ in a 3-Catalan tree. Every sibling off the path is a leaf.

Theorem 1.6 (Shuffle chain conjecture implies Jacobian conjecture). *If, for some $d \geqslant 3$, Conjecture 1.5 is true for all $p \geqslant 1$, then so is the Jacobian conjecture.*

We now turn to outlining some of our unconditional results. In this direction, we say a d-Catalan tree T is p-**perfect** if it contains a p-**perfect path**, that is, an ancestral path such that each sibling off the path is a leaf. See Figure 4.

In Section 5.2, we appeal to methods from the local limit theory of large random trees to show that, when the number of vertices is large, p-perfect trees are overwhelmingly common. Indeed, there is an underlying philosophy in the local behaviour of large random trees that distinct parts of the tree behave more or less independently of one another. With this picture in mind, the probability that a tree avoids having a p-perfect path somewhere decays exponentially in the number of vertices. This motivates the following result, which we will prove using a simple algorithm for generating uniform d-Catalan trees.

Theorem 1.7. *Let $p \geqslant 1$ and let $\mathbb{P}_k^{(d)}$ be the uniform measure on $C_k^{(d)}$. Then there exists a constant $\kappa_{p,d} > 0$ such that*

$$\mathbb{P}_k^{(d)}(T \textit{ is not } p\textit{-perfect}) \leqslant \mathrm{e}^{-\kappa_{p,d}(k-p)_+} \qquad \textit{for all } k \in \mathbb{N},$$

where $x_+ := \max(x, 0)$. In particular, we may set $\kappa_{p,d} := (2pd^p\mathrm{e}^p)^{-1}$.

If a tree is p-perfect, then rearranging the subtrees coming off its perfect path keeps the tree unchanged. Thus, p-perfect trees lie in a singleton shuffle class, and, in the setting of Conjecture 1.3, we can adjust the parameters λ_S so that (1.5) is true at least for all perfect trees, that is, for a fraction of at least $(1 - \mathrm{e}^{-\kappa_{p,d}(k-p)_+})$ of trees in $C_k^{(d)}$. In fact, we are also able to prove that, up to an error of the order $\mathrm{e}^{-O(\log^2 k)}$, Equation (1.5) is true for *all* trees in $C_k^{(d)}$. Let us summarise these facts in the following approximate version of Conjecture 1.3, which states that the constant function can be well approximated, in multiple ways, by a function in the span of shuffle indicators.

Theorem 1.8 (Approximating the constant function). *Let $p \in \mathbb{N}$ and $d \geqslant 2$. For $k \in \mathbb{N}$ sufficiently large, there exists a function ϕ_k in* $\mathrm{span}\{\mathbf{1}_S : S$ *is a length-p shuffle class in* $C_k^{(d)}\}$ *that well approximates the constant function* $\mathbf{1}$, *in the sense that the following bounds hold:*

- *(ℓ^1 bound) the exponential total variation bound*

$$||\phi_k - \mathbf{1}||_1 := \frac{1}{\#C_k^{(d)}} \sum_{T \in C_k^{(d)}} |\phi_k(T) - 1| \leqslant C\mathrm{e}^{-ck}; \tag{1.6}$$

- *(ℓ^∞ bound) the superpolynomial uniform bound*

$$||\phi_k - \mathbf{1}||_\infty := \sup_{T \in C_k^{(d)}} |\phi_k(T) - 1| \leqslant C'\mathrm{e}^{-c'(\log k)^2}. \tag{1.7}$$

Here, $C, c, C', c' \in (0, \infty)$ are constants that depend on the degree d of the trees and on the length p of the shuffle classes, but not on the number k of internal vertices in the trees.

Notice that the 1-norm of (1.6) can be also interpreted in terms of the uniform measure $\mathbb{P}_k^{(d)}$ on $C_k^{(d)}$ defined in Theorem 1.7. Indeed, if we denote by $\mathbb{E}_k^{(d)}$ the expectation with respect to $\mathbb{P}_k^{(d)}$, then $||f||_1 = \mathbb{E}_k^{(d)}|f(T)|$ for any function $f : C_k^{(d)} \to \mathbb{R}$.

We finally tie this result back to the Jacobian conjecture: we leverage Theorem 1.8 to provide quantitative information about the inverses of Keller maps, concluding that these have small power series coefficients for their high-degree terms. The precise statement of this approximate version of the Jacobian conjecture can be found in Theorem 5.4.

1.4 | Relevant literature

In this section, we touch on some related work.

To attack the Jacobian conjecture, numerous authors have studied the polynomial invertibility of Keller mappings of various degrees in various dimensions. Notably, Wang [30] proved in 1982 that the Jacobian conjecture is true in the special quadratic case:

Theorem 1.9 (Wang's Theorem [30]). *Every quadratic Keller mapping is a polynomial automorphism.*

See [8] for a simple proof. Wang's theorem contrasts strikingly with the degree $d \geqslant 3$ reduction of Bass, Connell and Wright, which states that, in order to prove the Jacobian conjecture, it is sufficient to study degree d Keller mappings for a fixed $d \geqslant 3$. See [9] for a reduction to the symmetric case.

Notably, Wang's theorem implies that the analogue of the combinatorial formulation of the Jacobian conjecture (Conjecture 1.2), *but with degree $d = 2$* rather than $d \geqslant 3$, is true! To this day, there is no known direct combinatorial derivation of this result. The authors are hopeful that a combinatorial proof of Wang's theorem may lead to information on how to prove (or disprove) the more general Jacobian conjecture by combinatorial means.

Combinatorial approaches to the Jacobian conjecture in the vein of Section 1.2 date back at least as far as the work of Bass, Connell and Wright [8]. Wright has continued work on the conjecture from the perspective of inversion formulas in [32–35]. The explicit formulation involving weighted trees given in Conjecture 1.2 was given explicitly in [17], but the idea is present in the aforementioned work of Wright, as well as in Singer [23, 24] and Zeilberger [37]. Let us also mention Abdesselam [1], whose combinatorial approach draws on quantum field theory.

Singer [23, 24] first utilised 'deweighting' techniques to convert the combinatorial formulation of the Jacobian problem involving weighted trees to one involving unweighted trees. In [23, 24], Singer was mainly concerned with studying the classical (i.e. 2-ary) Catalan trees to obtain a combinatorial derivation of Wang's theorem involving shuffle classes of 2-Catalan trees. As we will explain in Section 4.1, our methods and techniques draw inspiration from Singer's approach.

Johnston [16] recently used combinatorial methods to study the so-called strongly nilpotent automorphisms, that is, those of the form $F = I - H$, where

$$JH(X^{(1)}) \cdots JH(X^{(q)}) = 0 \qquad \forall X^{(1)}, \dots, X^{(q)} \in \mathbb{C}^n.$$

The above property is called strong nilpotency of index q, as it is a strengthening of the nilpotency condition $(JH)^q = 0$. Using combinatorial methods, [16] showed that, if $F = I - H$ has degree d and is strongly nilpotent of index q, then the inverse F^{-1} of F has degree at most d^{q-1}.

Finally, let us mention a conceptually related work of Le Jan and Sznitman [19], who used probabilistic expansions involving trees to study the three-dimensional Navier–Stokes equation.

1.5 | Outline of the article

The remainder of this article is structured as follows:

- In Section 2, we use arguments from branching processes to give a rapid proof of a combinatorial inversion formula for multivariate power series, which is in fact new in this generality.
- In Section 3, we show that the Keller property has a combinatorial interpretation in terms of shuffle classes of both labelled and unlabelled trees.
- In Section 4, we prove Theorems 1.4 and 1.6.
- In Section 5, we study perfect trees, proving Theorem 1.7. This leads to the ℓ^1 bound in Theorem 1.8 and to an approximate version of the Jacobian conjecture (Theorem 5.4).
- In Section 6, we construct the ℓ^∞ approximation in Theorem 1.8 and conclude the proof of the theorem.

Sections 2 and 3 contain new probabilistic proofs and insights, but are somehow preliminary to the rest of the work; they can be skipped on a first reading. Sections 4, 5 and 6 are the core of this article.

2 | COMBINATORIAL INVERSION VIA GALTON–WATSON TREES

In this section, we study combinatorial inversion of power series. The inversion formula we obtain is slightly more general than the one appearing in [17]. For the proof, we use a novel and efficient probabilistic approach based on multi-type Galton–Watson trees.

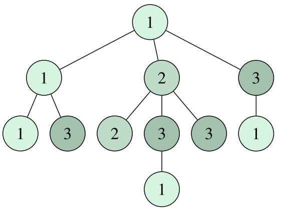

FIGURE 5 A ranked 3-type tree in $\vec{\mathbb{S}}_{1,(3,1,2)}$.

We first settle some notation. We will consider power series mappings $F = (F_1, \dots, F_n)$ with ith component of the form

$$F_i(X_1, \dots, X_n) = X_i - \sum_{\alpha \neq 0} h_{i,\alpha} X^\alpha = X_i - \sum_{\alpha \neq 0} \frac{H_{i,\alpha}}{\alpha!} X^\alpha, \tag{2.1}$$

where the sum is taken over multi-indices $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n_{\geqslant 0} \setminus \{0\}$, and we set $\alpha! := \alpha_1! \cdots \alpha_n!$ and $X^\alpha := X_1^{\alpha_1} \cdots X_n^{\alpha_n}$. As such, we canonically associate F with the set of complex coefficients $h = (h_{i,\alpha} : i \in [n], \alpha \neq 0)$ or $H = (H_{i,\alpha} : i \in [n], \alpha \neq 0)$. We will always use an upper case $H_{i,\alpha}$ when the coefficients are divided through by $\alpha!$. Working with the undivided coefficients is more amenable for probabilistic arguments, while working with the divided coefficients yields cleaner formulas when using the algebraic condition of nilpotency. Therefore, we initially use the $h_{i,\alpha}$ to prove the inversion formula; we then rephrase the latter in terms of the $H_{i,\alpha}$ and continue with the upper case coefficients in the study of nilpotency in Section 3.

2.1 | Inversion formula for ranked trees

A labelled planar rooted tree $\mathcal{T} = (T, \tau)$ is a planar rooted tree $T = (V, E)$ together with a type function $\tau : V \to [n]$ giving the vertices types in $[n]$; we also call this an n-**type tree**. We say such a tree is **ranked** if siblings of smaller types always lie to the left of siblings of larger types.

Given an n-type tree $\mathcal{T}$, we write $\mathrm{Leaftype}(\mathcal{T}) \in \mathbb{Z}^n_{\geqslant 0}$ for the multi-index with components

$$\mathrm{Leaftype}(\mathcal{T})_j := \#\{\text{type } j \text{ leaves of } \mathcal{T}\}, \qquad j = 1, \dots, n.$$

Given $i \in [n]$ and $\alpha \in \mathbb{Z}^n_{\geqslant 0}$, let $\mathbb{S}_{i,\alpha}$ and $\vec{\mathbb{S}}_{i,\alpha}$ denote, respectively, the set of (finite) unranked and ranked n-type trees with root type i and $\mathrm{Leaftype}(\mathcal{T}) := \alpha$ (i.e. the tree has α_j leaves of type j). See Figure 5 for an example. In what follows, for j in $[n]$, we will write $\mathbf{e}_j \in \mathbb{Z}^n_{\geqslant 0}$ for the multi-index whose jth component is a 1 and whose other components are 0. Let $(h_{i,\alpha} : i \in [n], \alpha \in \mathbb{Z}^n_{\geqslant 0})$ be a collection of complex coefficients. For an internal (i.e. non-leaf) vertex v, let $\mu(v) \in \mathbb{Z}^n_{\geqslant 0}$ be the multi-index whose jth component counts the number of type j children of v. Define a second collection of coefficients by

$$g_{i,\alpha} := \sum_{\mathcal{T} \in \vec{\mathbb{S}}_{i,\alpha}} \prod_{\substack{v \text{ internal} \\ \text{vertex of } \mathcal{T}}} h_{\tau(v),\mu(v)}. \tag{2.2}$$

We note that the sum in (2.2) is over an infinite set, since, for any fixed α, a tree in $\vec{\mathbb{S}}_{i,\alpha}$ may contain an arbitrarily long path of vertices, each of whom has exactly one child. We will assume for the remainder of this section that the linear coefficients $h_{i,\mathbf{e}_j}$ are such that

$$\sum_{k=0}^{\infty} A^k \text{ converges,} \quad \text{where } A = (A_{i,j})_{1\leqslant i,j\leqslant n}, \quad A_{i,j} := h_{i,\mathbf{e}_j}. \tag{2.3}$$

This assumption is sufficient to ensure that the sum in (2.2) converges; see the discussion on alternating trees in [17] for details.

We now state our main result of this section:

Theorem 2.1. *Suppose* $(h_{i,\alpha} : i \in [n], \alpha \in \mathbb{Z}^n_{\geqslant 0} \setminus \{0\})$ *is a collection of complex coefficients satisfying (2.3), and that the power series mapping* $F = (F_1, \dots, F_n)$ *with ith component*

$$F_i(X_1, \dots, X_n) := X_i - \sum_{\alpha \neq 0} h_{i,\alpha} X^\alpha \tag{2.4}$$

converges in an open neighbourhood of the origin in $\mathbb{C}^n$. *Then,* F *has an analytic inverse* $F^{-1} = (F_1^{-1}, \dots, F_n^{-1})$, *with ith component given by*

$$F_i^{-1}(X_1, \dots, X_n) := \sum_{\alpha \neq 0} g_{i,\alpha} X^\alpha,$$

where $g_{i,\alpha}$ *are defined in (2.2).*

Theorem 2.1 is new in this exact formulation, but it is similar to countless other power series inversion formulas in the literature. The special case of Theorem 2.1 in which each linear term $h_{i,\mathbf{e}_j}$ is zero is well trodden in the literature [8, 11, 32, 33]. An explicit formula for the case with non-zero linear terms only appeared recently in [17], but the formulation therein involves *alternating* trees and is more complicated. We thus believe that the statement in Theorem 2.1 represents the most general and cleanest inversion formula for multivariate power series so far.

The novelty here lies in our probabilistic derivation, which is substantially more efficient than other approaches. Before proving Theorem 2.1 in Section 2.3, we first define and discuss multi-type Galton–Watson trees.

2.2 | Multi-type Galton–Watson trees

In this section, we define multi-type Galton–Watson trees and provide some background.

Recalling that $\vec{\mathbb{S}}_{i,\alpha}$ denotes the set of (finite) ranked n-type trees with root type i and leaf type α, write

$$\vec{\mathbb{S}}_i := \text{Finite ranked } n\text{-type trees with root type } i = \bigcup_{\alpha \in \mathbb{Z}^n_{\geqslant 0}} \vec{\mathbb{S}}_{i,\alpha},$$

$$\vec{\mathbb{S}}_{i,*} := \text{Finite or infinite ranked } n\text{-type trees with root type } i.$$

Take a collection $(h_{i,\alpha} : i \in [n], \alpha \in \mathbb{Z}^n_{\geqslant 0})$ of complex numbers, indexed by $\alpha \in \mathbb{Z}^n_{\geqslant 0}$, as opposed to $\alpha \in \mathbb{Z}^n_{\geqslant 0} \setminus \{0\}$ as in the statement of Theorem 2.1. We say that this collection is a **multi-type offspring distribution** if $(h_{i,\alpha})_{\alpha \in \mathbb{Z}^n_{\geqslant 0}}$ is a probability mass function on $\mathbb{Z}^n_{\geqslant 0}$ for each i, that is, if

$$h_{i,\alpha} \in [0,1] \quad \text{for each } (i,\alpha) \qquad \text{and} \qquad \sum_{\alpha \in \mathbb{Z}^n_{\geqslant 0}} h_{i,\alpha} = 1. \tag{2.5}$$

A multi-type Galton–Watson tree with root type i is a random element of $\vec{\mathbb{S}}_{i,*}$, defined by the following simple rule: the probability that a vertex of type j has children of types β is $h_{j,\beta}$. More specifically, the tree may be constructed as follows. We begin in generation 0 with the root, which has type i. Thereafter, we grow the labelled tree outwards from the root. Suppose we have grown up until generation k. Then, for each vertex v of some type j in generation k, we choose a random element $\mu(v)$ in $\mathbb{Z}^n_{\geqslant 0}$ according to the probability mass function $(h_{j,\beta})_{\beta \in \mathbb{Z}^n_{\geqslant 0}}$; in particular, $h_{j,0}$ is the probability that v is a leaf. We then order the children of v so that those of smaller types are placed to the left of those of larger types. This creates a random ranked tree $\mathcal{T}$ in $\vec{\mathbb{S}}_{i,*}$. It is entirely possible (say, for instance, if $h_{j,0} = 0$ for each j) that with positive probability this process never terminates, and that the labelled tree we construct is infinite. We note that some authors do not require in their definition of a multi-type Galton–Watson tree that the children of a given vertex are ordered from left to right by their type; see, for example, [21, 26].

Let $\mathbf{P}^h_i$ denote the distribution of the Galton–Watson tree with root type i and offspring distribution $(h_{j,\beta})$. Then, the probability of a tree $\mathcal{T} \in \vec{\mathbb{S}}_{i,*}$ occurring is

$$\mathbf{P}^h_i(\mathcal{T}) = \prod_{v \text{ vertex of } \mathcal{T}} h_{\tau(v),\mu(v)}, \tag{2.6}$$

where the product is over all vertices of $\mathcal{T}$, including all internal vertices and leaves.

For $\mathcal{T} \in \vec{\mathbb{S}}_{i,*}$, define the associated process $(Z(k))_{k \geqslant 0}$ as the $\mathbb{Z}^n_{\geqslant 0}$-valued sequence given by

$$Z(k)_j := \#\{\text{vertices of } \mathcal{T} \text{ of type } j \text{ in generation } k\}, \tag{2.7}$$

where generation k of the tree refers to the set of vertices graph distance k from the root. Note that if $\mathcal{T}$ is finite, then $Z(k) = 0$ for all sufficiently large k.

Clearly, if a tree $\mathcal{T}$ in $\vec{\mathbb{S}}_{i,*}$ is chosen randomly according to $\mathbf{P}^h_i$, then $(Z(k))_{k \geqslant 0}$ is a vector-valued Markov process under $\mathbf{P}^h_i$. Moreover, $(Z(k))_{k \geqslant 0}$ is a branching process. Indeed, let $Z^{j,\ell}(k+1)$ be the multi-index coding the children of the ℓth vertex of type j in generation k, for all $k \geqslant 0$, $1 \leqslant j \leqslant n$ and $1 \leqslant \ell \leqslant Z(k)_j$. It is then clear that the sum of all $Z^{j,\ell}(k+1)$ over j and ℓ codes the whole generation $k+1$:

$$Z(k+1) = \sum_{j=1}^{n} \sum_{\ell=1}^{Z(k)_j} Z^{j,\ell}(k+1).$$

Note that, conditionally on the event $\{Z(k) = \alpha\}$, there are α_j vertices of each type j in generation k, and $(Z^{j,\ell}(k+1) : 1 \leqslant j \leqslant n, 1 \leqslant \ell \leqslant \alpha_j)$ are independent random variables such that each $Z^{j,\ell}(k+1)$ is distributed according to $(h_{j,\beta})_{\beta \in \mathbb{Z}^n_{\geqslant 0}}$. Equivalently, each $Z^{j,\ell}(k+1)$ has the same distribution as $\mu(v_{\text{root}})$ in a Galton–Watson tree in which the root has type j. The branching property will be integral to our probabilistic proof of Theorem 2.1.

While it will not be relevant to our goals in the remainder of the manuscript, let us close this section by discussing the asymptotic behaviour of multi-type Galton–Watson trees. Let $\mathbf{E}_i^h$ be the expectation associated with the probability measure $\mathbf{P}_i^h$. If we define

$$M_{i,j} := \mathbf{E}_i^h[Z(1)_j] = \sum_{\alpha \in \mathbb{Z}_{\geqslant 0}^n} h_{i,\alpha} \alpha_j$$

to be the expected number of type j children had by a type i vertex, then the asymptotic behaviour of $\mathbf{P}_i^h$ can be well-described in terms of spectral properties of the mean matrix $M := (M_{i,j})_{1 \leqslant i,j \leqslant n}$. Namely, if M has an eigenvalue strictly larger than 1 (**supercritical** case) and satisfies a certain irreducibility property, then a tree under $\mathbf{P}_i^h$ has a positive probability of being infinite and, accordingly, the sum of probabilities over finite trees satisfies

$$\sum_{\mathcal{T} \in \vec{\mathbb{S}}_i} \mathbf{P}_i^h(\mathcal{T}) < 1.$$

On the other hand, whenever every eigenvalue of $M_{i,j}$ is less than one (**subcritical** case), the tree is guaranteed to be almost surely finite, that is,

$$\sum_{\mathcal{T} \in \vec{\mathbb{S}}_i} \mathbf{P}_i^h(\mathcal{T}) = 1.$$

We refer the reader to [7] for further discussion of (multi-type) Galton–Watson trees.

2.3 | Proof of Theorem 2.1

Let us give a brief overview of our proof of Theorem 2.1. We begin by appealing to the inverse function theorem to establish that F has an analytic inverse F^{-1} in a neighbourhood of the origin. Thereafter, we observe that the coefficients of degree $\leqslant k$ for the inverse can be expressed as polynomials in the coefficients of degree $\leqslant k$ in the original power series F. We then appeal to analytic continuation, which allows us to assume that the coefficients of F form a multi-type offspring distribution. Under this assumption, our proof of Theorem 2.1 involves studying a generating function of the leaf type of a multi-type Galton–Watson tree. We utilise the branching property to show that this generating function satisfies a functional equation, which we then manipulate to prove Theorem 2.1.

Recall that Theorem 2.1 considers a general collection $(h_{i,\alpha} : i \in [n], \alpha \in \mathbb{Z}_{\geqslant 0}^n \setminus \{0\})$ of complex coefficients satisfying property (2.3). However, **until stated otherwise, we will assume for the remainder of this section that**

$$h_{i,\alpha} \in [0,1] \quad \text{for each } (i,\alpha) \qquad \text{and} \qquad \sum_{\alpha \in \mathbb{Z}_{\geqslant 0}^n \setminus \{0\}} h_{i,\alpha} < 1. \tag{2.8}$$

We may naturally extend the coefficients to a collection also indexed by $\alpha = 0$ by setting

$$d_i := h_{i,0} := 1 - \sum_{\alpha \in \mathbb{Z}_{\geqslant 0}^n \setminus \{0\}} h_{i,\alpha}. \tag{2.9}$$

It then follows from (2.8) and (2.9) that $d_i > 0$ and that (2.5) holds, so that the collection $(h_{i,\alpha} : i \in [n], \alpha \in \mathbb{Z}^n_{\geqslant 0})$ is a multi-type offspring distribution.

By (2.6) we have, for a given $\mathcal{T} \in \vec{\mathbb{S}}_{i,*}$,

$$\mathbf{P}_i^h(\mathcal{T}) = \prod_{\substack{v \text{ internal} \\ \text{vertex of } \mathcal{T}}} h_{\tau(v),\mu(v)} \prod_{\substack{v \text{ leaf} \\ \text{vertex of } \mathcal{T}}} d_{\tau(v)}.$$

Notice that

$$\prod_{\substack{v \text{ leaf} \\ \text{vertex of } \mathcal{T}}} d_{\tau(v)} = d_1^{\alpha_1} \cdots d_n^{\alpha_n} =: d^\alpha \qquad \text{for all } \mathcal{T} \in \vec{\mathbb{S}}_{i,\alpha}.$$

In particular, recalling the definition (2.2) of $g_{i,\alpha}$, the probability that a multi-type Galton–Watson tree with root type i is finite with leaf type α is given by

$$\mathbf{P}_i^h(\mathcal{T} \text{ is finite, Leaftype}(\mathcal{T}) = \alpha) = \sum_{\mathcal{T} \in \vec{\mathbb{S}}_{i,\alpha}} \mathbf{P}_i^h(\mathcal{T}) = d^\alpha g_{i,\alpha}. \tag{2.10}$$

The key object we need is the generating function $E = (E_1, \dots, E_n)$ for the leaf type:

$$E_i(s_1, \dots, s_n) := \mathbf{E}_i^h\left[\mathbf{1}_{\{\mathcal{T} \text{ finite}\}} s^{\text{Leaftype}(\mathcal{T})}\right] = \sum_{\alpha \neq 0} \mathbf{P}_i^h(\mathcal{T} \text{ is finite, Leaftype}(\mathcal{T}) = \alpha)\, s^\alpha,$$

which is analytically well defined for $(s_1, \dots, s_n)$ in a neighbourhood of the origin in $\mathbb{C}^n$. It follows from (2.10) that

$$E_i(s_1, \dots, s_n) = \sum_{\alpha \neq 0} g_{i,\alpha} d^\alpha s^\alpha = g_i(d_1 s_1, \dots, d_n s_n) = g_i(D(s)),$$

where $g = (g_1, \dots, g_n)$ and $D = (D_1, \dots, D_n)$ are the maps defined by

$$g_i(X) := \sum_{\alpha \neq 0} g_{i,\alpha} X^\alpha, \qquad D_i(X_1, \dots, X_n) := d_i X_i.$$

In short, we have

$$E = gD. \tag{2.11}$$

We now use the branching property to derive a functional equation for E.

Lemma 2.2. *Suppose that $h = (h_1, \dots, h_n)$, with $h_i(X_1, \dots, X_n) := \sum_{\alpha \neq 0} h_{i,\alpha} X^\alpha$, satisfies (2.8). Then, the leaf-type generating function $E = (E_1, \dots, E_n)$ satisfies the functional equation*

$$E = hE + D. \tag{2.12}$$

Proof. Recall the multi-type Galton–Watson process $(Z(k))_{k \geqslant 0}$ associated with a multi-type Galton–Watson tree and defined in (2.7).

By conditioning on $Z(1)$, we have

$$E_i(s_1, \dots, s_n) = \sum_{\alpha\in\mathbb{Z}^n_{\geqslant 0}} \mathbf{P}^h_i(Z(1)=\alpha)\,\mathbf{E}^h_i\left[\mathbf{1}_{\{\mathcal{T}\text{ finite}\}} s^{\mathrm{Leaftype}(\mathcal{T})}\,\middle|\,Z(1)=\alpha\right]. \tag{2.13}$$

Now, on the one hand, in the case $\alpha = 0$, we have

$$\mathbf{E}^h_i\left[\mathbf{1}_{\{\mathcal{T}\text{ finite}\}} s^{\mathrm{Leaftype}(\mathcal{T})}\,\middle|\,Z(1)=0\right] = s_i. \tag{2.14}$$

On the other hand, according to the branching property, for $\alpha \neq 0$ we have

$$\mathbf{E}^h_i\left[\mathbf{1}_{\{\mathcal{T}\text{ finite}\}} s^{\mathrm{Leaftype}(\mathcal{T})}\,\middle|\,Z(1)=\alpha\right] = \prod_{j=1}^n E_j(s_1,\dots,s_n)^{\alpha_j} =: E(s_1,\dots,s_n)^\alpha. \tag{2.15}$$

Using (2.14) and (2.15) in (2.13), we obtain

$$E_i(s_1,\dots,s_n) = \sum_{\alpha\neq 0} h_{i,\alpha} E(s_1,\dots,s_n)^\alpha + d_i s_i = h_i(E(s_1,\dots,s_n)) + d_i s_i,$$

and thus, (2.12) holds. □

Proof of Theorem 2.1. Let $F : \mathbb{C}^n \to \mathbb{C}^n$ be a map with components of the form (2.4). Since (2.3) guarantees that the matrix $I - A$ is invertible (with inverse given by $\sum_{k=0}^\infty A^k$), the Jacobian matrix of F at the origin

$$\left(\frac{\partial}{\partial X_j} F_i(X_1,\dots,X_n)\right)|_{(X_1,\dots,X_n)=0} = \delta_{i,j} - h_{i,\mathbf{e}_j} = (I-A)_{i,j}$$

is invertible. By the inverse function theorem (see, e.g. [28, Section 1.3]), F has an analytic inverse $F^{-1} : \mathbb{C}^n \to \mathbb{C}^n$ that satisfies $F^{-1}(0) = 0$ and converges in some neighbourhood of the origin. Thus, we may write $F^{-1}_i(X_1,\dots,X_n) = \sum_{\alpha\neq 0} c_{i,\alpha} X^\alpha$, for certain complex coefficients $c_{i,\alpha}$. We need to show that $F^{-1} = g$, that is, that $c_{i,\alpha} = g_{i,\alpha}$ is the coefficient defined in (2.2), for all i and α.

By considering the higher-order derivatives of $I = F \circ F^{-1}$ (or, say, using the multivariate Faà di Bruno formula [17]), one can see that the X^α coefficient $c_{i,\alpha}$ of F^{-1}_i may be expressed as a polynomial in the coefficients of F and $I - A$ of degree at most $|\alpha|$. Notice also that, if $(h_{i,\alpha} : i \in [n], \alpha \in \mathbb{Z}^n_{\geqslant 0} \setminus \{0\})$ satisfy (2.8), then $A := (h_{i,\mathbf{e}_j})_{1\leqslant i,j\leqslant n}$ is a strictly sub-stochastic matrix (i.e. it has non-negative entries and row sums strictly less than 1), and accordingly $I - A$ is invertible and (2.3) holds. It follows that, in order to characterise $c_{i,\alpha}$ as a polynomial in the variables $(h_{j,\beta} : j \in [n], \beta \neq 0, |\beta| \leqslant |\alpha|)$, it is sufficient, by analytic continuation, to do so under the constraint (2.8). Thus, for the remainder of the proof, we may assume (2.8) holds.

Under the assumption (2.8), by (2.12), we have $E = hE + D$, that is, $(I - h)E = D$. By (2.11), we also have $E = gD$, so that

$$(I-h)gD = D.$$

Since $d_i > 0$ for all i, we have that D is invertible and we can compose on the right by D^{-1} to obtain

$$(I - h)g = I.$$

This proves that g is a right inverse of $F = I - h$. Since the set of power series mappings in n variables with invertible linear term is a group under composition (see, e.g. [29, Section 1.1]), g is also a left inverse, so that $g = F^{-1}$, as desired. □

2.4 | Inversion formula for unranked trees

We now give an analogue of Theorem 2.1 for power series with the $\alpha!$ in the coefficients (see (2.1)). The modified formulas involve unranked trees (in $\mathbb{S}_{i,\alpha}$) rather than ranked trees (in $\vec{\mathbb{S}}_{i,\alpha}$).

Theorem 2.3. *Let $F : \mathbb{C}^n \to \mathbb{C}^n$ be a power series mapping with components of the form*

$$F_i(X_1, \dots, X_n) := X_i - \sum_{\alpha \neq 0} \frac{H_{i,\alpha}}{\alpha!} X^\alpha,$$

which converge in a neighbourhood of 0. Suppose that the series $\sum_{k=0}^{\infty} A^k$ converges, with $A_{i,j} = H_{i,\mathbf{e}_j}$. Then, F has an analytic inverse given by $F_i^{-1}(X_1, \dots, X_n) := \sum_{\alpha \neq 0} g_{i,\alpha} X^\alpha$, where

$$g_{i,\alpha} := \sum_{\mathcal{T} \in \mathbb{S}_{i,\alpha}} \prod_{\substack{v \text{ internal} \\ \text{vertex of } \mathcal{T}}} \frac{H_{\tau(v),\mu(v)}}{|\mu(v)|!}. \tag{2.16}$$

Here, $|\mu(v)|$ is the total number of children of an internal vertex v.

Proof. Setting $h_{i,\alpha} := H_{i,\alpha}/\alpha!$ and using Theorem 2.1, we see that the coefficients of F_i^{-1} are

$$g_{i,\alpha} = \sum_{\mathcal{T} \in \vec{\mathbb{S}}_{i,\alpha}} \prod_{\substack{v \text{ internal} \\ \text{vertex of } \mathcal{T}}} \frac{H_{\tau(v),\mu(v)}}{\mu(v)!}. \tag{2.17}$$

Consider the natural map from $\pi : \mathbb{S}_{i,\alpha} \to \vec{\mathbb{S}}_{i,\alpha}$ that takes a tree and reorders the children of each vertex so that the types increase from left to right, but otherwise maintains the order of child vertices of the same type. Note now that, given a vertex v with $\mu(v)$ children, there are $|\mu(v)|!$ ways of ordering these children in any order, and $\mu(v)!$ ways of ordering these children in such a way that the types of the children increases from left to right. In particular, for a given tree $\mathcal{T}$ in $\vec{\mathbb{S}}_{i,\alpha}$, we have

$$\#\{\mathcal{T}' \in \mathbb{S}_{i,\alpha} : \pi(\mathcal{T}') = \mathcal{T}\} = \prod_{\substack{v \text{ internal} \\ \text{vertex of } \mathcal{T}}} \frac{|\mu(v)|!}{\mu(v)!}. \tag{2.18}$$

We also note that the map π does not change the quantity $\prod_{\substack{v \text{ internal}\\ \text{vertex of } \mathcal{T}}} H_{\tau(v),\mu(v)}$, that is, if $\pi(\mathcal{T}') = \mathcal{T}$, then

$$\prod_{\substack{v \text{ internal}\\ \text{vertex of } \mathcal{T}'}} H_{\tau(v),\mu(v)} = \prod_{\substack{v \text{ internal}\\ \text{vertex of } \mathcal{T}}} H_{\tau(v),\mu(v)}. \tag{2.19}$$

Using (2.18) and (2.19) in (2.17), we obtain (2.16), completing the proof. □

Note that the inversion formula (1.2) is a consequence of Theorem 2.3. Indeed, if $H_{i,\alpha} = 0$ whenever $|\alpha| \neq d$, it follows from (2.16) that

$$g_{i,\alpha} := \frac{1}{(d!)^k} \sum_{\mathcal{T} \in \mathbb{S}_{i,\alpha},\, d\text{-ary}} \prod_{\substack{v \text{ internal}\\ \text{vertex of } \mathcal{T}}} H_{\tau(v),\mu(v)}.$$

Now note that a sum over all d-ary unranked labelled trees $\mathcal{T}$ in $\mathbb{S}_{i,\alpha}$ is equivalent to a sum over all (i, α) labellings of a d-Catalan tree. Thus, we obtain (1.2).

3 | KELLER PROPERTY AND COMBINATORIAL NILPOTENCY

In this section, we provide a more complete discussion of the combinatorial approach to the Jacobian conjecture, as outlined in Section 1.2.

We first appeal to the celebrated Bass–Connell–Wright [8] reduction of the Jacobian conjecture.

Theorem 3.1 (Bass, Connell and Wright [8]). *Suppose there exists $d \geqslant 3$ with the following property: for all n, every polynomial mapping of the form $F = I - H : \mathbb{C}^n \to \mathbb{C}^n$ with H homogeneous of degree d and JH nilpotent, is a polynomial automorphism. Then the Jacobian conjecture holds in full generality.*

We will then restrict from now on to polynomial mappings $F = I - H$, where

$$H : \mathbb{C} \to \mathbb{C}, \qquad H_i(X_1, \ldots, X_n) = \sum_{|\alpha|=d} \frac{H_{i,\alpha}}{\alpha!} X^\alpha \qquad \text{for all } i = 1, \ldots, n. \tag{3.1}$$

Combining Theorem 3.1 with Theorem 2.3 conveys the combinatorial approach to the Jacobian conjecture mentioned in the introduction: in order to prove the Jacobian conjecture, it is sufficient to verify that, whenever $F = I - H$ is a mapping with H of the form (3.1) and JH nilpotent, the inverse coefficients $g_{i,\alpha}$ defined in (2.16) are zero for all but finitely many α.

In this setting, we now work towards giving an even more explicit combinatorial statement of the Jacobian conjecture, with the nilpotency of JH (equivalent to the Keller property) also put in combinatorial terms, and more precisely in terms of shuffle classes of d-Catalan trees. Strictly speaking, such a combinatorial interpretation is not new, and variations of it have appeared in numerous places [17, 23, 24, 34]. In particular, we generalise from $C_k^{(2)}$ to $C_k^{(d)}$ tools developed by Singer in [23, 24], providing concise proofs.

3.1 | Shuffling d-Catalan trees

We begin by defining shuffle classes. Roughly speaking, a length-p shuffle class is a collection of trees that may be obtained from one another by shuffling the subtrees coming off a fixed length-p path of descendant vertices.

Recall from the introduction that, for $k \geqslant 1$, $C_k^{(d)}$ denotes the set of d-Catalan trees with k internal vertices; such trees have $dk+1$ vertices in total, and hence $(d-1)k+1$ leaves. An **(ancestral) path** of length p in a d-Catalan tree is a sequence of vertices $(v_0, \dots, v_p)$ such that each v_i is a child of v_{i-1}. The **height** of a vertex v is its graph distance from the root. Each vertex v of height at least p is associated to a unique path $(v_0, \dots, v_p)$ of length p ending with $v = v_p$.

Definition 3.2 (Shuffle class). Let $T \in C_k^{(d)}$. Let $p \geqslant 1$ be an integer and assume that the vertex v of T has height at least p. Let $(v_0, \dots, v_p)$ be the unique path associated with $v = v_p$. For each $i \geqslant 1$, v_i has $d-1$ siblings $w_{i,1}, \dots, w_{i,d-1}$; each of these siblings $w_{i,j}$ subtends a subtree $W_{i,j}$, which is itself a d-Catalan tree with fewer leaves than T. The **shuffle class** $\mathrm{Sh}(T; v_0, \dots, v_p)$ of the path $(v_0, \dots, v_p)$ is the subset of $C_k^{(d)}$ consisting of all Catalan trees T' that may be obtained from T by reshuffling the $(d-1)p$ subtrees $W_{i,j}$ across all p generations. See Figure 3 for an illustration.

Note that a shuffle class $\mathrm{Sh}(T; v_0, \dots, v_p)$ may contain up to $((d-1)p)!$ different trees in $C_k^{(d)}$. However, if some of the subtrees $W_{i,j}$ are the same, then the shuffle class will contain less than $((d-1)p)!$ elements. It is even possible that all the subtrees $W_{i,j}$ are identical; in this case, the shuffle class will consist of exactly one tree.

If a shuffle class contains more than one element, it can be represented in different ways. Indeed, if $T' \in \mathrm{Sh}(T; v_0, \dots, v_p)$, then there exists a path $(v_0', \dots, v_p')$ of vertices in T' such that $\mathrm{Sh}(T; v_0, \dots, v_p) = \mathrm{Sh}(T'; v_0', \dots, v_p')$.

Recall from Section 2.1 that, for $i \in [n]$ and $\alpha \in \mathbb{Z}_{\geqslant 0}^n$, $\mathbb{S}_{i,\alpha}$ is the set of (unranked) n-type trees $\mathcal{T} = (T, \tau)$ with root type i and α_j leaves of type j. Equivalently, as in Section 1.2, we say that $\mathcal{T}$ is an (i, α) **labelling** of T. Recall also that, for any internal vertex v, $\mu(v)$ is the multi-index whose jth component counts the number of type j children of v.

Suppose now $T \in C_k^{(d)}$, so that $|\mu(v)| = d$ for all internal vertices v of T, and the total number of leaves is $|\alpha| = (d-1)k + 1$. Let also $H = (H_{i,\alpha} : i \in [n], \alpha \in \mathbb{Z}_{\geqslant 0}^n, |\alpha| = d)$ be a collection of coefficients in $\mathbb{C}$. Recall from Section 1.2 that the H-**weight** of the labelling $\mathcal{T}$ of T is

$$\mathcal{E}_H(\mathcal{T}) := \prod_{\substack{v \text{ internal} \\ \text{vertex of } \mathcal{T}}} H_{\tau(v), \mu(v)}. \tag{3.2}$$

We also define the **average** H-**weight** of T as

$$E_{i,\alpha,H}(T) := \sum_{(i,\alpha) \text{ labellings } \mathcal{T} \text{ of } T} \mathcal{E}_H(\mathcal{T}). \tag{3.3}$$

The value of defining shuffle classes lies in the following lemma, which states that the nilpotency of JH guarantees that sums of average H-weights over shuffle classes are zero.

Lemma 3.3 (Shuffle lemma). *Let H be of the form (3.1). Suppose that the Jacobian matrix JH of H satisfies $(JH)^p = 0$ for some integer $1 \leqslant p \leqslant n$. Then*

$$\sum_{T\in S} E_{i,\alpha,H}(T) = 0 \qquad \textit{for all } i \in [n] \textit{ and } |\alpha| = (d-1)k+1$$

and for every shuffle class S of $C_k^{(d)}$ associated with a path of length p.

In Lemma 3.3, we have isolated the so-called index of nilpotency p from the underlying dimension n of the polynomial mapping H. Of course, if JH is nilpotent, then $(JH)^n = 0$ and we may set $p = n$.

We note again that Lemma 3.3 is not strictly new: the general idea is present in the works of Singer [23, 24], where an analogous result is proven in the $d = 2$ case. We will prove Lemma 3.3 in Section 3.3.

3.2 | Algebraic nilpotency

Our first step is to express the nilpotency of JH in terms of explicit algebraic equations in the coefficients $H_{i,\alpha}$.

Lemma 3.4. *Let H be of the form (3.1). Then, for any integer $1 \leqslant p \leqslant n$, we have $(JH)^p = 0$ if and only if*

$$\sum_{\beta^1+\cdots+\beta^p=\beta} \left(\prod_{\ell=1}^{p} \frac{(d-1)!}{\beta^\ell!}\right) \sum_{k_1,\ldots,k_{p-1}\in[n]} \prod_{\ell=1}^{p} H_{k_{\ell-1},\beta^\ell+\mathbf{e}_{k_\ell}} = 0$$

for all $i, j \in [n]$ and $|\beta| = (d-1)p$, where $k_0 := i$, $k_p := j$, and the outer sum is over all p-tuples $(\beta^1, \ldots, \beta^p)$ of multi-indices β^ℓ in $\mathbb{Z}^n_{\geqslant 0}$ with $|\beta^\ell| = d-1$ for all ℓ, whose sum is the multi-index β.

Proof. The (i, j)th entry of $JH = (\frac{\partial}{\partial X_j} H_i)_{1\leqslant i,j\leqslant n}$ is given by the polynomial

$$(JH)_{i,j}(X_1, \ldots, X_n) = \sum_{|\alpha|=d\,:\,\alpha_j\geqslant 1} \frac{H_{i,\alpha}\alpha_j}{\alpha!} X^{\alpha-\mathbf{e}_j} = \sum_{|\beta|=d-1} \frac{H_{i,\beta+\mathbf{e}_j}}{\beta!} X^\beta.$$

The (i, j)th entry of $(JH)^p$ is the polynomial

$$((JH)^p)_{i,j}(X_1, \ldots, X_n) := \sum_{|\beta^1|,\ldots,|\beta^p|=d-1} \sum_{k_1,\ldots,k_{p-1}\in[n]} \prod_{\ell=1}^{p} \frac{H_{k_{\ell-1},\beta^\ell+\mathbf{e}_{k_\ell}}}{\beta^\ell!} X^{\beta^\ell}, \tag{3.4}$$

where $k_0 := i$, $k_p := j$, and the outer sum is over all p-tuples of multi-indices β^ℓ in $\mathbb{Z}^n_{\geqslant 0}$ with $|\beta^\ell| = d-1$ for all ℓ.

Now, we have $(JH)^p = 0$ if and only if (3.4) is zero for all i, j. Since $(JH)^p$ is clearly a homogeneous polynomial of degree $(d-1)p$, we have $(JH)^p = 0$ if only if the coefficient of X^β in $((JH)^p)_{i,j}(X_1, \ldots, X_n)$ is zero for all i, j and for each $|\beta| = (d-1)p$. By (3.4), such a coefficient is given by

$$\sum_{\beta^1+\cdots+\beta^p=\beta}\ \sum_{k_1,\dots,k_{p-1}\in[n]}\ \prod_{\ell=1}^{p}\frac{H_{k_{\ell-1},\beta^\ell+\mathbf{e}_{k_\ell}}}{\beta^\ell!},$$

where the outer sum is over all p-tuples $(\beta^1,\dots,\beta^p)$ of multi-indices β^ℓ in $\mathbb{Z}^n_{\geqslant 0}$ with $|\beta^\ell| = d-1$ for all ℓ, whose sum is β. The result then follows by multiplying this coefficient by $((d-1)!)^p$. $\square$

3.3 | Combinatorial nilpotency: labelled version

From Lemma 3.4, we now deduce a 'labelled version' of the shuffle lemma (Lemma 3.3). We thus begin by defining the notion of labelled shuffle classes.

Definition 3.5. Let $\mathcal{T} = (T,\tau)$ be an n-type tree, where $T \in C_k^{(d)}$. Let $(v_0,\dots,v_p)$ be any length-p path in $\mathcal{T}$. The **labelled shuffle class** $\mathrm{Sh}(\mathcal{T}; v_0,\dots,v_p)$ is the set of n-type trees $\mathcal{T}'$ that may be obtained from $\mathcal{T}$ by:

- rearranging the labelled subtrees subtended by the $(d-1)p$ siblings of $v_1,\dots,v_p$;
- giving arbitrary new types in $\{1,\dots,n\}$ to the vertices $v_1,\dots,v_{p-1}$.

A labelled shuffle class $\mathrm{Sh}(\mathcal{T}; v_0,\dots,v_p)$ of length p may contain up to $n^{p-1}((d-1)p)!$ different labelled trees. If some of the labelled subtrees subtended by the siblings are identical (in both their topology and labelling), then the respective labelled shuffle class will contain fewer elements.

Lemma 3.6. *Let H be of the form (3.1). Suppose that the Jacobian matrix JH of H satisfies $(JH)^p = 0$ for some integer $1 \leqslant p \leqslant n$. Then, for all n-type tree $\mathcal{T} = (T,\tau)$, with $T \in C_k^{(d)}$, and every path $(v_0,\dots,v_p)$ in $\mathcal{T}$, we have*

$$\sum_{\mathcal{T}'\in \mathrm{Sh}(\mathcal{T};v_0,\dots,v_p)} \mathcal{E}_H(\mathcal{T}') = 0.$$

Proof. We can decompose the weight of $\mathcal{T}$ as

$$\mathcal{E}_H(\mathcal{T}) = \prod_{v \text{ internal}} H_{\tau(v),\mu(v)} = \mathcal{E}_H(\mathcal{R})\mathcal{E}_H(\mathcal{S})\left(\prod_{i=1}^{p}\prod_{j=1}^{d-1}\mathcal{E}_H(\mathcal{W}_{i,j})\right)\mathcal{E}_H(\mathcal{U}),$$

where

- $\mathcal{R}$ is the labelled subtree obtained from $\mathcal{T}$ by removing any vertex that is a descendant of v_0 (but keeping v_0 itself);
- $\mathcal{S}$ is the labelled subtree given by v_p and its descendants;
- $\mathcal{W}_{i,j}$ is the labelled subtree given by $w_{i,j}$ and its descendants, where, for each $1 \leqslant i \leqslant p$, $w_{i,1},\dots,w_{i,d-1}$ are the $d-1$ siblings of v_i;
- $\mathcal{U}$ is the labelled subtree consisting of the vertices $v_0,\dots,v_p$ and the siblings of $v_1,\dots,v_p$.

As a shorthand, we write

$$G_H(\mathcal{T}) := \mathcal{E}_H(\mathcal{R})\mathcal{E}_H(\mathcal{S})\prod_{i=1}^{p}\prod_{j=1}^{d-1}\mathcal{E}_H(\mathcal{W}_{i,j}).$$

Notice now that any $\mathcal{T}' \in \mathrm{Sh}(\mathcal{T}; v_0, \dots, v_p)$ contains the trees $\mathcal{R}$, $\mathcal{S}$, and each of the $\mathcal{W}_{i,j}$. Thus, its weight differs from $\mathcal{T}$ only in the labelling of $v_1, \dots, v_{p-1}$ and the rearrangement of the siblings. Therefore,

$$\mathcal{E}_H(\mathcal{T}') = G_H(\mathcal{T})\mathcal{E}_H(\mathcal{U}'),$$

where $\mathcal{U}'$ is the labelled subtree of $\mathcal{T}'$ consisting of the vertices $v_0, \dots, v_p$ and the siblings of $v_1, \dots, v_p$. In particular,

$$\sum_{\mathcal{T}' \in \mathrm{Sh}(\mathcal{T}; v_0, \dots, v_p)} \mathcal{E}_H(\mathcal{T}') = G_H(\mathcal{T}) \sum_{\mathcal{U}'} \mathcal{E}_H(\mathcal{U}'), \tag{3.5}$$

where the sum is over all labelled trees $\mathcal{U}'$ that can be obtained from $\mathcal{U}$ by relabelling $v_1, \dots, v_{p-1}$ and by rearranging the positions of its leaves (except v_p). Suppose that the total leaf type of $\mathcal{U}$ is β, that is, the total number of leaves of type k in $\mathcal{U}$ is the kth component of β; then, the total leaf type of each $\mathcal{U}'$ is also β.

Note that each such $\mathcal{U}'$ has (potentially) different labels $k_1, \dots, k_{p-1} \in [n]$ for the vertices $v_1, \dots, v_{p-1}$. Note also that, given any p-tuple $(\beta^1, \dots, \beta^p)$ of multi-indices of degree $d-1$ adding to β, and any $k_1, \dots, k_{p-1}$ in $[n]$, there are $\prod_{\ell=1}^{p} \frac{(d-1)!}{\beta^\ell !}$ different $\mathcal{U}'$ such that $\tau(v_\ell) = k_\ell$ for $1 \leqslant \ell \leqslant p-1$ and such that the types of the leaves in generation ℓ (excluding v_p) are encoded by the multi-index β^ℓ; this follows from the fact that there are $(d-1)!/\beta^\ell !$ different planar ways of rearranging the order of the $d-1$ leaves in generation ℓ such that, for each k, β^ℓ_k of them have type k. In short,

$$\sum_{\mathcal{U}'} \mathcal{E}_H(\mathcal{U}') = \sum_{\beta^1 + \dots + \beta^p = \beta} \left(\prod_{\ell=1}^{p} \frac{(d-1)!}{\beta^\ell !} \right) \sum_{k_1, \dots, k_{p-1} \in [n]} \prod_{\ell=1}^{p} H_{k_{\ell-1}, \beta^\ell + \mathbf{e}_{k_\ell}}, \tag{3.6}$$

where $k_0 := \tau(v_0)$ and $k_p := \tau(v_p)$ are fixed. Now plug (3.6) into (3.5), and use Lemma 3.4 to conclude. $\square$

The latter proof suggests that we may specialise Lemma 3.6 to a particular kind of trees, namely, the d-Catalan ferns introduced in Section 1.2. Recall that a d-**Catalan fern** is an element of $C_p^{(d)}$ of height p, with a designated 'sink' vertex in generation p, and the property that at most one vertex in each generation is not a leaf. Recall also that an (i, j, α) **labelling** of a d-Catalan fern U is any labelling of U giving the root type i, the sink type j and giving exactly α_ℓ of the non-sink leaves type ℓ, for all $\ell \in [n]$.

By taking $\mathcal{T}$ to be a labelled d-Catalan fern of height p in Lemma 3.6, we readily obtain the following result:

Corollary 3.7 ([17, Lemma 2.8]). *Let H be of the form (3.1). Suppose that the Jacobian matrix JH of H satisfies $(JH)^p = 0$ for some integer $1 \leqslant p \leqslant n$. Then, for any d-Catalan fern U of height p, we have*

$$\sum_{(i,j,\alpha) \text{ labellings } \mathcal{U} \text{ of } U} \mathcal{E}_H(\mathcal{U}) = 0.$$

Corollary 3.7 is precisely the third bullet point of the combinatorial approach to the Jacobian conjecture stated in Section 1.2.

3.4 | Combinatorial nilpotency: unlabelled version

We can now use Lemma 3.6 to prove Lemma 3.3.

Proof of Lemma 3.3. Let $\mathrm{Sh}(T; v_0, \dots, v_p)$ be a shuffle class associated with a path of length p, for some fixed $T \in C_k^{(d)}$. Let $i \in [n]$ and $|\alpha| = k(d-1)+1$. We need to show that

$$\sum_{T' \in \mathrm{Sh}(T; v_0, \dots, v_p)} \sum_{\substack{(i,\alpha) \text{ labellings} \\ \mathcal{T}' \text{ of } T'}} \mathcal{E}_H(\mathcal{T}') = 0. \tag{3.7}$$

The labelled tree $\mathcal{T}'$ in the expression above is obtained as follows:

(a) choose $T' \in \mathrm{Sh}(T; v_0, \dots, v_p)$;
(b) given such T', choose an (i, α) labelling $\mathcal{T}'$ of T'.

It is clear that any such $\mathcal{T}'$ can be obtained in only one way via the above procedure.

Consider now the alternative procedure:

(a') choose an (i, α) labelling $\mathcal{T}$ of T;
(b') given such $\mathcal{T}$, choose $\mathcal{T}' \in \mathrm{Sh}(\mathcal{T}; v_0, \dots, v_p)$.

In general, there are several different ways of obtaining any $\mathcal{T}'$ via the alternative procedure. We then have

$$\sum_{T' \in \mathrm{Sh}(T; v_0, \dots, v_p)} \sum_{\substack{(i,\alpha) \text{ labellings} \\ \mathcal{T}' \text{ of } T'}} \mathcal{E}_H(\mathcal{T}') = \sum_{\substack{(i,\alpha) \text{ labellings} \\ \mathcal{T} \text{ of } T}} \sum_{\mathcal{T}' \in \mathrm{Sh}(\mathcal{T}; v_0, \dots, v_p)} \frac{\mathcal{E}_H(\mathcal{T}')}{N(\mathcal{T}')}, \tag{3.8}$$

where $N(\mathcal{T}') \geqslant 1$ is the number of ways of obtaining $\mathcal{T}'$ via the second procedure. Let us now analyse such a number.

For $1 \leqslant i \leqslant p$, denote by $\mathcal{W}'_{i,j}$, $1 \leqslant j \leqslant d-1$, the labelled trees subtended by the $d-1$ siblings of v_i in $\mathcal{T}'$. The greater flexibility provided by the second procedure amounts to two facts: first, in the step (a') one may choose to label the vertices $v_1, \dots, v_{p-1}$ *arbitrarily*, since these can then be given new types in the labelled tree shuffling of step (b'); second, if in $\mathcal{T}'$ some of the $\mathcal{W}'_{i,j}$ have identical tree structure but different labellings, then there is more than one way to label such trees in step (a') so that an appropriate reshuffling in step (b') yields exactly $\mathcal{T}'$. More precisely, we have

$$N(\mathcal{T}') = n^{p-1} \prod_{W \text{ planar tree}} N_W(\mathcal{T}'),$$

where $N_W(\mathcal{T}')$ is the multinomial coefficient counting the number of *distinct* ways of rearranging all the $\mathcal{W}'_{i,j}$ that have tree structure W (but potentially different labellings); by convention, $N_W(\mathcal{T}')$ is set to be 1 when no $\mathcal{W}'_{i,j}$ has tree structure W. Alternatively, we have the formula

$$N(\mathcal{T}') = \frac{\#\mathrm{Sh}(\mathcal{T}'; v_0, \dots, v_p)}{\#\mathrm{Sh}(T; v_0, \dots, v_p)},$$

where we stress that the numerator is the cardinality of a labelled shuffle class, while the denominator is the cardinality of an unlabelled shuffle class. It is clear from both the above formulas that $N(\mathcal{T}')$ is constant on all trees $\mathcal{T}'$ in $\mathrm{Sh}(\mathcal{T}; v_0, \dots, v_p)$; equivalently, $N(\mathcal{T}') = N(\mathcal{T})$. Therefore, (3.8) reads

$$\sum_{T' \in \mathrm{Sh}(T; v_0, \dots, v_p)} \sum_{\substack{(i,\alpha)\text{ labellings} \\ \mathcal{T}'\text{ of }T'}} \mathcal{E}_H(\mathcal{T}') = \sum_{\substack{(i,\alpha)\text{ labellings} \\ \mathcal{T}\text{ of }T}} \frac{1}{N(\mathcal{T})} \sum_{\mathcal{T}' \in \mathrm{Sh}(\mathcal{T}; v_0, \dots, v_p)} \mathcal{E}_H(\mathcal{T}').$$

We now use Lemma 3.6 to conclude the proof of (3.7). □

4 | THE SHUFFLE CONJECTURES

In this section, we prove Theorems 1.4 and 1.6, which state that certain conjectures involving shuffling subtrees of d-Catalan trees imply the Jacobian conjecture.

4.1 | Subtree shuffling conjecture

While our work so far in Sections 2 and 3 has been based on reshaping (and finding new approaches to) known results in the literature, in this section we introduce our new and foundational idea based on the orthogonality of indicators of shuffle classes to the average energy function. This converts the problem from one involving weighted trees to a simpler one involving unweighted trees.

The basic idea, which lies at the heart of our approach, is the following. Recall first that, to prove the Jacobian conjecture, one would need to show that, if JH is nilpotent, the quantity $\Phi_i^\alpha(H)$ defined in (1.2) is zero for large $|\alpha|$. On the other hand, as we showed in Lemma 3.3, the nilpotency condition implies that variants of the sum occurring in (1.2), with the T-sum over a shuffle class $\mathcal{S} \subseteq C_k^{(d)}$ rather than the whole $C_k^{(d)}$, are zero. Putting the two observations together, we conclude: *if* the constant function can be written as a linear combination of indicators of shuffle classes $\mathcal{S}$, then $\Phi_i^\alpha(H) = 0$, as desired.

To make the above argument rigorous, let us define an inner product on the set of functions $(C_k^{(d)})^* := \{f : C_k^{(d)} \to \mathbb{C}\}$ by setting

$$\langle f, g\rangle := \sum_{T \in C_k^{(d)}} \overline{f(T)} g(T).$$

Given integers $k \geqslant 1$ and $d \geqslant 2$, the average H-weight function $E_{i,\alpha,H}$, defined in (3.3), is an element of $(C_k^{(d)})^*$. With this at hand, (1.2) reads

$$\Phi_i^\alpha(H) = \frac{1}{(d!)^k} \langle \mathbf{1}, E_{i,\alpha,H}\rangle, \tag{4.1}$$

where $\mathbf{1} \in (C_k^{(d)})^*$ is the constant function defined by $\mathbf{1}(T) := 1$ for every T.

Consider now the indicator function of a shuffle class $\mathcal{S}$ of length p, that is, the element $\mathbf{1}_{\mathcal{S}}$ of $(C_k^{(d)})^*$ defined by

$$\mathbf{1}_{\mathcal{S}}(T) := \begin{cases} 1 & \text{if } T \in \mathcal{S}, \\ 0 & \text{if } T \notin \mathcal{S}. \end{cases}$$

The (shuffle) Lemma 3.3 now reads

$$(JH)^p = 0 \quad \Longrightarrow \quad \langle 1_{\mathcal{S}}, E_{i,\alpha,H}\rangle = 0 \text{ for every length-}p \text{ shuffle class } \mathcal{S} \subseteq C_k^{(d)}. \tag{4.2}$$

We are now able to show that the subtree shuffling conjecture posed in the introduction implies the Jacobian conjecture.

Proof of Theorem 1.4. Fix $d \geqslant 3$. Suppose that, for all $p \geqslant 1$, there exists some $k_0(d,p)$ such that, for all $k \geqslant k_0(d,p)$, $\mathbf{1} = \sum_{\mathcal{S}} \lambda_{\mathcal{S}} \mathbf{1}_{\mathcal{S}}$ can be written as a linear combination of the length-p shuffle indicators, for some constants

$$\{\lambda_{\mathcal{S}} : \mathcal{S} \text{ is a length-}p \text{ shuffle class in } C_k^{(d)}\}.$$

Then, using the conjugate linearity of $\langle \cdot, \cdot \rangle$ in the first slot, we have

$$\langle \mathbf{1}, E_{i,\alpha,H}\rangle = \sum_{\mathcal{S}} \overline{\lambda_{\mathcal{S}}} \langle \mathbf{1}_{\mathcal{S}}, E_{i,\alpha,H}\rangle.$$

If $(JH)^p = 0$, then, according to (4.2), we have $\langle \mathbf{1}_{\mathcal{S}}, E_{i,\alpha,H}\rangle = 0$ for every length-p shuffle class $\mathcal{S}$. It follows from (4.1) that $\Phi_i^\alpha(H) = 0$ for all $i \in [n]$ and $|\alpha| = (d-1)k + 1$.

Set now $p = n$. In light of (1.2), we have that, whenever H is a homogeneous degree d polynomial mapping such that JH is nilpotent (i.e. with $(JH)^n = 0$), the coefficient $g_{i,\alpha}$ of X^α in the ith component of the inverse F^{-1} of $F = I - H$ is zero, for all α such that $|\alpha| \geqslant (d-1)k_0(d,n) + 1$. In particular, F^{-1} is a polynomial mapping. By the Bass–Connell–Wright reduction (Theorem 3.1), this implies that the Jacobian conjecture is true. □

While Theorem 1.4 states that Conjecture 1.3 implies the Jacobian conjecture, we are not sure about the reverse implication, that is, whether or not the Jacobian conjecture implies Conjecture 1.3. Nonetheless, we believe that Conjecture 1.3 is a valuable line of investigation, since it represents a dramatic simplification of the combinatorial formulation of the Jacobian conjecture set out in Conjecture 1.2: the latter involves complicated weighted sums over labellings of trees, whereas the former deals only with unweighted trees.

Furthermore, as we have already stated in the introduction and will prove in the next sections, it turns out that an approximate version of the (subtree shuffling) Conjecture 1.3 is true.

To conclude our discussion in this section, it is worth comparing the subtree shuffling conjecture with the following conjecture, posed by Singer in the $d = 2$ setting.

Conjecture 4.1 (Singer [23]). *Let $d \geqslant 2$, $p \geqslant 1$. Then there exists $k_1(d,p)$ such that, whenever $k \geqslant k_1(d,p)$, the indicator functions of the length-p shuffle classes in $C_k^{(d)}$ span the vector space of functions $C_k^{(d)} \to \mathbb{C}$.*

Singer arrived at the latter while attempting to provide a combinatorial derivation of (Wang's) Theorem 1.9. He used an innovative argument to prove Conjecture 4.1 in the cases $p = 1, 2, 3$. In the case $p = 3$, for instance, he proved that taking $k \geqslant 7$ is sufficient, thereby deducing that $F = I - H$ with H quadratic and $(JH)^3 = 0$ has an inverse of degree at most 6, *regardless of the underlying dimension* n.

We remark that the $d = 2$ case of the (subtree shuffling) Conjecture 1.3 is substantially weaker than (Singer's) Conjecture 4.1. Conjecture 1.3 states that the constant function lies in the span of the shuffle indicators, while Conjecture 4.1 states that *all* functions lie in the span of the shuffle indicators.

4.2 | Shuffle chain conjecture

We now discuss how to reformulate the machinery of Section 4.1 in terms of Markov chains on trees, with the final goal of proving Theorem 1.6 from the introduction. We believe this formulation potentially offers a tractable path to proving Conjecture 1.3, which, as we saw in Section 4.1, implies the Jacobian conjecture.

In this section, we consider a Markov chain that takes values in $C_k^{(d)}$, and updates its state by randomly shuffling the subtrees along a randomly chosen path.

To define this process explicitly, recall that, for each vertex v of height at least p in a d-Catalan tree, there is a unique path $(v_0, \dots, v_p)$ of length p leading to $v = v_p$. In this section, we write $\mathrm{Sh}_{T,p}(v) := \mathrm{Sh}(T; v_0, \dots, v_p)$ as a shorthand for the shuffle class associated with this path. We also denote by $V(T)$ the set of vertices of a tree T.

For each T in $C_k^{(d)}$, consider now a probability measure P_T supported on the set of vertices of height at least p in T, that is, a [0,1]-valued function on the vertices v of T such that

$$\sum_{v \in V(T)} P_T(v) = 1,$$

and $P_T(v) > 0$ if and only if v has height at least p. By the correspondence between vertices of T of height at least p and length-p shuffle classes containing T, P_T may also be regarded as a probability measure on the length-p shuffle classes containing T.

Definition 4.2. Fix a collection $\{P_T : T \in C_k^{(d)}\}$, where each P_T is a probability measure supported on the set of vertices of T of height at least p. The p-**shuffle chain** on $C_k^{(d)}$ associated with $\{P_T : T \in C_k^{(d)}\}$ is the Markov chain $(T_j)_{j \geqslant 0}$ with state space $C_k^{(d)}$ and transition probabilities

$$\mathbb{P}\big(T_{j+1} = T' \,\big|\, T_j = T\big) = \sum_{v \in V(T)} P_T(v) \frac{\mathbf{1}\{T' \in \mathrm{Sh}_{T,p}(v)\}}{\#\mathrm{Sh}_{T,p}(v)}. \tag{4.3}$$

In other words, conditional on the event $\{T_j = T\}$, in order to obtain T_{j+1} we first choose a random vertex v of T according to $P_T(v)$, and then uniformly choose a tree $T_{j+1} = T'$ from those in $\mathrm{Sh}_{T,p}(v)$.

Let us remark that choosing a tree T' uniformly from those in $\mathrm{Sh}_{T,p}(v)$ is equivalent to taking a uniform permutation σ of the $(d-1)p$ subtrees in the path of length p leading up to v.

It is clear that, in the p-shuffle chain, communicating classes are closed: if there is a positive probability of going from T to T' in a certain number of steps, then there is also a positive probability of going from T' to T by just reversing these steps. Furthermore, for fixed d and p and sufficiently large k, every state in a p-shuffle chain communicates with any other state. Therefore, for large enough k, any p-shuffle chain is an irreducible Markov chain on a finite state space and, as such, it has a unique stationary distribution.

The framework of Section 4.1 allows us now to show that the shuffle chain conjecture posed in the introduction implies the Jacobian conjecture.

Proof of Theorem 1.6. We first prove that Conjecture 1.5 implies Conjecture 1.3, and then appeal to Theorem 1.4.

Let $d \geqslant 3$. Suppose that, for all $p \geqslant 1$, there exists some $k_0(d,p)$ such that, for all $k \geqslant k_0(d,p)$, there exist a collection of probability measures $\{P_T : T \in C_k^{(d)}\}$ with the property that the associated p-shuffle chain has the uniform distribution as its stationary distribution. Then, for every tree $T' \in C_k^{(d)}$, we have

$$\frac{1}{\# C_k^{(d)}} = \sum_{T \in C_k^{(d)}} \frac{1}{\# C_k^{(d)}} \mathbb{P}\big(T_{j+1} = T' \,\big|\, T_j = T\big). \tag{4.4}$$

Multiplying (4.4) by $\# C_k^{(d)}$ and using the transition probabilities (4.3), we obtain

$$1 = \sum_{T \in C_k^{(d)}} \sum_{v \in V(T)} P_T(v) \frac{\mathbf{1}\{T' \in \mathrm{Sh}_{T,p}(v)\}}{\# \mathrm{Sh}_{T,p}(v)}. \tag{4.5}$$

Given a shuffle class $\mathcal{S} \subseteq C_k^{(d)}$, define the constant

$$\lambda_{\mathcal{S}} := \sum_{T \in C_k^{(d)}} \sum_{v \in V(T)} \frac{P_T(v)}{\# \mathrm{Sh}_{T,p}(v)} \mathbf{1}\{\mathrm{Sh}_{T,p}(v) = \mathcal{S}\}.$$

Then, (4.5) reads

$$1 = \sum_{\mathcal{S}} \lambda_{\mathcal{S}} \mathbf{1}_{\mathcal{S}}(T')$$

for all T' in $C_k^{(d)}$. In other words, the constant function $\mathbf{1} : C_k^{(d)} \to \mathbb{C}$ lies in the span of the length-p shuffle indicators.

In particular, if, for some $d \geqslant 3$ and all $p \geqslant 1$, Conjecture 1.5 is true, then so is Conjecture 1.3, and, by virtue of Theorem 1.4, so is the Jacobian conjecture. □

Conjecture 1.5 is, strictly speaking, stronger than Conjecture 1.3, and corresponds to a special case of Conjecture 1.3 in which all of the weights $\lambda_{\mathcal{S}}$ are non-negative. However, we believe it is a valuable reformulation, as it recasts the problem in terms of Markov chains on trees, for which there is a rich literature (see, e.g. [4, 6, 31, 36]).

5 | PERFECT TREES

In Section 1, we briefly introduced the notion of p-perfect tree.

Definition 5.1. A d-Catalan tree T is p-**perfect** if it contains a path of vertices $(v_0, \dots, v_p)$, as in Definition 3.2, such that each of the $(d-1)p$ siblings of the vertices $v_1, \dots, v_p$ is a leaf.

See Figure 4 for an illustration.

The main aim of this section is to prove that a large uniform d-Catalan tree is extremely likely to be p-perfect (Theorem 1.7). We then exploit this crucial fact to prove:

1. the ℓ^1 approximation of the constant function in Theorem 1.8, which can be regarded as an approximate version of Conjecture 1.3;
2. certain quantitative estimates for the coefficients of the inverse of a Keller mapping, which can be regarded as an approximate version of the Jacobian conjecture itself.

In recent decades, there has been a large body of literature on the local structure of random trees. The interested reader might consult work by Aldous [5], Janson [14], Holmgren and Janson [13] and in particular Janson's survey article [15]. Nonetheless, it appears the exact question we study (namely, estimating the probability that a large uniform d-Catalan tree is p-perfect) has not been examined in the literature. Furthermore, for our statements we require absolute bounds on the probabilities involved, rather than convergence in distribution, hence we develop the requisite estimates from scratch.

5.1 | Uniform tree algorithm

We start by collecting some basic facts about d-Catalan numbers. Denote by $C_k^{(d)} := \#\mathcal{C}_k^{(d)}$ the d-**Catalan number**, that is, the number of d-Catalan trees with k internal vertices (and, thus, $dk+1$ vertices in total). It is a combinatorial exercise (see, e.g. [10, Section 7.5]) to prove that these numbers possess the closed-form expression

$$C_k^{(d)} = \frac{1}{(d-1)k+1}\binom{dk}{k}. \tag{5.1}$$

Using Stirling's formula, it is easily verified that

$$C_k^{(d)} = (1+o(1))C_d k^{-3/2} \mathrm{e}^{A_d k} \qquad \text{as } k \to \infty, \tag{5.2}$$

where $C_d := \frac{1}{\sqrt{2\pi}} \frac{\sqrt{d}}{(d-1)^{3/2}}$ and $A_d := d \log d - (d-1)\log(d-1)$.

The d-Catalan numbers also satisfy the recursion

$$C_k^{(d)} = \sum_{\substack{k_1, \dots, k_d \geqslant 0 : \\ k_1 + \dots + k_d = k-1}} C_{k_1}^{(d)} \cdots C_{k_d}^{(d)}.$$

This immediately sets up a way of generating a uniform Catalan tree in $\mathcal{C}_k^{(d)}$:

- If $k = 0$, make the root a leaf, thus obtaining the unique d-Catalan tree with 0 internal vertices. If $k \geqslant 1$, sample a random vector $(k_1, \dots, k_d)$, satisfying $k_1, \dots, k_d \geqslant 0$ and $k_1 + \cdots + k_d = k - 1$, from the probability distribution

$$P_k^{(d)}(k_1, \dots, k_d) := \frac{C_{k_1}^{(d)} \cdots C_{k_d}^{(d)}}{C_k^{(d)}}. \tag{5.3}$$

- For all $i = 1, \dots, d$, tag the ith child (going left-to-right) of the root with the integer k_i. If the ith child of the root is tagged with $k_i = 0$, make that child a leaf. If the ith child is tagged with $k_i > 0$, then sample a new random vector $(k_{i,1}, \dots, k_{i,d})$, satisfying $k_{i,1}, \dots, k_{i,d} \geqslant 0$ and $k_{i,1} + \cdots + k_{i,d} = k_i - 1$, from the probability distribution

$$P_{k_i}^{(d)}(k_{i,1}, \dots, k_{i,d}) := \frac{C_{k_{i,1}}^{(d)} \cdots C_{k_{i,d}}^{(d)}}{C_{k_i}^{(d)}}.$$

- Repeat the procedure with all the vertices in the second generation, and so on.

We remark that this procedure for generating a uniform Catalan tree is essentially a (planar) special case of a Gibbs fragmentation tree [20].

5.2 | Likelihood of p-perfect trees

Let us now sketch how one can utilise the algorithm of Section 5.1 to provide a description of the local structure of a d-Catalan tree at the root. As in Section 1, let $\mathbb{P}_k^{(d)}$ be the uniform probability distribution on $\mathcal{C}_k^{(d)}$. Sample T from $\mathbb{P}_k^{(d)}$ and let $K_1, \dots, K_d$ denote the respective numbers of internal vertices in the d subtrees subtended by the d children of the root. The random vector $K := (K_1, \dots, K_d)$ is then distributed according to (5.3). Now write $(X_1, \dots, X_{d-1})$ for the vector K with the largest coordinate removed (if there are joint largest coordinates, remove any largest coordinate). Whenever $x_1 + \cdots + x_{d-1} < k/2$ (in which case there certainly were not joint largest coordinates), we have

$$\mathbb{P}_k^{(d)}((X_1, \dots, X_{d-1}) = (x_1, \dots, x_{d-1})) = dC_{x_1}^{(d)} \cdots C_{x_{d-1}}^{(d)} \frac{C_{k-(x_1+\cdots+x_{d-1}+1)}^{(d)}}{C_k^{(d)}},$$

with the factor d in front accounting for the fact that any given coordinate of K is equally likely to be the largest (on the event that there is a unique largest coordinate). For fixed $(x_1, \dots, x_{d-1})$, using (5.2) we see that

$$\begin{aligned} Q_\infty(x_1, \dots, x_{d-1}) &:= \lim_{k\to\infty} \mathbb{P}_k^{(d)}((X_1, \dots, X_{d-1}) = (x_1, \dots, x_{d-1})) \\ &= dC_{x_1}^{(d)} \cdots C_{x_{d-1}}^{(d)} \mathrm{e}^{-A_d(x_1+\cdots+x_{d-1}+1)}. \end{aligned} \tag{5.4}$$

In other words, if we pick a uniform Catalan tree from $C_k^{(d)}$, when k is very large, we expect that, of the d children subtended by the root, the subtrees subtended by exactly $d-1$ of these children usually have size $O(1)$, and one of the d children subtends a subtree of size $k - O(1)$. The distribution of the sizes of the $d-1$ order 1 subtrees is given by (5.4).

Let us remark that such a local convergence holds in much more generality: see, for example, work by Abraham and Delmas [2] or Stufler [27].

We hope this sketch calculation supplies the reader with an idea of the local behaviour of large d-Catalan trees. Let us note in particular that the probability that, in a giant uniform d-Catalan tree (i.e. in the limit as $k \to \infty$), all but one of the children of the root is a leaf, is given by

$$Q_\infty(0, \dots, 0) = d\mathrm{e}^{-A_d} = \left(1 + \frac{1}{d-1}\right)^{-(d-1)} \geqslant 1/e$$

for all $d \geqslant 2$. In turns out that the latter bound is in fact absolute across all d and across uniform Catalan trees with any number of vertices:

Lemma 5.2. *For all $d \geqslant 2$ and $k \geqslant 1$, the probability that at least $d-1$ of the children of the root of a uniformly chosen d-Catalan tree with k internal vertices are leaves is at least $1/e$.*

Proof. If $k = 1$, then all of the children are leaves with probability 1, and we are done.

We may then suppose $k \geqslant 2$. We have

$$\mathbb{P}_k^{(d)}(\text{at least } d-1 \text{ of the rootchildren are leaves}) = dP_k^{(d)}(0, 0, \dots, 0, k-1),$$

where the factor of d comes from the fact that there are d different ways of choosing which of the d children of the root is going to be the non-leaf.

Note that $C_0^{(d)} = 1$. Using (5.3) and (5.1), we then have

$$\mathbb{P}_k^{(d)}(\text{at least } d-1 \text{ of the rootchildren are leaves}) = d\frac{C_{k-1}^{(d)}}{C_k^{(d)}} = d\frac{(d-1)k+1}{(d-1)(k-1)+1}\frac{\binom{d(k-1)}{(k-1)}}{\binom{dk}{k}},$$

which implies

$$\mathbb{P}_k^{(d)}(\text{at least } d-1 \text{ of the rootchildren are leaves}) \geqslant d\frac{\binom{d(k-1)}{k-1}}{\binom{dk}{k}}.$$

Expanding the binomial coefficients and rearranging, we see that

$$d\frac{\binom{d(k-1)}{k-1}}{\binom{dk}{k}} = \frac{(dk-k)!}{((dk-k)-(d-1))!}\frac{((dk-1)-(d-1))!}{(dk-1)!} = \prod_{i=1}^{d-1}\frac{dk-k-i+1}{dk-i}.$$

Noticing that the function $x \mapsto \frac{x-a}{x}$ is increasing on $(0, \infty)$ for any fixed $a > 0$, we deduce

$$\mathbb{P}_k^{(d)}(\text{At least } d-1 \text{ of the rootchildren are leaves}) \geqslant \left(\frac{dk-k-d+2}{dk-d+1}\right)^{d-1}.$$

Now note that $\frac{dk-k-d+2}{dk-d+1} = 1 - \frac{(k-1)}{d(k-1)+1} \geqslant 1 - 1/d$. It follows that

$$\mathbb{P}_k^{(d)}(\text{At least } d-1 \text{ of the rootchildren are leaves}) \geqslant (1-1/d)^{d-1}.$$

For $d \geqslant 2$, we have $(1-1/d)^{d-1} = (1+1/(d-1))^{-(d-1)} \geqslant 1/e$, completing the proof. □

Corollary 5.3. *Let $k > p$. Then the $\mathbb{P}_k^{(d)}$-probability that there is at most one internal vertex in each of the generations 1 through p is at least e^{-p}.*

Proof. Use the uniform tree algorithm and Lemma 5.2 p times in a row. □

We now estimate the probability that T is p-perfect. Let T be any tree in $C_k^{(d)}$. For $d \geqslant 2$, T has at most $1 + d + \cdots + d^{p-1} \leqslant d^p$ vertices in its generations 0 through $p-1$. Let $v_1, \ldots, v_r$ be a list of the vertices in generation p of the tree, and let $n_1, \ldots, n_r$ denote the number of internal vertices in the trees they subtend. Then

$$n_1 + \cdots + n_r \geqslant k - d^p. \tag{5.5}$$

If we let $Q_{p,d}(k)$ be the probability that a uniformly chosen d-Catalan tree with k internal vertices is not p-perfect. To obtain an estimate on $Q_{p,d}(k)$, note that we have

$$\begin{aligned} Q_{p,d}(k) \leqslant\ & \mathbb{P}_k^{(d)}(\{T \text{ does not have a } p\text{-perfect path starting at the root}\} \\ & \cap \{T \text{ does not have a } p\text{-perfect path starting at generation } p\}). \end{aligned}$$

By Corollary 5.3, the probability of the first event is bounded above by $1 - e^{-p} \leqslant e^{-e^{-p}}$. By the uniform tree algorithm, the probability of the second event, conditionally on r and $n_1, \ldots, n_r$, is bounded above by $\prod_{i=1}^r Q_{p,d}(n_i)$. Therefore, for any $k > p$ we have

$$Q_{p,d}(k) \leqslant e^{-e^{-p}} \sup_{\substack{r \leqslant d^p, \\ n_1+\cdots+n_r \geqslant k-d^p}} \prod_{i=1}^r Q_{p,d}(n_i), \tag{5.6}$$

where the supremum is taken over all $r \leqslant d^p$ (the number of vertices r in generation p of the tree satisfies $r \leqslant d^p$) and r-tuplets of non-negative integers $n_1, \ldots, n_r$ adding up to at least $k - d^p$ (by (5.5)). Using this inequality, we are now ready to prove Theorem 1.7.

Proof of Theorem 1.7. We need to show that $Q_{p,d}(k) \leqslant e^{-\kappa_{p,d}(k-p)_+}$ for all $k \in \mathbb{N}$. The result is clearly true for all $k \leqslant p$, because $Q_{p,d}(k)$ is a probability, and hence certainly $\leqslant 1$. We now prove the result by induction for $k > p$.

Suppose that $Q_{p,d}(m) \leqslant e^{-\kappa_{p,d}(m-p)_+}$ for all $m < k$. Then, using (5.6), we have

$$Q_{p,d}(k) \leqslant \sup_{\substack{r \leqslant d^p, \\ n_1+\cdots+n_r \geqslant k-d^p}} e^{-e^{-p} - \kappa_{p,d} \sum_{i=1}^r (n_i - p)_+}.$$

Using the simple fact that $(n_i - p)_+ \geqslant n_i - p$ and the constraints, we obtain

$$Q_{p,d}(k) \leqslant \exp\left\{-\mathrm{e}^{-p} - \kappa_{p,d}(k - d^p - d^p p)\right\}.$$

To prove the result, it is now sufficient to show that

$$\exp\left\{-\mathrm{e}^{-p} - \kappa_{p,d}(k - (p+1)d^p)\right\} \leqslant \mathrm{e}^{-\kappa_{p,d}(k-p)},$$

which amounts to the inequality

$$\mathrm{e}^{-p} + \kappa_{p,d}(k - (p+1)d^p) \geqslant \kappa_{p,d}(k-p).$$

The latter is easily verified using the definition of $\kappa_{p,d} := (2pd^p\mathrm{e}^p)^{-1}$. □

Let us remark that, if $(v_0, \dots, v_p)$ is a path of vertices in a tree, then $v_0, \dots, v_{p-1}$ are certainly internal. It follows that, when $p > k$, it is impossible to find a path of length p in a tree with k internal vertices. Therefore, when $p > k$, $C_k^{(d)}$ contains no p-perfect trees (or length-p shuffle classes, for that matter) and, in this sense, the inequality of Theorem 1.7 is sharp.

5.3 | The ℓ^1 bound

The existence of a function in the span of the length-p shuffle indicators satisfying the total variation bound (1.6) in Theorem 1.8 is an immediate consequence of Theorem 1.7, as we now see.

Proof of (1.6). If T is p-perfect, then T lies in a singleton length-p shuffle class. It follows that the indicator function of any p-perfect tree lies in the span of the length-p shuffle indicators. Therefore, the function

$$\phi(T) := \mathbf{1}\{T \text{ is } p\text{-perfect}\}$$

also lies in the span of the length-p shuffle indicators. Recall that $\mathbb{E}_k^{(d)}$ denotes the expectation with respect to $\mathbb{P}_k^{(d)}$. It then follows from Theorem 1.7 that

$$||\phi - \mathbf{1}||_1 := \mathbb{E}_k^{(d)}|1 - \phi(T)| = \mathbb{P}_k^{(d)}(T \text{ is not } p\text{-perfect}) \leqslant \mathrm{e}^{-\kappa_{p,d}(k-p)},$$

as desired. □

The full proof of Theorem 1.8, that is, the exhibition of a function ϕ_k satisfying both (1.6) and (1.7), can be found in Section 6.

5.4 | Quantitative estimates for inverse coefficients of Keller mappings

The ℓ^1 bound in Theorem 1.8 yields a bound for the inverse coefficients $g_{i,\alpha}$ from (1.2) of maps F of the form (1.1). More precisely, we will prove that the coefficients $g_{i,\alpha}$ satisfy a sharp bound, which, in the case where F is a Keller mapping, can be improved upon by an exponential factor.

Theorem 5.4. *Let F be a polynomial map of the form (1.1), where JH may or may not be nilpotent. Let $L := \max_{i,\alpha} |H_{i,\alpha}|$. Then, for all $i \in [n]$ and all multi-indices α of degree $|\alpha| = (d-1)k+1$, we have*

$$|g_{i,\alpha}| \leqslant \frac{1}{(d!)^k} \frac{1}{(d-1)k+1} \binom{dk}{k} n^{k-1} \frac{|\alpha|!}{\alpha!} L^k. \tag{5.7}$$

This bound is sharp, in the sense that there exists an F such that the equality holds in (5.7) for all i and α. If we additionally assume that JH is nilpotent, with $(JH)^p = 0$ for some $p \in \mathbb{N}$, then for any ϕ lying in $\mathrm{span}\{\mathbf{1}_{\mathcal{S}} :\ \mathcal{S}$ is a length-p shuffle class in $C_k^{(d)}\}$ we have

$$|g_{i,\alpha}| \leqslant \frac{1}{(d!)^k} \frac{1}{(d-1)k+1} \binom{dk}{k} n^{k-1} \frac{|\alpha|!}{\alpha!} L^k \times ||\phi - \mathbf{1}||_1. \tag{5.8}$$

In particular, when $(JH)^p = 0$, we have

$$|g_{i,\alpha}| \leqslant \frac{1}{(d!)^k} \frac{1}{(d-1)k+1} \binom{dk}{k} n^{k-1} \frac{|\alpha|!}{\alpha!} L^k \times C e^{-ck}, \tag{5.9}$$

where C and c are the same constants appearing in (1.6).

Proof. Let H be a degree d homogeneous polynomial mapping, where JH may or may not be nilpotent. Notice that $L := \max_{i,\alpha} |H_{i,\alpha}|$ is finite, since $H_{i,\alpha} = 0$ for all but finitely many i and α. We thus have the simple bound

$$|\mathcal{E}_H(\mathcal{T})| \leqslant L^k \tag{5.10}$$

on the H-weight of a labelling $\mathcal{T}$ of a d-Catalan tree with k internal vertices (see definition (3.2)). In general, this bound is sharp, and is attained for any labelled tree $\mathcal{T}$ with k internal vertices by setting $H_{i,\alpha} := L$ for all $i \in [n]$ and $|\alpha| = d$.

Given $\alpha \in \mathbb{Z}^n_{\geqslant 0}$ and a d-Catalan tree T with $|\alpha| = (d-1)k+1$ leaves, there are

$$\#\{(i,\alpha) \text{ labellings of } T\} = n^{k-1} \frac{|\alpha|!}{\alpha!} \tag{5.11}$$

ways of (i,α)-labelling T. The factor n^{k-1} is due to the number of ways of labelling the $k-1$ non-root internal vertices with any type in $[n]$, while $|\alpha|!/(\alpha_1! \cdots \alpha_n!)$ is the multinomial coefficient counting the distinct ways of labelling the leaves with labels in $[n]$ such that exactly α_j leaves are given label j. We remark that the number of (i,α) labellings does not depend on the particular d-Catalan tree T, but only on its total number of vertices/leaves.

Using (5.11) and (5.10) in (3.3), we obtain the bound

$$|E_{i,\alpha,H}(T)| \leqslant n^{k-1} \frac{|\alpha|!}{\alpha!} L^k \tag{5.12}$$

for all d-Catalan tree T with $|\alpha| = (d-1)k+1$ leaves. Using the bound (5.12) in (1.2) in conjunction with d-Catalan number formula (5.1), we obtain

$$|g_{i,\alpha}| \leqslant \frac{1}{(d!)^k} \frac{1}{(d-1)k+1} \binom{dk}{k} n^{k-1} \frac{|\alpha|!}{\alpha!} L^k, \tag{5.13}$$

which accounts for the first bound (5.7). The sharpness of (5.13) (and (5.12)) may be seen by setting again $H_{i,\alpha} := L$ for all i, α; one can verify that, for such a choice of H, JH is *not* nilpotent.

Let us now suppose that JH is nilpotent, with $(JH)^p = 0$, and turn to the proof of (5.8). Similarly as in the proof of Theorem 1.4, one can see that, whenever $(JH)^p = 0$, in light of (4.2), any function ϕ lying in the span of the length-p shuffle indicators satisfies $\langle \phi, E_{i,\alpha,H} \rangle = 0$. Accordingly, we have

$$\Phi_i^\alpha(H) = \frac{1}{(d!)^k} \langle \mathbf{1}, E_{i,\alpha,H} \rangle = \frac{1}{(d!)^k} \langle \mathbf{1} - \phi, E_{i,\alpha,H} \rangle.$$

In particular, using (5.12), for any ϕ in the span of the length-p shuffle indicators we have the total variation bound

$$\begin{aligned} |\Phi_i^\alpha(H)| &\leqslant \frac{1}{(d!)^k} n^{k-1} \frac{|\alpha|!}{\alpha!} L^k \sum_{T \in C_k^{(d)}} |1 - \phi(T)| \\ &= \frac{1}{(d!)^k} \frac{1}{(d-1)k+1} \binom{dk}{k} n^{k-1} \frac{|\alpha|!}{\alpha!} L^k ||\phi - \mathbf{1}||_1, \end{aligned}$$

which completes the proof of (5.8). Finally, (5.9) follows immediately from (5.8) and (1.6). □

Let us take a moment to elaborate on this result. First, we stress that the bound (5.7) concerns the inverse coefficients of *any* degree d homogeneous polynomial mapping and is not a priori concerned with the Jacobian conjecture. We believe this bound, which is also new to the best of our knowledge, might be of independent interest.

As discussed in Section 3, Bass, Connell and Wright showed in [8] that, in order to prove the Jacobian conjecture, it is sufficient to consider maps $F = I - H$, with H homogeneous of degree d and JH nilpotent (see Theorem 3.1). Indeed, the nilpotency condition is equivalent to the Keller condition for such maps (see [17, Section 2.5]).

We remark that both bounds in (5.7) and (5.9) grow exponentially in k. Theorem 5.4 may be regarded as an approximate version of the Jacobian conjecture: the inverses of Keller maps are closer to being polynomial than their non-Keller counterparts, in that the coefficients of their high-degree terms are smaller by an exponential factor.

Finally, we note that (5.8) implies Theorem 1.4. Indeed, if Conjecture 1.3 is true, then, for large enough k, we can set $\phi = \mathbf{1}$ in (5.8), so that the bound yields $g_{i,\alpha} = 0$ whenever $|\alpha| = (d-1)k+1$.

6 | CONSTRUCTION OF THE ℓ^∞ BOUND

In this section, we construct a function that satisfies the ℓ^∞ bound of Theorem 1.8 and conclude its proof.

6.1 | Indexing of shuffle classes by vertices and a first approximation

Through this section, we will use the notation

$$N_j(T) := \#\{\text{vertices in generation } j \text{ of } T\}$$

for the number of vertices at graph distance j from the root of a tree T. We also write

$$N_{\leqslant q}(T) := \sum_{0\leqslant j\leqslant q} N_j(T) \qquad \text{and} \qquad N_{\geqslant q}(T) := \sum_{j\geqslant q} N_j(T),$$

where the latter sum is clearly finite for any finite tree T.

Recall that, given a tree T and a path of vertices $(v_0, \ldots, v_p)$ in T, we let the shuffle class $\mathrm{Sh}(T; v_0, \ldots, v_p)$ be the set of trees T' that may be obtained from T by shuffling the subtrees subtended by the siblings of $v_1, \ldots, v_p$. Consider now examining the set of shuffle classes containing a given tree T. Each vertex v lying at a height of at least p in T may be associated with a canonical path of $(v_0, \ldots, v_p)$ of vertices in T by letting $v_p := v$ and $v_{p-(i+1)}$ be the unique parent of v_{p-i}, for each $0 \leqslant i \leqslant p-1$. Therefore, vertices of height at least p are in bijection with paths of length p. In particular, for each tree T, we might informally expect in a suitable sense that

$$\#\{\text{length-}p \text{ shuffle classes containing } T\} \approx N_{\geqslant p}(T). \tag{6.1}$$

We need to be careful however: it is entirely possible that two different choices v and v' of vertices of height at least p, and their associated paths, give rise to the *same* shuffle class. As a simple example, consider a tree T containing two different perfect p-paths $(v_0, \ldots, v_p)$ and $(v'_0, \ldots, v'_p)$: then, both paths (or equivalently, both vertices $v = v_p$ and $v' = v'_p$) give rise to the same shuffle class, that is, the singleton $\{T\}$.

It is however possible to make precise sense of (6.1) after taking on a suitable indexing of the shuffle classes. To do so, we begin by noting that each vertex v in a d-Catalan tree may be associated with a code in $\cup_{j\geqslant 0}\{1, \ldots, d\}^j$ (this is often called Ulam–Harris labelling [12] or Neveu formalism for trees [22]): the root is given by the empty code, the d children of the root, listed from left to right, are given by the 1-tuples $(1), \ldots, (d)$, and, more generally, the children of an internal vertex with code u are given by the concatenations $(u1), \ldots, (ud)$. The length of the code of a vertex is the same as its height in the tree. A d-Catalan tree itself may, as such, be regarded as a subset of $\cup_{j\geqslant 0}\{1, \ldots, d\}^j$ such that, if $(ui) \in T$ for some code u, then $(u) \in T$ and $(uj) \in T$ for all $j = 1, \ldots, d$. Let $(v_0, \ldots, v_p)$ be a path in a d-Catalan tree T, which we may then associate with a collection of codes of lengths $j, \ldots, j+p$ for some $j \geqslant 0$. We note that every tree T' in the shuffle class $\mathrm{Sh}(T; v_0, \ldots, v_p)$ contains this collection of codes.

We now define a function $\tilde{\psi}$ lying in the span of the indicator functions of the length-p shuffle classes by letting

$$\tilde{\psi} := \sum_{j\geqslant p} \sum_{v\in\{1,\ldots,d\}^j} \sum_{\substack{S=\mathrm{Sh}(T;v_0,\ldots,v_p=v) \\ \text{for some } T\in C_k^{(d)}}} \mathbf{1}_S. \tag{6.2}$$

We now note that

$$\tilde{\psi}(T) = N_{\geqslant p}(T). \tag{6.3}$$

Indeed, for every tree T, every vertex v in T of height at least p (which has a unique code of length at least p) scores a 1 in the sum $\tilde{\psi}$. It follows that (6.2) and (6.3) make sense of (6.1).

In the following, our goal is to show that, when k is large compared to p, the vast majority of vertices in any tree T with k internal vertices have height at least p. In fact, we can be far more specific: every tree T in $C_k^{(d)}$ has at most $1 + d + \cdots + d^{p-1}$ vertices in its generations 0 through $p-1$, hence

$$N_{\geqslant p}(T) \geqslant (dk+1) - (1 + d + \cdots + d^{p-1}). \tag{6.4}$$

In particular, we are equipped to give the following preliminary construction in the direction of the uniform bound of Theorem 1.8.

Proposition 6.1. *Define*

$$\psi := \frac{1}{dk+1} \sum_{j \geqslant p} \sum_{v \in \{1,\dots,d\}^j} \sum_{\substack{S = \mathrm{Sh}(T; v_0, \dots, v_p = v) \\ \text{for some } T \in C_k^{(d)}}} \mathbf{1}_S.$$

Then ψ lies in the span of the length-p shuffle indicators, and satisfies the uniform bound

$$\sup_{T \in C_k^{(d)}} |\psi(T) - 1| \leqslant C_{d,p}/(dk+1), \tag{6.5}$$

where we take $C_{d,p} := (d^p - 1)(d-1)^{-1}$.

Proof. Clearly ψ is a scalar multiple of $\tilde{\psi} = N_{\geqslant p}$, and hence lies in the span of the length-p shuffle indicators. The bound (6.5) then follows from (6.4). □

The previous result exhibits a function ϕ in the span of the length-p shuffle indicators, for which the supremum norm $||\phi - \mathbf{1}||_\infty$ decays like $1/k$. In the next section, we develop a more sophisticated approach to find a function ϕ in the span of the length-p shuffle indicators for which $||\phi - \mathbf{1}||_\infty$ is bounded superpolynomially in $1/k$.

6.2 | Construction of the width product function

We refine the ideas of Section 6.1 with the following lemma.

Lemma 6.2. *Let $\chi : C_k^{(d)} \to \mathbb{R}$ be a function that only depends on generations $0, 1, \dots, j$ of a d-Catalan tree; namely, if T and T' have the same generations $0, \dots, j$, then $\chi(T) = \chi(T')$. Then, the function $T \mapsto N_{j+p}(T)\chi(T)$ lies in the span of the length-p shuffle indicators.*

Proof. Let v be a vertex of height $j+p$ in a tree T, and consider the path $(v_0, \dots, v_p)$, with $v_p = v$, and the associated shuffle class $S := \mathrm{Sh}(T; v_0, \dots, v_p)$. Then, v_0 has height j. We note that shuffling the subtrees off the path $(v_0, \dots, v_p)$ leaves generations $0, \dots, j$ of the tree unaffected and, accordingly, χ is constant on S. In particular, the function $\mathbf{1}_S(T)\chi(T)$ lies in the span of the length-p shuffle indicators. Summing over all such shuffle classes obtained this way, we obtain the function $N_{j+p}(T)\chi(T)$. □

Corollary 6.3. *For each integer $m \geqslant 1$, the function*

$$J_m(T) := N_{\geqslant m(p-1)+1}(T) \prod_{i=1}^{m-1} N_{\leqslant i(p-1)}(T)$$

lies in the span of the length-p shuffle indicators.

Proof. Let $\chi(T) := \prod_{i=1}^{m-1} N_{\leqslant i(p-1)}(T)$. Then, by Lemma 6.2, $N_{j+p}(T)\chi(T)$ lies in the span of the length-p shuffle indicators whenever $j \geqslant (m-1)(p-1)$. Equivalently, $N_k(T)\chi(T)$ lies in the span of the shuffle indicators for all $k \geqslant m(p-1)+1$. Summing over all such k, it follows that

$$\sum_{k \geqslant m(p-1)+1} N_k(T)\chi(T) = N_{\geqslant m(p-1)+1}(T) \prod_{i=1}^{m-1} N_{\leqslant i(p-1)}(T)$$

lies in the span of the length-p shuffle indicators. □

Note that, since each tree in $C_k^{(d)}$ has $dk+1$ vertices, we have the crude bound

$$J_m(T) \leqslant (dk+1)^m \mathbf{1}\{\text{the height of } T \text{ is at least } m(p-1)+1\}.$$

Since a tree in $C_k^{(d)}$ has height at most k, $J_m(T)$ is identically zero for large enough m. In particular, we have a well-defined infinite sum

$$\begin{aligned}
\phi(T) &:= \sum_{m=1}^{\infty} \frac{J_m(T)}{(dk+1)^m} \\
&= \sum_{m=1}^{\infty} \frac{(dk+1-N_{\leqslant m(p-1)}) \prod_{i=1}^{m-1} N_{\leqslant i(p-1)}(T)}{(dk+1)^m} \\
&= \sum_{m=1}^{\infty} \left(\prod_{i=1}^{m-1} \frac{N_{\leqslant i(p-1)}(T)}{dk+1} - \prod_{i=1}^{m} \frac{N_{\leqslant i(p-1)}(T)}{dk+1} \right).
\end{aligned}$$

Since the latter is a telescopic sum, we obtain the expression

$$\phi(T) = 1 - \prod_{m=1}^{\infty} \frac{N_{\leqslant m(p-1)}(T)}{dk+1}. \tag{6.6}$$

Again, the product may be taken over finitely many terms: since each tree T in $C_k^{(d)}$ has height at most k, we have that $N_{\leqslant m(p-1)}(T)/(dk+1) = 1$ whenever $m(p-1) \geqslant k$. Notice also that, since all J_m lie in the span of the length-p shuffle indicators by Corollary 6.3, ϕ also lies in the same span. We now provide an upper bound on the quantity $||\phi - \mathbf{1}||_\infty$.

Lemma 6.4. *Let $d \geqslant 2$ and $k \in \mathbb{N}$ such that $k \geqslant d^{6p}$. Then, for every $T \in C_k^{(d)}$, we have $|1-\phi(T)| \leqslant \exp(-c_{d,p}\log^2 k)$, where $c_{d,p} \in (0,\infty)$ may be taken as $(4p\log d)^{-1}$.*

Proof. If T in $C_k^{(d)}$, then $N_j(T) \leqslant d^j$. In particular,

$$N_{\leqslant m(p-1)}(T) = \sum_{j=0}^{m(p-1)} N_j(T) \leqslant \sum_{j=0}^{m(p-1)} d^j = \frac{d^{m(p-1)+1}-1}{d-1} \leqslant d^{mp} \qquad \text{for } d \geqslant 2.$$

Also, trivially we have $N_{\leqslant m(p-1)}(T) \leqslant dk+1$. It follows that

$$|1-\phi(T)| = \prod_{m=1}^{\infty}\left|\frac{N_{\leqslant m(p-1)}(T)}{dk+1}\right| \leqslant \prod_{m=1}^{\infty}\left|\frac{d^{mp}}{dk+1}\wedge 1\right| \leqslant \prod_{m=1}^{\infty}\left|\frac{d^{mp}}{k}\wedge 1\right| = \prod_{m=1}^{m_0}\frac{d^{mp}}{k},$$

where in the final expression above, m_0 denotes the largest integer m such that $d^{mp} \leqslant k$. Then

$$m_0 = \left\lfloor \frac{\log k}{p\log d}\right\rfloor.$$

A calculation tells us that $\prod_{m=1}^{m_0}\frac{d^{mp}}{k} = \exp\left(-m_0\log k + p(\log d)m_0(m_0+1)/2\right)$. Now, using the simple bound $\frac{\log k}{p\log d} - 1 \leqslant m_0 \leqslant \frac{\log k}{p\log d}$, we obtain

$$\begin{aligned}-m_0\log k + p\log d\frac{m_0(m_0+1)}{2} &\leqslant -\frac{\log^2 k}{p\log d} + \log k + \frac{1}{2}\frac{(\log k)^2}{p\log d} + \frac{1}{2}\log k\\ &= -\frac{1}{2}\frac{\log^2 k}{p\log d} + \frac{3}{2}\log k.\end{aligned}$$

Now, since by assumption $\log(k) \geqslant 6p\log(d)$, we arrive at

$$-m_0\log k + p\log d\frac{m_0(m_0+1)}{2} \leqslant -\frac{\log^2 k}{4p\log d},$$

thus completing the proof. □

While Lemma 6.4 provides a uniform upper bound on $|1-\phi(T)|$ on all trees T in $C_k^{(d)}$, we remark that, for the typical tree T in $C_k^{(d)}$, the quantity $|1-\phi(T)|$ is much smaller. Sketching a few of the key steps here, it is known that, when k is large, the typical tree T in $C_k^{(d)}$ has both height and width of order $\sqrt{k}$ (see, e.g. Addario-Berry, Devroye and Janson [3]). Roughly speaking, this entails that

$$|1-\phi(T)| = \prod_{m=1}^{\infty}\frac{N_{\leqslant m(p-1)}(T)}{dk+1} \sim \prod_{m=1}^{O(\sqrt{k}/p)}\frac{O(\sqrt{k})}{dk+1} = \mathrm{e}^{-O\left(\frac{\sqrt{k}}{p}\log k\right)}. \tag{6.7}$$

It is possible to make (6.7) rigorous using the absolute bounds in [3]. We leave the details to the interested reader.

6.3 | Proof of Theorem 1.8

In Section 5, we saw that a proportion of at most $\mathrm{e}^{-\kappa_{d,p}(k-p)_+}$ of the trees in $C_k^{(d)}$ are not p-perfect. So far in this section, we have seen that the function ϕ given by (6.6) has the property that $|1-\phi(T)| \leqslant \mathrm{e}^{-c_{d,p}(\log k)^2}$. We now show it is possible to combine these two results to prove Theorem 1.8.

Proof of Theorem 1.8. We begin by noting that, given *any* function $\rho : C_k^{(d)} \to \mathbb{R}$, the function

$$\rho(T)\mathbf{1}\{T \text{ is } p\text{-perfect}\}$$

lies in the span of the length-p shuffle indicators, since it is a linear combination of the indicator functions of individual p-perfect trees, which lie in singleton shuffle classes.

Therefore, taking ϕ as in (6.6), both $(1-\phi(T))\mathbf{1}\{T \text{ is } p\text{-perfect}\}$ and ϕ lie in the span of the length-p shuffle indicators. In particular, so does the sum of these two functions, which is

$$\phi_*(T) := \mathbf{1}\{T \text{ is } p\text{-perfect}\} + \mathbf{1}\{T \text{ is not } p\text{-perfect}\}\phi(T). \tag{6.8}$$

We clearly have

$$||\phi_*(T) - \mathbf{1}||_\infty \leqslant ||\phi(T) - \mathbf{1}||_\infty \leqslant \mathrm{e}^{-c_{d,p}(\log k)^2} \qquad \text{for } k \geqslant d^{6p}$$

by Lemma 6.4, proving that ϕ_* satisfies (1.7).

Moreover, for all $k \geqslant d^{6p}$, we have

$$\begin{aligned} ||\phi_*(T) - \mathbf{1}||_1 &:= \frac{1}{\#C_k^{(d)}} \sum_{T \in C_k^{(d)}} |1-\phi_*(T)| \\ &= \frac{1}{\#C_k^{(d)}} \sum_{T \in C_k^{(d)} \text{ not } p\text{-perfect}} |1-\phi(T)| \\ &\leqslant \frac{\#\{T \in C_k^{(d)} : T \text{ is not } p\text{-perfect}\}}{\#C_k^{(d)}} \mathrm{e}^{-c_{d,p}(\log k)^2} \\ &\leqslant \mathrm{e}^{-c_{d,p}(\log k)^2 - \kappa_{d,p}(k-p)}. \end{aligned}$$

In the second equality above we used the definition (6.8), in the following inequality we used Lemma 6.4, whereas in the final inequality we used Theorem 1.7. Since $k-p$ diverges to ∞ faster than $(\log k)^2$ as $k \to \infty$, we have that ϕ_* also satisfies the ℓ^1 bound (1.6). □

ACKNOWLEDGEMENTS

This research was supported by the Heilbronn Institute for Mathematical Research (HIMR) as part of the HIMR Focused Research workshop titled: *A graph-theoretic approach to the Jacobian con-*

jecture. Piotr Dyszewski was partially supported by the National Science Centre, Poland (Sonata, Grant number 2020/39/D/ST1/00258). Samuel G. G. Johnston was partially supported in the early stages by the EPSRC funded Project EP/S036202/1 *Random fragmentation-coalescence processes out of equilibrium*. Joscha Prochno was supported by the Austrian Science Fund (FWF) Project P32405 *Asymptotic geometric analysis and applications*. Dominik Schmid acknowledges the DAAD PRIME program for financial support. We thank an anonymous referee for insightful comments.

Open access funding enabled and organized by Projekt DEAL.

JOURNAL INFORMATION

The *Journal of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

REFERENCES

1. A. Abdesselam, *The Jacobian conjecture as a problem of perturbative quantum field theory*, Ann. Henri Poincaré **4** (2003), 199–215.
2. R. Abraham and J.-F. Delmas, *Local limits of conditioned Galton-Watson trees: the infinite spine case*, Electron. J. Probab. **19** (2014), 1–19.
3. L. Addario-Berry, L. Devroye, and S. Janson, *Sub-Gaussian tail bounds for the width and height of conditioned Galton–Watson trees*, Ann. Probab. **41** (2013), no. 2, 1072–1087.
4. D. Aldous, *The random walk construction of uniform spanning trees and uniform labelled trees*, SIAM J. Discrete Math. **3** (1990), no. 4, 450–465.
5. D. Aldous, *Asymptotic fringe distributions for general families of random trees*, Ann. Appl. Probab. **1** (1991), no. 2, 228–266.
6. D. Aldous and J. Pitman, *Tree-valued Markov chains derived from Galton-Watson processes*, Ann. Inst. Henri Poincaré Probab. Stat. **34** (1998), no. 5, 647–686.
7. S. Asmussen and H. Hering, *Branching processes*, Birkhäuser, Boston, MA, 1983.
8. H. Bass, E. H. Connell, and D. Wright, *The Jacobian conjecture: reduction of degree and formal expansion of the inverse*, Bull. Amer. Math. Soc. **7** (1982), no. 2, 287–330.
9. M. de Bondt and A. van den Essen, *A reduction of the Jacobian conjecture to the symmetric case*, Proc. Amer. Math. Soc. **133** (2005), no. 8, 2201–2205.
10. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete mathematics: a foundation for computer science*, Addison-Wesley, 1994.
11. M. Haiman and W. Schmitt, *Incidence algebra antipodes and Lagrange inversion in one and several variables*, J. Combin. Theory Ser. A **50** (1989), no. 2, 172–185.
12. T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Math. Proc. Cambridge Philos. Soc. **56** (1960), no. 1, 13–20.
13. C. Holmgren and S. Janson, *Fringe trees, Crump–Mode–Jagers branching processes and m-ary search trees*, Probab. Surv. **14** (2017), 53–154.
14. S. Janson, *Random cutting and records in deterministic and random trees*, Random Structures Algorithms **29** (2006), no. 2, 139–179.
15. S. Janson, *Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation*, Probab. Surv. **9** (2012), 103–252.
16. S. G. G. Johnston, *A degree bound for strongly nilpotent polynomial automorphisms*, J. Algebra **608** (2022), 259–271.
17. S. G. G. Johnston and J. Prochno, *Faà di Bruno's formula and inversion of power series*, Adv. Math. **395** (2022), 108080.
18. O.-H. Keller, *Ganze Cremona-Transformationen*, Monatsch. Math. Phys. **47** (1939), 299–306.

19. Y. Le Jan and A. S. Sznitman, *Stochastic cascades and 3-dimensional Navier-Stokes equations*, Probab. Theory Related Fields **109** (1997), 343–366.
20. P. McCullagh, J. Pitman, and M. Winkel, *Gibbs fragmentation trees*, Bernoulli **14** (2008), no. 4, 988–1002.
21. G. Miermont, *Invariance principles for spatial multitype Galton-Watson trees*, Ann. Inst. Henri Poincaré Probab. Stat. **44** (2008), no. 6, 1128–1161.
22. J. Neveu, *Arbres et processus de Galton–Watson*, Ann. Inst. Henri Poincaré Probab. Stat. **22** (1986), no. 2, 199–207.
23. D. Singer, *On Catalan trees and the Jacobian conjecture*, Electron. J. Combin. **8** (2001), no. 1, R2.
24. D. Singer, *Toward a combinatorial proof of the Jacobian conjecture!*, Electron. J. Combin. **18** (2011), no. 2, (The Zeilberger Festschrift volume), P27.
25. S. Smale, *Mathematical problems for the next century*, Math. Intelligencer **20** (1998), 7–15.
26. R. Stephenson, *Local convergence of large critical multi-type Galton–Watson trees and applications to random maps*, J. Theoret. Probab. **31** (2018), 159–205.
27. B. Stufler, *Local limits of large Galton–Watson trees rerooted at a random vertex*, Ann. Inst. Henri Poincaré Probab. Stat. **55** (2019), no. 1, 155–183.
28. M. E. Taylor, *Partial differential equations I, Basic theory*, Springer, 1996.
29. A. van den Essen, *Polynomial automorphisms and the Jacobian conjecture*, Birkhäuser Basel, 2000.
30. S. Wang, *A Jacobian criterion for separability*, J. Algebra **65** (1980), no. 2, 453–494.
31. D. B. Wilson, *Generating random spanning trees more quickly than the cover time*, STOC '96: Proceedings of the twenty-eighth annual ACM symposium on Theory of Computing, 1996, pp. 296–303.
32. D. Wright, *Formal inverse expansion and the Jacobian conjecture*, J. Pure. Appl. Algebra **48** (1987), no. 1-2, 199–219.
33. D. Wright, *Tree formulas for reversion of power series*, J. Pure Appl. Algebra **57** (1989), no. 2, 191–211.
34. D. Wright, *The Jacobian conjecture as a problem in combinatorics*, arXiv:math/0511214, 2005.
35. D. Wright, *The Jacobian conjecture: ideal membership questions and recent advances*, Contemp. Math. **369** (2005), 261–276.
36. S. Zachary, *Countable state space Markov random fields and Markov chains on trees*, Ann. Probab. **11** (1983), no. 4, 894–903.
37. D. Zeilberger, *Toward a combinatorial proof of the Jacobian conjecture?*, G. Labelle, P. Leroux (eds.), Combinatoire énumérative, Lecture Notes in Mathematics, vol. 1234, Springer, Berlin, Heidelberg, 1986.
38. W. Zhao, *Recurrent inversion formulas*, arXiv:math/0305162, 2003.
39. W. Zhao, *Inversion problem, Legendre transform and inviscid Burgers' equations*, J. Pure. Appl. Algebra **199** (2005), no. 1-3, 299–317.