The uniform Kruskal theorem: between finite combinatorics and strong set existence

Anton Freund and Patrick Uftring

Department of Mathematics, Technical University of Darmstadt, Schloßgartenstr. 7, 64289 Darmstadt, Germany

The uniform Kruskal theorem extends the original result for trees to general recursive data types. As shown by Freund, Rathjen and Weiermann (Freund, Rathjen, Weiermann 2022 *Adv. Math.* 400, 108265 (doi:10.1016/j.aim.2022.108265)), it is equivalent to $\Pi^1_1$-comprehension, over $\text{RCA}_0$ with the ascending descending sequence principle (ADS). This result provides a connection between finite combinatorics and abstract set existence. The present article sheds further light on this connection. Firstly, we show that the original Kruskal theorem is equivalent to the uniform version for data types that are finitely generated. Secondly, we prove a dichotomy result for a natural variant of the uniform Kruskal theorem. On the one hand, this variant still implies $\Pi^1_1$-comprehension over $\text{RCA}_0$ extended by the chain antichain principle (CAC). On the other hand, it becomes weak when CAC is removed from the base theory.

This article is part of the theme issue ‘Modern perspectives in Proof Theory’.

1. Introduction

Kruskal’s theorem [1] asserts that any infinite sequence $t_0, t_1, \ldots$ of finite rooted trees admits $i < j$ such that $t_i$ embeds into $t_j$. Precise definitions for the case of ordered trees can be found in §2. The result for unordered trees has the same strength, due to results by Rathjen & Weiermann [2] as well as van der Meeren [3, chapter 2].

As famously shown by Friedman [4], Kruskal’s theorem is unprovable in predicative axiom systems, so that we have a concrete mathematical example for the incompleteness phenomenon from Gödel’s theorems. Specifically, Kruskal’s theorem has consistency strength strictly between $\text{ATR}_0$ and $\Pi^1_1\text{-CA}_0$, the two strongest of...
five particularly prominent axiom systems from reverse mathematics (see the textbook by Simpson [5] for general background).

In [6, it is shown that $\Pi_1^1$-comprehension (the central axiom of $\Pi_1^1\text{-}CA_0$) is equivalent to a uniform version of Kruskal’s theorem, which extends the latter from trees to general recursive data types. In our view, the equivalence sheds some light on a fascinating phenomenon that is typical for modern mathematics, namely, on connections between the abstract and the concrete. To refine our understanding of these connections, the present article investigates how the strength of the uniform Kruskal theorem is affected by certain finiteness conditions.

Our first aim is to recall the precise formulation of the uniform Kruskal theorem. As the following discussion is rather condensed, we point out that more explanations and examples can be found in the introduction of [6]. A function $f : X \to Y$ between partial orders is called a quasi-embedding if it reflects the order, i.e. if $f(x) \leq y$ entails $x \leq x'$. It is an embedding if the order is also preserved, i.e. if the converse implication holds as well. By $\mathcal{PO}$, we denote the category with the partial orders as objects and the quasi-embeddings as morphisms. A functor $\mathcal{W} : \mathcal{PO} \to \mathcal{PO}$ is said to preserve embeddings if

$$\mathcal{W}(X)$$

is well partial order whenever the same holds for $X$. If $\mathcal{W}(X)$ is a well partial order whenever the same holds for $X$, then $\mathcal{W}$ is called a $\text{WP0}$-dilator. We say that $\mathcal{W}$ is normal if

$$\sigma \leq \mathcal{W}(X) \tau \quad \Rightarrow \quad \text{for any } x \in \text{supp}_X(\sigma) \text{ there is an } x' \in \text{supp}_Y(\tau) \text{ with } x \leq x'.$$

holds for any partial order $X$ and all $\sigma, \tau \in \mathcal{W}(X)$.

One may think of $\sigma \in \mathcal{W}(X)$ as a finite structure with labels from $\text{supp}_X(\sigma) \subseteq X$. In typical examples, we have $\sigma \leq \mathcal{W}(X) \sigma$ precisely if there is an embedding that sends each label to a larger one, which witnesses that $\mathcal{W}$ is normal.

**Definition 1.1.** A P0-dilator consists of a functor $\mathcal{W} : \text{PO} \to \text{PO}$ that preserves embeddings and a natural transformation $\text{supp} : \mathcal{W} \Rightarrow [\cdot]^{<\omega}$ such that we have

$$\text{supp}_Y(\sigma) \subseteq \text{rng}(f) \quad \Rightarrow \quad \sigma \in \text{rng}(\mathcal{W}(f)),$$

for any embedding $f : X \to Y$ and any $\sigma \in \mathcal{W}(Y)$. If $\mathcal{W}(X)$ is a well partial order whenever the same holds for $X$, then $\mathcal{W}$ is called a $\text{WP0}$-dilator. We say that $\mathcal{W}$ is normal if

$$\sigma \leq \mathcal{W}(X) \tau \quad \Rightarrow \quad \text{for any } x \in \text{supp}_X(\sigma) \text{ there is an } x' \in \text{supp}_Y(\tau) \text{ with } x \leq x'.$$

holds for any partial order $X$ and all $\sigma, \tau \in \mathcal{W}(X)$.

**Definition 1.2.** A Kruskal fixed point of a normal P0-dilator $\mathcal{W}$ consists of a partial order $X$ and a bijection $\kappa : \mathcal{W}(X) \to X$ such that all $\sigma, \tau \in \mathcal{W}(X)$ validate

$$\kappa(\sigma) \leq x \kappa(\tau) \quad \Leftrightarrow \quad \sigma \leq \mathcal{W}(X) \tau \text{ or } \kappa(\sigma) \leq x \kappa(\tau) \text{ for some } x \in \text{supp}_X(\tau).$$

We say that the given Kruskal fixed point is initial if any other Kruskal fixed point $\pi : \mathcal{W}(Y) \to Y$ admits a unique quasi-embedding $f : X \to Y$ with $f \circ \kappa = \pi \circ \mathcal{W}(f)$.

The finite ordered trees without labels form the initial Kruskal fixed point of the transformation that maps $X$ to the collection of finite sequences in $X$ (with the order from Higman’s lemma). Here, $\kappa$ implements the usual recursive construction that builds a tree $\kappa(\sigma)$ from the list $\sigma$ of its immediate subtrees. The aforementioned equivalence says that a tree $\kappa(\sigma)$ embeds into $\kappa(\tau)$ when the immediate subtrees $\sigma$ embed into corresponding subtrees $\tau$ or the entire tree $\kappa(\sigma)$ embeds into some subtree $x \in \text{supp}_X(\tau)$. In [6, section 3], it is shown that each normal P0-dilator $\mathcal{W}$ has an initial Kruskal fixed point, which is unique up to isomorphism and will be denoted by $\mathcal{T} \mathcal{W}$.

**Definition 1.3.** By the uniform Kruskal theorem for a collection $\Gamma$ of P0-dilators, we mean the statement that $\mathcal{T} \mathcal{W}$ is a well partial order for every normal $\mathcal{W} \in \Gamma$. 
Let us stress that normality is assumed by default (and will not be mentioned explicitly), as the very construction of $TW$ may fail when $W$ is not normal. The uniform Kruskal theorem is true when $Γ$ is the collection of all $WPO$-dilators, which is the case studied in [6]. When $W$ is not a $WPO$-dilator, $TW$ may fail to be a well partial order (consider constant $W$ with empty supports). To get the original Kruskal theorem for unlabelled trees, one declares that $Γ$ contains the single $WPO$-dilator that maps $X$ to the sequences in $X$ (and a similar $WPO$-dilator yields trees with labels from a given well partial order).

We now recall a more concrete approach to $P0$-dilators, which allows for a formalization in reverse mathematics. Let $P0_0 \subseteq P0$ contain a unique representative from each isomorphism class of finite partial orders. For each such order $a$, we pick an isomorphism $en_a : |a| \to a$ with $|a| \in P0_0$. When $a$ is a finite suborder of some $X$, we define $en^X_a := t \circ en_a$ with $t : a \leftrightarrow X$. Owing to the support condition from definition 1.1, each $σ \in W(X)$ has a unique normal form

$$σ =_{nf} W(en^X_a)(σ_0) \quad \text{with} \quad a = supp_σ(σ),$$

where we have $σ_0 \in W(|a|)$. The condition $a = supp_σ(σ)$ is equivalent to $|a| = supp_{|a|}(σ_0)$, as we have $supp_σ(σ) = [en^X_a]_{\omega} \circ supp_{|a|}(σ_0)$ by the naturality of supports. This motivates the following analogue of a definition due to Girard [7].

**Definition 1.4.** The trace of a $P0$-dilator $W$ is given by

$$Tr(W) := \{(c, ρ) | c \in P0_0 \text{ and } ρ \in W(c) \text{ with } supp_ρ(ρ) = c\}.$$  

We say that $W$ is finite if $Tr(W)$ is a finite set.

As mentioned earlier, the elements of $W(X)$ can be uniquely represented by pairs $(a, σ_0)$ with $a \in [X]^{<ω}$ and $(|a|, σ_0) \in Tr(W)$. One can infer that $W$ is determined (up to isomorphism) by its restriction to $P0_0$. Assuming that $Tr(W)$ is countable, this makes it possible to code $W$ in the framework of reverse mathematics. The details given in [6] reveal that being a (coded) $P0$-dilator is an arithmetical property, while being a $WPO$-dilator is $Π^1_2$. Girard has shown that the notion of dilator on linear orders is $Π^1_2$-complete (see the proof by Normann given in [8], appendix 8.E). The analogous result for $WPO$-dilators follows by section 5 of [6] (where theorem 5.11 provides the crucial reduction). Let us also point out that $W$ is finite precisely if $W(a)$ is finite for each finite order $a$ and there is a number $n \in \mathbb{N}$ such that the support $supp_σ(σ)$ of any $σ \in W(X)$ has at most $n$ elements (cf. the proof of proposition 2.2).

Let us now consider an initial Kruskal fixed point $κ : W(TW) → TW$. Given $a \in [TW]^{<ω}$ and $(|a|, σ_0) \in Tr(W)$, we abbreviate

$$Φ(a, σ_0) := κ(σ) \quad \text{for} \quad σ =_{nf} W(en^TW_a)(σ_0).$$

Since $κ$ is bijective, each element of $TW$ can be uniquely written as $Φ(a, σ_0)$. Also, the fact that $κ$ is initial ensures that we get a height function

$$h : TW → \mathbb{N} \quad \text{with} \quad h(Φ(a, σ_0)) = \max \{0 \cup \{h(t) + 1 | t \in a\} \},$$

as shown by Theorem 3.5 of [9] (see also the paragraph before Lemma 3.5 of [6]). This suggests a recursive construction of $TW$ as a set of terms, which is available in the weak base theory $RCA_0$ from reverse mathematics. The latter shows that $TW$ is an initial Kruskal fixed point, as verified in [6, section 3]. At the same time, the principle that initial Kruskal fixed points are well partial orders is very strong, by the following theorem that was already mentioned. We recall that the ascending descending sequence principle (ADS) asserts that every infinite linear order contains an infinite sequence that is either strictly increasing or strictly decreasing. In [6], the theorem was stated with the chain antichain principle (CAC) in place of ADS. The latter is weaker (see [10]) and suffices according to Remark 2.7 below.

**Theorem 1.5 [6].** The following are equivalent over $RCA_0 + ADS$:

(i) the uniform Kruskal theorem for $WPO$-dilators,

(ii) the principle of $Π^1_1$-comprehension.
In the first half of the present article, we study the uniform Kruskal theorem for \( \mathsf{PO} \)-dilators that are finite. The initial Kruskal fixed points of these \( \mathsf{PO} \)-dilators may be seen as finitely generated data types, since we can view \( \sigma(a, \sigma_0) \in TW \) as the value of the constructor \( \sigma_0 \) on arguments \( a \). To summarize our results, we introduce one further notion:

**Definition 1.6.** For quasi-embeddings \( f, g : X \to Y \) between partial orders, we write \( f \leq g \) if we have \( f(x) \leq Y g(x) \) for all \( x \in X \). A \( \mathsf{PO} \)-dilator \( W \) is called monotone if \( f \leq g \) entails \( W(f) \leq W(g) \).

As we show in the next section, each \( \mathsf{WPO} \)-dilator is monotone, provably in \( \mathsf{RCA}_0 \). Conversely, any finite monotone \( \mathsf{PO} \)-dilator is a \( \mathsf{WPO} \)-dilator, but this cannot be shown in \( \mathsf{RCA}_0 \). Indeed, this theory proves it equivalent to the fact that \( X \mapsto X^n \) preserves well partial orders for all \( n \in \mathbb{N} \). Essentially over \( \mathsf{RCA}_0 \) (but modulo a result that Rathjen and Weiermann \([2]\) prove over \( \mathsf{ACA}_0 \)), we show that the uniform Kruskal theorem for finite monotone \( \mathsf{PO} \)-dilators is equivalent to the original Kruskal theorem for unlabelled trees (which is considerably weaker than the statements from theorem 1.5). The proof idea comes from the final section of Kruskal’s original paper \([1]\), but one needs to show that it applies in the general setting of dilators. This is not self-evident, as related statements remain strong under finiteness conditions (consider binary trees with gap condition in \([11]\), weakly finite dilators in \([12]\) and the paragraph after corollary 2.6).

In the second half of our article, we investigate versions of the uniform Kruskal theorem in the absence of \( \mathsf{ADS} \). Let us write \( \psi \) for the uniform Kruskal theorem for \( \mathsf{WPO} \)-dilators \( W \) such that \( W(1) \) is finite (where 1 is the order with a single element). As we show in \( \S 3 \), this variant leads to a surprising dichotomy: On the one hand, \( \psi \) does still imply \( \Pi^1_1 \)-comprehension over \( \mathsf{RCA}_0 + \mathsf{CAC} \). On the other hand, any \( \Pi^1_1 \)-theorem of \( \mathsf{RCA}_0 + \psi \) can already be proved in a theory that is weaker than \( \mathsf{ACA}_0 \) (and hence much weaker than \( \Pi^1_1 \)-comprehension). As a consequence, \( \mathsf{RCA}_0 \) cannot prove that the original Kruskal theorem for binary trees follows from the uniform Kruskal theorem for finite \( \mathsf{WPO} \)-dilators, in contrast with the previous paragraph (recall that the finite \( \mathsf{WPO} \)-dilators and the finite monotone \( \mathsf{PO} \)-dilators do not coincide over \( \mathsf{RCA}_0 \)). We do not know whether a similar dichotomy holds in the case of theorem 1.5. At any rate, it seems remarkable that some natural version of the uniform Kruskal theorem can have either predicative or impredicative strength, depending on weak principles in the base theory.

2. Trees and finitely generated data types

In this section, we prove an equivalence between the usual Kruskal theorem for trees and the uniform Kruskal theorem for finite monotone \( \mathsf{PO} \)-dilators. Before, we show that a finite \( \mathsf{PO} \)-dilator is monotone precisely if it is a \( \mathsf{WPO} \)-dilator.

The first implication holds over the usual weak base theory. We note that Girard has proved the analogous result for dilators on linear orders, using the fact that \( \omega^\omega \) is well founded (see \([7]\), proposition 2.3.10 and also \([6]\), lemma 5.3).

**Proposition 2.1.** Any \( \mathsf{WPO} \)-dilator is monotone, provably in \( \mathsf{RCA}_0 \).

**Proof.** We identify \( n \in \mathbb{N} \) with the set \( \{0, \ldots, n - 1\} \) ordered as an antichain and equip \( n \times \mathbb{N} \) with the usual product order, so that \( (c, k) \leq (d, l) \) holds precisely when we have \( c = d \) and \( k \leq l \) as natural numbers. To see that \( \mathsf{RCA}_0 \) recognizes \( n \times \mathbb{N} \) as a well partial order, we consider an infinite sequence

\[(c_0, k_0), (c_1, k_1), \ldots \subseteq n \times \mathbb{N},\]

and assume that it is bad, i.e. that \( (c_i, k_i) \not\leq (c_j, k_j) \) holds for all \( i < j \). For fixed \( i \), there are only finitely many \( (c_i, k_i) \) with \( k_i < k_i \). After passing to a recursive subsequence, we may thus assume that \( i \mapsto k_i \) is increasing. But we easily find indices \( i < j \) with \( c_i = c_j \). This yields \( (c_i, k_i) \not\leq (c_j, k_j) \), against our assumption.

Given a \( \mathsf{WPO} \)-dilator \( W \), we conclude that \( W(n \times \mathbb{N}) \) is a well partial order for all \( n \in \mathbb{N} \). To deduce that \( W \) is monotone, we consider two quasi-embeddings \( f, g : X \to Y \) with \( f \leq g \). We first prove \( W(f) \leq W(g) \) under the following assumptions:
(A1) the partial order $X$ is finite,
(A2) any element of $Y$ lies in the range of $f$ or $g$,
(A3) we have $f(x) \neq g(x')$ for $x \neq x'$.

In view of (A1), we fix an enumeration of the $n$-element set $X = \{x_0, \ldots, x_{n-1}\}$. For each $k \in \mathbb{N}$, we have a quasi-embedding

$$h_k : X \to n \times \mathbb{N} \quad \text{with} \quad h_k(x_i) := \begin{cases} (i, 0) & \text{if } f(x_i) = g(x_i), \\ (i, k) & \text{otherwise}. \end{cases}$$

Given $\sigma \in W(X)$, we find $k < l$ with $W(h_k(\sigma)) \leq W(h_l(\sigma))$ in $W(n \times \mathbb{N})$, as the latter is a well partial order. Let us recall that any quasi-embedding is injective (use anti-symmetry in $X$ to derive $x = x'$ from $f(x) = f(x')$). Combined with conditions (A2) and (A3), this allows us to consider

$$h : Y \to n \times \mathbb{N} \quad \text{with} \quad h(y) := \begin{cases} (i, 0) & \text{if } y = f(x_i) = g(x_i), \\ (i, k) & \text{if } y = f(x_i) \neq g(x_i), \\ (i, l) & \text{if } y = g(x_i) \neq f(x_i). \end{cases}$$

To show that $h$ is a quasi-embedding, we consider an inequality $h(y) = (i, p) \leq (i, q) = h(y')$. If we have $p, q \in \{0, k\}$, we get $y = f(x_i) = y'$. Similarly, $p = l = q$ entails $y = g(x_i) = y'$. Given $p \leq q$, the only other possibility amounts to $p \in \{0, k\}$ and $q = l$. Here, $f \leq g$ ensures $y = f(x_i) \leq g(x_i) = y'$. Now observe that we have

$$W(h \circ f)(\sigma) = W(h_k(\sigma)) \leq W(h_l(\sigma)) = W(h \circ g)(\sigma).$$

Since $W(h)$ is a quasi-embedding, we get $W(f)(\sigma) \leq W(g)(\sigma)$, as required.

Let us now consider quasi-embeddings $f, g : X \to Y$ with $f \leq g$ that validate (A3) but not necessarily (A1) and (A2). Given $\sigma \in W(X)$, we set

$$X_0 := \supp_X(\sigma) \subseteq X \quad \text{and} \quad Y_0 := [f]^{\leq}\sigma(X_0) \cup [g]^{\leq}\sigma(X_0) \subseteq Y.$$  

Write $\iota_X : X_0 \hookrightarrow X$ and $\iota_Y : Y_0 \to Y$ for the inclusions. We define $f_0, g_0 : X_0 \to Y_0$ by $\iota_Y \circ f_0 = f \circ \iota_X$ and $\iota_Y \circ g_0 = g \circ \iota_X$. This clearly yields quasi-embeddings $f_0 \leq g_0$ that validate (A1) to (A3). Owing to the support condition from definition 1.1, we may write $\sigma = W(\iota_X)(\sigma_0)$ with $\sigma_0 \in W(X_0)$. We obtain

$$W(f)(\sigma) = W(f \circ \iota_X)(\sigma_0) = W(\iota_Y \circ f_0)(\sigma_0) \leq W(\iota_Y \circ g_0)(\sigma_0) = W(g)(\sigma),$$

by the result for $f_0, g_0$ and the fact that $W(\iota_Y)$ is an embedding (also by definition 1.1).

Finally, we also drop (A3) and consider arbitrary quasi-embeddings $f, g : X \to Y$ with $f \leq g$. Let us equip $Y \times 2$ with the lexicographic order, in which $(y, 0) < (y, 1)$ and $(y, i) < (y', j)$ hold whenever we have $y < y'$ in $Y$ (note that 2 is no longer an antichain). We consider the embedding

$$\iota : Y \to Y \times 2 \quad \text{with} \quad \iota(y) := (y, 0) \quad \text{and the quasi-embedding}$$

$$f^+ : X \to Y \times 2 \quad \text{with} \quad f^+ \colon (x, 0) := \begin{cases} (f(x), 0) & \text{if } f(x) = g(x), \\ (f(x), 1) & \text{otherwise}. \end{cases}$$

Note that we have $\iota \circ f \leq f^+$ and that (A3) holds for $\iota \circ f$ and $f^+$. To see that we have $f^+ \leq \iota \circ g$, assume $f(x) \neq g(x)$ and note that $f \leq g$ forces $f(x) < g(x)$, so that we get

$$f^+(x) = (f(x), 1) < (g(x), 0) = \iota \circ g(x).$$

Furthermore, $f^+$ and $\iota \circ g$ validate (A3). Indeed, if we have

$$f^+(x) = (f(x), i) = (g(x'), 0) = \iota \circ g(x'),$$

we get $i = 0$ and hence $f(x) = g(x)$, which yields $x = x'$ by the injectivity of $g$. For $\sigma \in W(X)$, the result under (A3) will now yield

$$W(\iota \circ f)(\sigma) \leq W(f^+)(\sigma) \leq W(\iota \circ g)(\sigma).$$
As \( W(i) \) is a (quasi) embedding, we once again get \( W(f)(\sigma) \leq W(g)(\sigma) \).

For dilators on linear orders, Girard has also proved a converse result [7, proposition 4.3.8]. We establish the analogue for partial orders and include a reversal. Let us point out that \( X^n \) carries the usual product order, in which we have \( (x_0, \ldots, x_{n-1}) \leq (x'_0, \ldots, x'_{n-1}) \) precisely when \( x_i \leq x'_i \) holds for all \( i < n \). Concerning the following statement (ii), it is known that the closure of well quasi-orders under binary products follows from \( \mathcal{CC} \) and entails \( \mathcal{ADS} \) (see [13] and the paragraph before theorem 2.22 of [14]). Both implications are strict (even in the presence of weak König’s lemma), as shown by Towsner [15]. Also, \( \Sigma^0_2 \)-induction follows from the statement that \( \alpha^n \) is a well partial order for any well order \( \alpha \) and all \( n \in \mathbb{N} \) [16, theorem 5].

**Proposition 2.2.** The following are equivalent over \( \mathcal{RCA}_0 \):

(i) any finite monotone \( 0 \)-dilator is a \( \mathcal{WPO} \)-dilator,

(ii) if \( X \) is a well partial order, then so is \( X^n \) for every \( n \in \mathbb{N} \).

**Proof.** To see that (i) implies (ii), we consider the monotone \( 0 \)-dilator \( W \) with \( W(X) := X^n \) and

\[
W(f)((x_0, \ldots, x_{n-1})) := (f(x_0), \ldots, f(x_{n-1}))
\]

and

\[
\text{supp}_X((x_0, \ldots, x_{n-1})) := \{x_0, \ldots, x_{n-1}\}.
\]

It suffices to show that (the trace of) \( W \) is finite. Let us recall that \( (c, \rho) \in \text{Tr}(W) \) amounts to \( \rho \in W(c) \) and \( \text{supp}_c(\rho) = c \in \mathbb{N}_0 \). The latter entails that \( c \) can have at most \( n \) elements. But then there are only finitely many possibilities for \( c \), since \( \mathbb{N}_0 \) contains a single element from each isomorphism class of finite partial orders. Given that \( W(c) \) is finite for each finite \( c \), we could conclude by the principle that finite unions of finite sets are finite. However, the latter is not available in \( \mathcal{RCA}_0 \). Indeed, this theory proves it equivalent to \( \Sigma^0_2 \)-collection (see [17] for a proof). To circumvent this issue, we fix an enumeration \( c_0, \ldots, c_{N-1} \) of the orders in \( \mathbb{N}_0 \) that have at most \( n \) elements. For each \( i < N \), we also fix an injection \( e_i : c_i \to \{0, \ldots, n-1\} \). Let us write \( N \times n \) for the finite antichain that consists of the pairs \((i, k)\) with \( i < n \) and \( k < n \). Then the function \( e_i' : c_i \to N \times n \) with \( e_i'(x) := (i, e_i(x)) \) is a quasi-embedding. We note that \( i \) can be recovered from \( \text{rng}(e_i') \). Now consider

\[
e : \text{Tr}(W) \to W(N \times n) \quad \text{with} \quad e(c, \rho) := W(e_i') (\rho) \quad \text{for} \quad c = c_i.
\]

Given that \( W(N \times n) \) is finite, it suffices to show that \( e \) is injective. By the naturality of supports, we have \( \text{supp}_{N \times n}(e(c, \rho)) = \text{rng}(e_i') \). Thus, \( e(c, \rho) \) determines the unique \( i < N \) with \( c = c_i \). But then \( \rho \) is also determined by \( e(c, \rho) \), as \( W(e_i') \) is a quasi-embedding and hence injective.

We now show that (ii) implies (i). Let \( W \) be any finite monotone \( 0 \)-dilator. To view finite partial orders as sequences, we fix a bijection

\[
\{0, \ldots, \text{len}(c) - 1\} \ni i \mapsto c[i] \in c
\]

with \( \text{len}(c) \in \mathbb{N} \) for each \( c \in \mathbb{N}_0 \). We then pick an \( n \in \mathbb{N} \) such that \( (c, \rho) \in \text{Tr}(W) \) entails \( \text{len}(c) \leq n \). Given a well partial order \( X \), we consider

\[
\text{Tr}(W) + X := \{0, \xi \mid \xi \in \text{Tr}(W)\} \cup \{(1, x) \mid x \in X\},
\]

with the summand \( \text{Tr}(W) \) ordered as an antichain and no comparisons between the two summands, so that the only strict inequalities are \((1, x) < (1, x')\) for \( x < x' \) in \( X \). Since this yields a well partial order, it suffices to construct a quasi-embedding

\[
g : W(X) \to (\text{Tr}(W) + X)^{n+1}.
\]

As we have seen in §1, each element of \( W(X) \) has a unique normal form \( \sigma =_n W(\text{en}_X^X)(\sigma_0) \) with \((|a|, \sigma_0) \in \text{Tr}(W) \), for a certain embedding \( \text{en}_X^X \) that maps \(|a| \in \mathbb{N}_0 \) onto \( a \subseteq X \). When \( \sigma \in W(X) \) has
normal form as given, we put

\[ g(\sigma) := (y_0, \ldots, y_n) \text{ with } y_i := \begin{cases} (1, e^{\chi}_{\kappa}(|a|)|i|) & \text{when } i < \text{len } (|a|), \\ (0, (|a|, o_0)) & \text{otherwise.} \end{cases} \]

Let us note that \( y_n \) is always determined by the second case due to the choice of \( n \). Given \( g(\sigma) \leq g(\tau) \), we thus obtain \( \sigma =_{ml} W(e^{\chi}_{\kappa})(\rho) \) and \( \tau =_{ml} W(e^{\chi}_{\kappa})(\rho) \) for a single \( \rho \) and with \( |a| = |b| \).

The components with index \( i < \text{len}(|a|) \) ensure that we have \( e^{\chi}_{\kappa} \leq e^{\chi}_{\kappa} \). Now monotonicity yields \( \sigma \leq \tau \), as required. \qed

To connect with the original Kruskal theorem, we define certain collections of finite ordered trees. In the following, \( l \star (t_0, \ldots, t_k) \) stands for the tree with root label \( l \) and immediate subtrees \( t_i \) (which recursively yields labels at all vertices). Let us point out that \( m = 1 \) in the following definition amounts to the case without labels (as a single label \( l = 0 \) has no effect). For \( n = \omega \), the condition \( k < n \) is equivalent to \( k \in \mathbb{N} \).

**Definition 2.3.** For \( m \in \mathbb{N} \) and \( n \in \mathbb{N} \cup \{\omega\} \), we declare that \( T^m_n \) is recursively generated as follows: Whenever \( t_0, \ldots, t_{k-1} \in T^m_n \) with \( k < n \) are already constructed, we add a new element \( l \star (t_0, \ldots, t_{k-1}) \in T^m_n \) for every \( l < m \). Let us put \( T^\omega_n := T^1_n \).

An embedding of a tree \( t \) into \( t' \) will either send the root to the root and immediate subtrees into corresponding subtrees or it will send all of \( t \) into a proper subtree of \( t' \). We also demand that labels are preserved. The embeddability relation \( \leq \) can thus be characterized as follows (where we write \([k] := (0, \ldots, k - 1)\)).

**Definition 2.4.** For trees \( l \star \tau, l' \star \tau' \in T^m_n \) with \( \tau = (t_0, \ldots, t_{j-1}) \) and \( \tau' = (t'_{0}, \ldots, t'_{k-1}) \), we recursively declare

\[ l \star \tau \leq l' \star \tau' : \iff (l = l' \text{ and } \tau \leq \tau') \text{ or } l \star \tau \leq l'_i \text{ for some } i < k \]

with

\[ \tau \leq \tau' : \iff \begin{cases} \text{there is a strictly increasing } f : [j] \to [k] \\ \text{with } t_i \leq t'_i(f(i)) \text{ for all } i < j. \end{cases} \]

We come to the main result of this section. The condition that branchings are bounded will be removed in a corollary.

**Theorem 2.5.** The following are equivalent over \( \mathsf{RO}^0 \):

(i) the uniform Kruskal theorem for finite monotone \( \mathsf{PO} \)-dilators,

(ii) the original Kruskal theorem for finite ordered trees with bounded branchings, which is the statement that \( (T_n, \leq) \) is a well partial order for all \( n \in \mathbb{N} \).

**Proof.** To show that (i) implies (ii), we fix \( n \in \mathbb{N} \) and put

\[ W(X) := \{ \langle x_0, \ldots, x_{k-1} \rangle \mid x_i \in X \text{ for } i < k < n \}. \]

When \( X \) is a partial order, we declare that \( \langle x_0, \ldots, x_{j-1} \rangle \leq W(X) \langle x'_0, \ldots, x'_{k-1} \rangle \) holds precisely if there is a strictly increasing \( f : [j] \to [k] \) with \( x_i \leq x'_{f(i)} \) for all \( i < j \). Let us define \( W(f) \) and \( \text{supp}_X \) analogous to the proof of proposition 2.2, to obtain a finite monotone \( \mathsf{PO} \)-dilator that is clearly normal. It suffices to construct a quasi-embedding of \( T_n \) into the initial Kruskal fixed point \( TW \) of \( W \). As seen in §1, the latter comes with a map \( \kappa : W(TW) \to TW \).

We consider

\[ f : T_n \to TW \text{ with } f(0 \star (t_0, \ldots, t_{k-1})) := \kappa((f(t_0), \ldots, f(t_{k-1}))), \]

which amounts to a recursion over trees. A straightforward induction over the number of vertices confirms that \( f(t) \leq TW(f(t')) \) entails \( t \leq t' \), as the disjuncts in the definition of \( \leq \) correspond to those in definition 1.2.

In order to prove that (ii) implies (i), we first show that it entails
(ii') the relation \( \preceq \) is a well partial order on \( \mathbb{T}^m_n \) for all \( m, n \in \mathbb{N} \).

The idea is to simulate labels \( l < m \) by trees \( t(l) \) that are incomparable under \( \preceq \). Write \( t_k^l \in \mathbb{T}_{k+1} \) for the full \( k \)-branching of tree height \( j \), given by \( t_0^k := 0 \) and \( t_{i+1}^k := 0 \cdot (t_0^k, \ldots, t_{i-1}^k) \) with \( t_i^k = t_i \) for all \( i < k \). A straightforward induction shows that \( t_k^l \preceq t_i^l \) with \( i, k > 0 \) entails \( i \leq j \) and \( k \leq l \). Our incomparable trees can thus be given by \( t(l) := \mathbb{T}^n_{l+1} \) for \( l < m \). Aiming at (ii'), we may assume \( m < n \), which ensures \( t(l) \in \mathbb{T}_{n+1} \). Let us now define \( g : \mathbb{T}^m_n \to \mathbb{T}_{n+1} \) by

\[
g(l \cdot (t_0, \ldots, t_{k-1})) := 0 \cdot (t'_0, \ldots, t'_{k-1}) \quad \text{with} \quad t'_i := \begin{cases} g(t_i) & \text{for} \ i < k, \\ t(l) & \text{for} \ k \leq i < n. \end{cases}
\]

To show that \( g(s) \preceq g(t) \) entails \( s \preceq t \), we write \( s = l \ast \sigma \) and \( t = l' \ast \tau \) with \( \sigma = \langle s_0, \ldots, s_{j-1} \rangle \) and \( \tau = \langle t_0, \ldots, t_{j-1} \rangle \). First assume that \( g(s) \preceq g(t) \) holds by the first disjunct from definition 2.4. We then have \( s'_j \preceq t'_i \) for all \( i < n \), with \( s'_j \) and \( t'_i \) as in the definition of \( g \). As the definition of \( \mathbb{T}^m_n \) ensures \( j, k < n \), we can deduce \( t(l) = s'_{n-1} \preceq t'_{n-1} = t(l') \) and hence \( l = l' \). For \( i < \min(j, k) \), we also obtain \( g(s_i) = s'_i \preceq t'_i = g(t_i) \), so that we may inductively infer \( s_i \preceq t_i \). To get \( s \preceq \ast \tau \) and with that \( s \preceq t \), we need only show \( j \leq k \). If the latter was false, we would get

\[
g(s_k) = s'_k \preceq t'_k = t(l) = t'_{k+1}.
\]

However, as \( g(s_k) \) has \( n \) immediate subtrees with \( n > m \geq l + 1 \), it is straightforward to refute \( g(s_k) \preceq t'_{k+1} \) by induction on \( k \). If \( g(s) \preceq g(t) \) holds by the second disjunct from definition 2.4, we have \( g(s) \preceq t'_i \) for some \( i < n \). Given that \( g(s) \preceq t(l) \) was just refuted, we must have \( i < k \) and \( t'_i = g(t_i) \). So we inductively get \( s \preceq t_i \) and then \( s \preceq t \).

To complete the proof, we show that (ii') entails (i). Given a finite monotone \( \mathbb{P}_0 \)-dilator \( W \) that is normal, we fix an injective function \( e : \text{Tr}(W) \to \mathbb{N} \) for a suitable \( m \in \mathbb{N} \). Recall the bijections \( [\text{len}(c)] \ni i \mapsto c[i] \in e \) for \( c \in \mathbb{P}_0 \) that were fixed in the proof of proposition 2.2. As in the latter, each finite \( a \subseteq TW \) gives rise to an embedding \( e^W_T \) with domain \( |a| \in \mathbb{P}_0 \) and range \( a \subseteq TW \). Pick an \( n \in \mathbb{N} \) such that \( (c, \rho) \in \text{Tr}(W) \) entails \( \text{len}(c) < n \). From the paragraph after definition 1.4, we recall the notation \( o(a, \sigma) \) for elements of \( TW \) as well as the height function \( h : TW \to \mathbb{N} \). The latter justifies the recursive definition of a function \( j : TW \to \mathbb{T}^m_n \) with

\[
j(o(a, \sigma)) := e(|a|, \sigma) \ast (t_0, \ldots, t_{\text{len}(|a|)-1}) \quad \text{for} \quad t_i := j \left( e^W_T(|a|[i]) \right).
\]

In view of [6, Definition 3.4], this can be cast as a recursion over subterms, which permits formalization in \( \mathbb{R}_0 \). It remains to show that \( j(s) \preceq j(t) \) entails \( s \preceq TW \), for which we use induction on \( h(s) + h(t) \). Write \( s = o(a, \sigma) \) and \( t = o(b, \tau) \). We first assume that \( j(s) \preceq j(t) \) holds by the first disjunct in definition 2.4, so that we have \( e(|a|, \sigma) = e(|b|, \tau) \) and hence \( (|a|, \sigma) = (|b|, \tau) \) as well as \( j \circ e^W_T(|a|[i]) \preceq j \circ e^W_T(|b|[i]) \) for \( i < \text{len}(|a|) = \text{len}(|b|) \). Inductively, we can conclude that we have \( e^W_T \preceq e^W_T \). Owing to \( \sigma = \tau \), we can invoke monotonicity and the equivalence from definition 1.2 to conclude

\[
s = o(a, \sigma) = \kappa \left( W(e^W_T(|a|)) \right) \leq_{TW} \kappa \left( W(e^W_T(|b|)) \right) = o(b, \tau) = t.
\]

Note that the second and penultimate equalities hold by the definition of the notation \( o(a, \sigma) \) in §1. If \( j(s) \preceq j(t) \) holds by the second disjunct from definition 2.4, then we have \( j(s) \preceq j(e^W_T(|b|[i])) \) for some \( i < \text{len}(|b|) \). Given that \( (|b|, \tau) \in \text{Tr}(W) \) entails \( \text{supp}_{|b|}(\tau) = |b| \), we can use the induction hypothesis to get

\[
s \preceq TW e^W_T(|b|[i]) \in \left[ e^W_T \right]^{<\omega} \circ \text{supp}_{|b|}(\tau) = \text{supp}_{TW} \circ W \left( e^W_T \right)(\tau).
\]

We get \( s \preceq TW \) by the second disjunct in the equivalence from definition 1.2.

Owing to the bound on branchings, statement (ii) of theorem 2.5 can in fact be deduced from a result of Higman [18] (see also the survey by Pouzet [19]). What is often called Higman’s lemma is a weaker special case of this result. Kruskal [1] has shown that the result remains valid when the
bound on branchings is removed. Perhaps surprisingly, the extension by Kruskal does not have higher logical strength (in the case of trees without labels): Rathjen & Weiermann have shown that (ii) is equivalent to statement (iii) (see [2], corollary 2.1 and [3], chapter 2). We keep the base theory $\text{ACA}_0$ from [2], even though $\text{RCA}_0$ seems to suffice (see the first line of [2], section 2, as well as [20, section 6]).

**Corollary 2.6.** Over $\text{ACA}_0$, statement (i) of theorem 2.5 is equivalent to

(iii) the original Kruskal theorem for the collection of all finite trees without labels, which is the statement that $(T_{\alpha,\omega}, \leq)$ is a well partial order.

Parallel to the equivalence between (ii) and (iii), one might believe that the uniform Kruskal theorem has similar strength for finite and infinite trace. This, however, is firmly refuted by the equivalence with $\Pi^1_1$-comprehension in theorem 1.5 (which is proved in [6]). The stark contrast between the finite and infinite case can be explained in terms of quantifier complexity: As mentioned in Section 1, the notion of WPO-dilator is $\Pi^1_2$-complete. For finite P0-dilators, on the other hand, preservation of well partial orders is equivalent to monotonicity (see propositions 2.1 and 2.2), which is determined on $P_0$ and hence an arithmetical condition (by the proof of the linear case in section 5 of [6]). Together with the fact that $TW$ is uniformly computable in $W$ (see section 3 of [6]), this means that the uniform Kruskal theorem for finite monotone P0-dilators is a $\Pi^1_1$-statement. It is known that $\Pi^1_1$-comprehension cannot be equivalent to a statement of this quantifier complexity (nor to a $\Pi^1_2$-statement). To conclude this section, we justify a slight improvement in theorem 1.5 that was mentioned earlier:

**Remark 2.7.** In [6], the equivalence from theorem 1.5 has been established over $\text{RCA}_0$ extended by the chain antichain principle. We argue that the latter may be weakened to ADS. For a well partial order $Z$, put $W_Z(X) := 1 + Z \times X$ and extend this into a normal P0-dilator. Details can be found in examples 2.3 and 3.3 of [6]. In the cited article, the chain antichain principle was assumed to ensure that $W_Z$ is a WPO-dilator, so that the uniform Kruskal theorem could be applied. Working in $\text{RCA}_0 + ADS$, one can show that $W_\alpha$ is a WPO-dilator whenever $\alpha$ is a well order: Given an infinite sequence $(\alpha_0, x_0), (\alpha_1, x_1), \ldots$ in $\alpha \times X$, the crucial task is to find an infinite subsequence in which the components $\alpha_i$ are weakly increasing. If there are only finitely many different $\alpha_i$, we can conclude by the infinite pigeonhole principle, which follows from ADS (via [21], proposition 4.5). Otherwise we achieve $\alpha_0 <_N \alpha_1 <_N \ldots$ by passing to a subsequence. Then ADS yields $i(0) < i(1) < \ldots$ with $\alpha(i(0)) < \alpha(i(1)) < \ldots$, as desired. Owing to example 3.9 of [6], the initial Kruskal fixed-point $TW_\alpha$ coincides with $\alpha^{<\omega}$, the set of finite sequences in $\alpha$ with the order from Higman’s lemma. So over $\text{RCA}_0 + ADS$, the uniform Kruskal theorem for WPO-dilators implies that $\alpha^{<\omega}$ is a well partial order for any well order $\alpha$. The latter entails arithmetical comprehension and in particular the chain antichain principle, via a result of Girard & Hirst (see [22], section II.5 and [23], theorem 2.6). Thus, the base theory of theorem 1.5 can be weakened as promised. Note that Tr($W_\alpha$) is infinite when the same holds for $\alpha$. The chain antichain principle cannot be derived from the uniform Kruskal theorem for finite monotone P0-dilators, as the latter is a $\Pi^1_1$-statement.

3. On Kruskal fixed points in the absence of ADS

In this section, we show that the class of WPO-dilators is heavily restricted in systems that reject the axiom ADS. Moreover, in theorems 3.5 and 3.9, we prove that a certain variant of the uniform Kruskal theorem is weak over $\text{RCA}_0$, while it entails $\Pi^1_1$-comprehension in the presence of CAC. The following notion will play a crucial role.

**Definition 3.1.** A P0-dilator $W$ is called unary if $\text{supp}_X(\sigma)$ has at most one element, for every partial order $X$ and all $\sigma \in X$. 

Note that the following result involves a theory that is not sound. Normal \( WP_0 \)-dilators that are not unary do in fact exist (see, e.g., the previous section). The result will later be applied inside \( \omega \)-models of \( \neg \text{ADS} \).

**Lemma 3.2.** In the theory \( \text{RCA}_0 + \neg \text{ADS} \), one can derive that any normal \( WP_0 \)-dilator \( W \) is unary.

**Proof.** Towards a contradiction, assume that \( a := \text{supp}_X(\sigma) \) contains two different elements \( x_0 \) and \( x_1 \). We invoke \( \neg \text{ADS} \) to pick an infinite well order \( Z \) that has no infinite ascending sequence, i.e. such that the inverse order \( Z^* \) is also well founded. Let us consider the well partial order

\[
Y := Z + Z^* + a := \{(i, y) \mid \text{either } i \in \{0, 1\} \text{ and } y \in Z \text{ or } i = 2 \text{ and } y \in a\},
\]

where the only inequalities are \( (0, z) \leq \gamma(0, z') \) and \( (1, z) \geq \gamma(1, z') \) for \( z \leq z' \) as well as \( (2, x) \leq \gamma(2, x) \) for \( x \in a \) (one could also use the induced order on \( a \subseteq X \)). For each \( z \in Z \), we have a quasi-embedding

\[
f_z : a \rightarrow Y \quad \text{with } f_z(x) := \begin{cases} 
(0, z) & \text{if } x = x_0, \\
(1, z) & \text{if } x = x_1, \\
(2, x) & \text{otherwise}.
\end{cases}
\]

Note that the assumption \( x_0 \neq x_1 \) is needed to make this well defined. Owing to the support condition from definition 1.1, we may write \( \sigma = W(i)(\sigma_0) \) with \( \sigma_0 \in W(a) \), where \( i : a \leftrightarrow X \) is the inclusion. The naturality of supports ensures that we have

\[
[i]^{<\omega} \circ \text{supp}_a(\sigma_0) = \text{supp}_X \circ W(i)(\sigma_0) = \text{supp}_X(\sigma) = a
\]

and hence \( \text{supp}_a(\sigma_0) = a \). For \( z \in Z \) we put \( \tau_z := W(f_z)(\sigma_0) \in W(Y) \) and note

\[
\text{supp}_Y(\tau_z) = \text{supp}_Y \circ W(f_z)(\sigma_0) = [f_z]^{<\omega} \circ \text{supp}_a(\sigma_0) = \{ (i, y) \mid \text{either } i \in \{0, 1\} \text{ and } y = z \text{ or } i = 2 \text{ and } y \in a \setminus \{x_0, x_1\}\}.
\]

Given that \( Z \) is infinite, we have an injection \( \mathbb{N} \rightarrow Z \). Let us compose the latter with the map \( z \mapsto \tau_z \), to get an infinite sequence in \( W(Y) \). The latter is a well partial order, as the same holds for \( Y \) and since \( W \) is a \( WP_0 \)-dilator. So \( Z \) must contain some \( y \neq z \) with \( \tau_y \leq W(Y)\tau_z \). By the normality condition from definition 1.1, we know that \( (0, y) \) and \( (1, y) \) in \( \text{supp}_Y(\tau_y) \) must be bounded by some elements of \( \text{supp}_Y(\tau_z) \), which can only be true if we have both \( (0, y) \leq \gamma(0, z) \) and \( (1, y) \leq \gamma(1, z) \). However, this yields \( y \leq z \), as well as \( y \geq z \) in contradiction with \( y \neq z \).

Our next result confirms that being unary is a serious restriction (cf. theorem 1.5). We recall that 1 denotes the order with a single element.

**Lemma 3.3.** The following are equivalent over \( \text{RCA}_0 \):

(i) the uniform Kruskal theorem for unary monotone \( \mathcal{P} \)-dilators \( W \) with the property that \( W(1) \) is a well partial order,

(ii) the principle of arithmetical comprehension.

**Proof.** The implication from (i) to (ii) holds by an argument from [6], which has been recalled in remark 2.7. In contrast with the remark, our implication does not rely on \( \text{ADS} \), because (i) does not require that \( W \) is a \( WP_0 \)-dilator (even though this follows over a sufficient base theory). For the implication from (ii) to (i), we consider the partial order

\[
Y := W(0) + W(1) = \{(i, \sigma) \mid i \in \{0, 1\} \text{ and } \sigma \in W(i)\},
\]

in which the two summands are incomparable, so that \( (i, \sigma) \leq \gamma(i, \tau) \) for \( \sigma \leq W(i)^\tau \) are the only inequalities. Given that \( W(1) \) is a well partial order, the same holds for \( Y \), as the empty function \( i : 0 \leftrightarrow 1 \) induces an embedding \( W(i) : W(0) \rightarrow W(1) \). Assuming (ii), we get Higman’s lemma in the
form that makes $Y^{<\omega}$ a well partial order (see, e.g., [5, theorem X.3.22]). For unary $W$, elements of Tr($W$) have the form $(i, a, \sigma)$ with $\sigma \in W(i)$ for $i \in \{0, 1\}$. We can thus consider

$$j : TW \to Y^{<\omega} \text{ with } j((a, \sigma)) := \begin{cases} (0, \sigma) & \text{if } a = \emptyset, \\ (1, \sigma, \neg) & \text{if } a = \{t\}. \end{cases}$$

Closely parallel to the proof of theorem 2.5, one inductively shows that $j$ is a quasi-embedding. It follows that $TW$ is a well partial order, as needed for the unique Kruskal theorem in (i).

In order to prove that the uniform Kruskal theorem is weak over $\mathcal{R}A_0$, one might try to find an $\omega$-model of $\neg \mathcal{D}S$ that validates statement (i). However, no such model can exist, as the equivalent statement (ii) entails $\mathcal{D}S$ In the following, we look at $\Pi^1_1$-consequences of (i), which are guaranteed to reflect into $\omega$-models. Let us recall that $\mathcal{P}_0$-dilators are officially coded as subsets of $\mathbb{N}$. The powerset of $\mathbb{N}$ will be denoted by $\mathcal{P}(\mathbb{N})$.

**Theorem 3.4.** Consider a theory $\Gamma \supseteq \mathcal{A}0$ and a class $\Gamma \subseteq \mathcal{P}(\mathbb{N})$ defined by a $\Sigma_1^1$-formula. Assume $\Gamma$ proves that $W(1)$ is a well partial order for any unary monotone $\mathcal{P}_0$-dilator $W \in \Gamma$. Write $\psi$ for the uniform Kruskal theorem for $\mathcal{P}_0$-dilators in $\Gamma$. Then any $\Pi^1_2$-theory of $\mathcal{R}A_0 + \psi$ is provable in $\Gamma$.

**Proof.** Assume $\mathcal{R}A_0 + \psi$ proves $XX \subseteq \mathbb{N}, \theta(X)$ for a $\Sigma_1^1$-formula $\theta$. To conclude by completeness, we consider an arbitrary model $M \models \Gamma$ and show that $M \models \theta(X)$ holds for any $X$ in its second-order part. Let $N \subseteq M$ be the $\omega$-submodel of all sets that $M$ believes to be computable in $X$. We recall that we have $N \models \mathcal{R}A_0$ (see, e.g., [5, theorem VIII.1.3]). A classical result of Tennenbaum asserts that $\mathcal{D}S$ fails in the realm of computable sets. The usual proof of this result (see, e.g., the textbook by Rosenstein [24]) relativizes and is readily formalized in $\mathcal{A}0 \subseteq \Gamma$ (see lemma 3.7 for a less direct argument over $\mathcal{R}A_0$). We thus get $N \models \neg \mathcal{D}S$. In view of proposition 2.1 and lemma 3.2, we can infer $N \models \varphi_0 \rightarrow \varphi$, where $\varphi_0$ is the uniform Kruskal theorem for unary monotone $\mathcal{P}_0$-dilators in $\Gamma$. By the previous lemma and the assumption on $\Gamma$, we have $M \models \varphi_0$. Being a unary monotone $\mathcal{P}_0$-dilator is an arithmetical property, as mentioned in §1 and the paragraph after corollary 2.6. Furthermore, the initial Kruskal fixed point $TW$ is uniformly computable in $W$, due to a construction from [6]. The computation of this fixed point yields the same result in both models $M$ and $N$, since the latter is an $\omega$-submodel of the former. Given that $\Gamma$ is $\Sigma_1^1$, it follows that $\varphi_0$ is $\Pi^1_1$, so that we get $N \models \varphi_0$ and hence $N \models \varphi$. By assumption, we get $N \models \theta(X)$ and then $M \models \theta(X)$, as $\theta$ is $\Sigma^1_1$. □

Part (a) of the following dichotomy result is obtained as a direct application of the previous theorem. With some additional work, one can replace $\mathcal{A}0$ by a weaker theory, as we shall see later. In order to indicate further applications of the previous theorem, we note that one can take $\Gamma$ to be the class of $\mathcal{P}_0$-dilators $W$ such that $W(1)$ admits a quasi-embedding into a fixed $X$ that $\Gamma$ proves to be a well partial order.

**Theorem 3.5.** Write $\psi$ for the uniform Kruskal theorem for $\mathcal{P}_0$-dilators $W$ such that $W(1)$ is finite.

(a) Any $\Pi^1_2$-theorem of $\mathcal{R}A_0 + \psi$ is provable in $\mathcal{A}0$.

(b) In $\mathcal{R}A_0 + \mathcal{C}0$, one can prove that $\psi$ is equivalent to $\Pi^1_1$-comprehension.

**Proof.** To obtain (a), one applies the previous theorem to the class $\Gamma$ of $\mathcal{P}_0$-dilators $W$ with the property that $W(1)$ is finite. The latter is a $\Sigma_1^1$-condition (in fact arithmetical) that entails that $W(1)$ is a well partial order. To derive (b) from theorem 1.5 (originally proved in [6]), we need only show that a normal $\mathcal{P}_0$-dilator $W$ can be transformed into a normal $\mathcal{P}_0$-dilator $W'$ such that $W'(1)$ is finite and we have a quasi-embedding $j : TW \to T'W$ between the initial Kruskal fixed points. For a partial order $X$, we set

$$W'(X) := \{\bullet, +\} \cup \{(x, y, \sigma) \in X \times X \times W(X) \mid x \neq y\}. $$

Here $\bullet$ and $+$ represent new elements that are incomparable with any others, while we have

$$(x, y, \sigma) \leq_{W'(X)} (x', y', \sigma') \iff x \leq x'$ and $y \leq y'$ and $\sigma \leq_{W(X)} \sigma'. $$


To get a functor, we recall that any quasi-embedding \( f : X \to Y \) is injective. This allows us to put
\[
W'(f)(\bullet) := \bullet, \quad W'(f)(+) := +, \quad W'(f)(x, y, \sigma) := (f(x), f(y), W(f)(\sigma)).
\]
Let us also define functions \( \text{supp}'_X : W'(X) \to [X]^{<\omega} \) by setting
\[
\text{supp}'_X(\bullet) := \emptyset, \quad \text{supp}'_X(+) := \emptyset, \quad \text{supp}'_X((x, y, \sigma)) := \{x, y\} \cup \text{supp}_X(\sigma),
\]
where \( \text{supp}_X : W(X) \to [X]^{<\omega} \) is the support function that comes with the WPO-dilator \( W \). One readily checks that this turns \( W' \) into a normal PO-dilator. Crucially, \( \emptyset \) ensures that the well partial orders are closed under binary products (see the paragraph before proposition 2.2), which implies that \( W' \) is a WPO-dilator. Let us also note that \( W'(1) = (\bullet, +) \) is finite.

To construct our quasi-embedding \( j : TW \to TW' \), we first note that \( TW' \) contains distinct elements \( \bar{\star} := \circ(\emptyset, \bullet) \) and \( \bar{\top} := \circ(\emptyset, +) \). As explained in §1, each \( \circ(a, \sigma) \in TW \) gives rise to an element \( W(\text{en}^{TW}_a(\sigma)) \in W(TW) \). Let us also recall that the Kruskal fixed point of \( W' \) comes with a bijection \( \kappa' : W(TW') \to TW' \). As in the proof of theorem 2.5, we can use recursion over terms to define
\[
j(\circ(a, \sigma)) := \kappa'(\langle \bar{\star}, \bar{\top}, W(j \circ \text{en}^{TW}_a(\sigma)) \rangle).
\]
To be more precise, we note that the finite function \( j \circ \text{en}^{TW}_a \) is available in the recursion step. In the following inductive proof, the induction hypothesis will imply that \( j \circ \text{en}^{TW}_a \) is a quasi-embedding, which is needed to ensure that \( W(j \circ \text{en}^{TW}_a) \) is defined. To disentangle the recursion and the induction, one can assign a default value such as \( j(\circ(a, \sigma)) := \bar{\star} \) for the hypothetical case that \( j \circ \text{en}^{TW}_a \) is not a quasi-embedding (which is decidable). As preparation for our induction, we define a length function \( l : TW \to \mathbb{N} \) with \( l(\circ(a, \sigma)) := 1 + \sum_{\tau \in \mathbb{N}} l(\tau) \), by another recursion over terms. To check that \( j(s) \leq TW j(t) \) entails \( s \leq TW t \), one argues by induction on \( l(s) + l(t) \). For \( s = \circ(a, \sigma) \) and \( t = \circ(b, \tau) \), the induction hypothesis ensures that \( j \) is a quasi-embedding on \( a \cup b \), due to our choice of length function. Therefore, \( W(j \circ \text{en}^{TW}_a) \), \( W(j \circ \text{en}^{TW}_b) \) and \( W(j \circ \text{en}^{TW}_{a \cup b}) \) are well defined. We assume \( j(s) \leq TW j(t) \) and proceed by case distinction (for this, recall the order properties of \( \kappa \) and \( \kappa' \) from definition 1.2). In the first case, we successively get
\[
\langle \bar{\star}, \bar{\top}, W(j \circ \text{en}^{TW}_a(\sigma)) \rangle \leq_{W(TW')} \langle \bar{\star}, \bar{\top}, W(j \circ \text{en}^{TW}_b(\tau)) \rangle,
\]
\[
W(j \circ \text{en}^{TW}_a(\sigma)) \leq_{W(TW')} W(j \circ \text{en}^{TW}_b(\tau)),
\]
\[
W(j \circ \text{en}^{TW}_a(\sigma)) \leq_{W(TW')} W(\text{en}^{TW}_{a \cup b}(\sigma)) \leq_{W(TW')} W(j \circ \text{en}^{TW}_{a \cup b}(\tau)),
\]
\[
W(\text{en}^{TW}_{a \cup b}(\sigma)) \leq_{W(TW')} W(\text{en}^{TW}_{a \cup b}(\tau)),
\]
\[
W(\text{en}^{TW}_a(\sigma)) \leq_{W(TW')} W(\text{en}^{TW}_b(\tau)),
\]
where \( \text{en}^{TW}_{a \cup b} \) is the canonical embedding from \( a \cup b \) to \( TW \). In the other case, we have
\[
j(s) \leq TW r \quad \text{for some } r \in \text{supp}'_{TW} \left( \langle \bar{\star}, \bar{\top}, W(j \circ \text{en}^{TW}_a(\sigma)) \rangle \right) = \langle \bar{\star}, \bar{\top} \rangle \cup \langle j \rangle^{<\omega}(b).
\]
Assume that \( j(s) \) is less or equal to \( \bar{\star} \) (or \( \bar{\top} \)). By using the definition of \( \kappa' \), we consider the following cases: In the first case, we have \( \langle \bar{\star}, \bar{\top}, W(j \circ \text{en}^{TW}_a(\sigma)) \rangle \leq_{W(TW')} W(\text{en}^{TW}_a(\sigma)) = \bullet \). This cannot be by definition of \( W' \). In the other case, we have that \( j(s) \) lies below an element from the support of \( W(\text{en}^{TW}_a(\sigma)) = \bullet \), which cannot be since it is empty. We conclude \( j(s) \leq TW j(x) \) for some \( x \in b \). By induction hypothesis, this yields \( s \leq TW x \) and, finally, our claim \( s \leq TW t \).

According to theorem 2.5, the uniform Kruskal theorem for finite monotone PO-dilators entails the original Kruskal theorem for trees with bounded branchings, provably in \( \text{RCA}_0 \) (see corollary 2.6 for the case with arbitrary branchings). Over a sufficient base theory, the finite monotone PO-dilators coincide with the finite WPO-dilators, by propositions 2.1 and 2.2. The following implies that the resulting versions of the uniform Kruskal theorem differ over \( \text{RCA}_0 \).
Corollary 3.6. If $\varphi$ is a $\Sigma^1_2$-theorem of $\mathcal{CA}_0$, then $\mathcal{RCA}_0 + \varphi$ does not show that the uniform Kruskal theorem for finite WPO-dilators entails the original Kruskal theorem for binary trees, i.e. the statement that $(T_3, \preceq)$ is a well partial order (cf. definitions 2.3 and 2.4).

Proof. As shown by Friedman, the original Kruskal theorem for finite trees with arbitrary branchings is unprovable in predicative theories [4]. The version for binary trees is unprovable in the theory $\mathcal{CA}_0$. In order to see this, one should first recall that $\mathcal{CA}_0$ is conservative over Peano arithmetic (see e.g. [25], theorem III.1.16). The latter cannot prove the well foundedness of a certain term system $\varepsilon_0$ that represents its proof theoretic ordinal, due to classical work of Gentzen [26] (see, e.g., [27] for a modern presentation). Finally, the well foundedness of $\varepsilon_0$ follows from the original Kruskal theorem for binary trees, by unpublished work of de Jongh (see [28] for the attribution and, e.g., [20] for a detailed proof). Now if $W$ is a finite P0-dilator, then $W(1)$ is finite. This holds because any $\sigma \in W(1)$ has a normal form $\sigma = W(\varepsilon_0)(\sigma_0)$ with $\sigma \leq 1$ and $(|\sigma|, \sigma_0) \in \text{Tr}(W)$, as explained in §1. Hence, the uniform Kruskal theorem for finite WPO-dilators is entailed by statement $\psi$ from the previous theorem. So if the corollary was false, $\mathcal{RCA}_0 + \psi$ would prove that $\varphi$ implies the original Kruskal theorem for binary trees. This implication is $\Pi^1_2$, so that the previous theorem would make it provable in $\mathcal{CA}_0$. As the latter proves $\varphi$, it would thus prove the binary Kruskal theorem, which we have seen to be false. \hfill \blacksquare

In the following, we show how to improve part (a) of theorem 3.5. To see that Tennenbaum’s result about $\mathcal{ADS}$ is available over $\mathcal{RCA}_0$, we combine some results from the literature. By $\varphi^X$ we denote the formula that results from $\varphi$ when all second-order quantifiers are restricted to the sets that are computable relative to $X \subseteq \mathbb{N}$.

**Lemma 3.7.** We have $\neg \text{ADS}^X$ for all $X \subseteq \mathbb{N}$, provably in $\mathcal{RCA}_0$.

Proof. By work of Hirschfeldt & Shore [21] and their joint work with Slaman [29], we have $\mathcal{RCA}_0 \vdash \text{ADS} \rightarrow \text{HYP}$ and hence $\mathcal{RCA}_0 \vdash \text{ADS}^X \rightarrow \text{HYP}^X$. Here, $\text{HYP}$ is the hyperimmunity principle which asserts that, for any $Y \subseteq \mathbb{N}$, there is a function that is not dominated by any $Y$-computable function. It is clear that $\mathcal{RCA}_0$ proves $\neg \text{HYP}^X$, as needed. \hfill \blacksquare

The following result characterizes the fragment of lemma 3.3 that is needed for theorem 3.5. To explain the notation $\omega^{\omega^2}$ in part (iii), we recall that $\omega^X$ denotes the lexicographic order on the non-increasing finite sequences in a given linear order $X$.

**Lemma 3.8.** The following are equivalent over $\mathcal{RCA}_0$:

(i) the uniform Kruskal theorem for unary monotone P0-dilators $W$ such that $W(1)$ is finite,

(ii) Higman’s lemma for sequences with entries from a finite partial order,

(iii) the statement that $\omega^{\omega^2}$ is a well order.

Proof. Given that $W$ is unary, $W(1)$ is finite precisely if the same holds for the trace $\text{Tr}(W)$. Furthermore, Higman’s lemma for finite orders is equivalent to the statement that $(\text{Tr}^n, \preceq)$ is a well partial order for any $n \in \mathbb{N}$, in the notation from the previous section. Thus, the equivalence between (i) and (ii) holds by the proof of theorem 2.5 (also consider remark 2.7 with finite $Z$). The equivalence between (ii) and (iii) holds because $\omega^{\omega^2}$ is well founded if and only if $\omega^\omega$ is well founded for any $n \in \mathbb{N}$. The latter holds for $n$ if and only if Higman’s lemma holds for sequences with entries from a partial order with $n + 1$ elements, due to a result of de Jongh & Parikh [30] (see, e.g., [31,32] for the formalization in $\mathcal{RCA}_0$). \hfill \blacksquare

As promised, part (a) of the following theorem improves part (a) of theorem 3.5. Concerning part (b), we note that the infinite pigeonhole principle is a $\Pi^1_2$-statement, which one may include as premise of a $\Pi^1_2$-theorem as in (a).

**Theorem 3.9.** Let $\psi$ express the uniform Kruskal theorem for WPO-dilators $W$ such that $W(1)$ is finite.

(a) Any $\Pi^1_2$-theorem of $\mathcal{RCA}_0 + \psi$ is provable in $\mathcal{RCA}_0 + \text{‘}\omega^{\omega^2}\text{’ well founded’}$. 


(b) The well foundedness of $\omega^{\omega^\omega}$ is provable in the extension of $\text{RCA}_0$ by statement $\psi$ and the infinite pigeonhole principle.

Proof. Part (a) can be proved exactly like theorem 3.4, since lemmas 3.7 and 3.8 ensure that the relevant ingredients remain available, i.e., that $\neg\text{AD}^\omega$ holds in our $\omega$-submodel $N$ (this is ensured by the former lemma and the fact that $M$ models $\text{RCA}_0$) and that the uniform Kruskal theorem for unary monotone $\mathbb{P}0$-dilators $W$ with finite $W(1)$ is available to $N$ (this is provided by the latter lemma together with the fact that $\omega^{\omega^\omega}$ being well founded is a $\Pi^1_2$-formula and can, therefore, be inherited from the original model $M$). In order to reduce part (b) to the previous lemma, it suffices to observe that any unary monotone $\mathbb{P}0$-dilator $W$ with finite trace is a $\mathbb{W}0$-dilator, over $\text{RCA}_0$ with the infinite pigeonhole principle. To confirm this, we consider a well partial order $X$ and an infinite sequence $a_0, a_1, \ldots$ in $W(X)$. Each entry has a normal form $a_i = W(\Pi^1_i)$ with $(\langle a(i), a'_i \rangle) \in \text{Tr}(W)$, as explained in §1. Given that $W$ has finite trace, the infinite pigeonhole principle allows us to assume that $(\langle a(i), a'_i \rangle)$ is independent of $i$. Since $W$ is unary, the sets $a(i) \subseteq X$ contain at most one element. In the non-trivial case, they are non-empty, so that we may write $a(i) = \{x_j\}$ with $x_j \in X$. Because $X$ is a well partial order, we find $i < j$ with $x_i \leq x_j$. Monotonicity yields $a_i \leq W(x_i)a_j$, as needed to make $W$ a $\mathbb{W}0$-dilator.

To conclude, we mention one further variation of our dichotomy result.

Remark 3.10. Let $\psi'$ be the uniform Kruskal theorem for $\mathbb{W}0$-dilators $W$ such that $W(1)$ coincides with $W(0)$, i.e., such that $W(1) = \text{rng}(W(0))$ for the empty inclusion $\iota : 0 = \emptyset \rightarrow 1$. If $W$ satisfies this condition and is unary, then all elements of its trace have the form $\langle \emptyset, a \rangle$. The latter entails $TW \equiv W(0)$, so that $\text{RCA}_0$ alone proves the corresponding instance of the uniform Kruskal theorem. It follows that $\text{RCA}_0 + \psi'$ is a $\Pi^1_2$-conservative extension of $\text{RCA}_0$. Over $\text{RCA}_0 + \text{AC}$, however, statement $\psi'$ is still equivalent to $\Pi^1_2$-comprehension, by the proof of theorem 3.5 (since in that proof, we have $W'(0) = W'(1)$).

Data accessibility. This article has no additional data.

Authors’ contributions. A.F.: conceptualization, investigation, writing—original draft, writing—review and editing; P.U.: conceptualization, investigation, writing—original draft, writing—review and editing.

Both authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project no. 460597863.

Acknowledgements. A variant of theorem 2.5 had been proved in joint work with Michael Rathjen and Andreas Weiermann [33]. We are grateful that they have allowed us to adapt the result for the present paper. We also thank Gabriele Buriola and Peter Schuster, who pursue a related project and have pointed out the original formulation of a result due to G. Higman (cf. the paragraph before corollary 2.6). Furthermore, we thank Leszek Kołodziejczyk for helpful advice on the reverse mathematics of $\text{AD}^\omega$. Finally, we are grateful for the excellent suggestions made by the referees.

References

1. Kruskal J. 1960 Well-quasi-ordering, the tree theorem, and Vazsonyi’s conjecture. *Trans. Am. Math. Soc.* 95, 210–225. (doi:10.1090/S0002-9947-1960-0111704-1)
2. Rathjen M, Weiermann A. 1993 Proof-theoretic investigations on Kruskal’s theorem. *Ann. Pure Appl. Logic* 60, 49–88. (doi:10.1016/0168-0072(93)90192-G)
3. van der Meeren J. 2015 *Connecting the two worlds: well-partial-orders and ordinal notation systems*. PhD thesis. Ghent, Belgium: Ghent University.
4. Simpson S. 1985 Nonprovability of certain combinatorial properties of finite trees. In *Harvey Friedman’s research on the foundations of mathematics* (eds L Harrington, M Morley, A Sčedrov, S Simpson), vol. 117. Studies in Logic and the Foundations of Mathematics, pp. 87–117. Amsterdam, The Netherlands: North-Holland.
5. Simpson S 2009 *Subsystems of second order arithmetic*. Perspectives in Logic. Cambridge, UK: Cambridge University Press.

6. Freund A, Rathjen M, Weiermann A. 2022 Minimal bad sequences are necessary for a uniform Kruskal theorem. *Adv. Math.* **400**, 108265. (doi:10.1016/j.aim.2022.108265)

7. Girard JY. 1981 $\Pi^1_1$-logic, part 1: dilators. *Ann. Pure Appl. Logic* **21**, 75–219. (doi:10.1016/0003-4843(81)90016-4)

8. Girard JY. 1982 *Proof theory and logical complexity*, vol. 2. See http://girard.perso.math.cnrs.fr/Archives4.html (accessed on 21 November 2017), manuscript.

9. Freund A. 2020 From Kruskal’s theorem to Friedman’s gap condition. *Math. Struct. Comput. Sci.* **8**, 952–975. (doi:10.1017/S0960129520000298)

10. Lerman M, Solomon R, Towsner H. 2013 Separating principles below Ramsey’s theorem for pairs. *J. Math. Logic* **13**, 1350007. (doi:10.1142/S0219061313500074)

11. van der Meeren J, Rathjen M, Weiermann A. 2017 An order-theoretic characterization of the Howard-Bachmann-hierarchy. *Arch. Math. Logic* **56**, 79–118. (doi:10.1007/s00153-016-0515-6)

12. Aguilera JP, Pakhomov F, Weiermann A. 2022 Functorial fast-growing hierarchies. (http://arXiv.org/abs/2201.04536)

13. Cholak P, Marcone A, Solomon R. 2004 Reverse mathematics and the equivalence of definitions for well and better quasi-orders. *J. Symbolic Logic* **69**, 683–712. (doi:10.2178/jsl/1096901762)

14. Marcone A. 2020 The reverse mathematics of wqos and bqos. In *Well-quasi orders in computation, logic, language and reasoning* (eds P Schuster, M Seisenberger, A Weiermann), vol. 53. Trends in Logic (Studia Logica Library), pp. 189–219. Cham, Switzerland: Springer.

15. Towsner H. 2020 Constructing sequences one step at a time. *J. Math. Logic* **20**, article no. 2050017. (doi:10.1142/S0219061320500178)

16. Uftring P. 2023 Weak and strong versions of effective transfinite recursion. *Ann. Pure Appl. Logic* **174**, 103232. (doi:10.1016/j.apal.2022.103232)

17. Frittaion E, Marcone A. 2012 Linear extensions of partial orders and reverse mathematics. *Math. Logic Q.* **58**, 417–423. (doi:10.1002/malq.201200025)

18. Higman G. 1952 Ordering by divisibility in abstract algebras. *Proc. Lond. Math. Soc.* **3**, 326–336. (doi:10.1112/plms/s3-2.1.326)

19. Pouzet M. 1985 Applications of well quasi-ordering and better quasi-ordering. In *Graphs and Order. Proceedings of the NATO Advanced Study Institute, Banff 1984* (ed. I Rival), vol. 147. NATO ASI Series C: Mathematical and Physical Sciences, pp. 503–519. Dordrecht, The Netherlands: Reidel.

20. Hirst JL. 1994 Reverse mathematics and ordinal exponentiation. *Ann. Pure Appl. Logic* **66**, 1–18. (doi:10.1016/0168-0072(94)90076-0)

21. Hirschfeldt D, Shore R. 2007 Combinatorial principles weaker than Ramsey’s theorem for pairs. *J. Symbolic Logic* **72**, 171–206. (doi:10.2178/jsl/1174668391)

22. Girard JY. 1987 *Proof theory and logical complexity*, vol. 1. Studies in Proof Theory. Napoli, Italy: Bibliopolis.

23. Hirst JL. 1994 Reverse mathematics and ordinal exponentiation. *Ann. Pure Appl. Logic* **66**, 1–18. (doi:10.1016/0168-0072(94)90076-0)

24. Gentzen G. 1934 Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie. *Math. Annalen* **119**, 149–161. (doi:10.1007/BF01564760)

25. Buchholz W. 1991 Notation systems for infinitary derivations. *Arch. Math. Logic* **30**, 277–296. (doi:10.1007/BF01621472)

26. de Jongh D, Parikh R. 1977 Well-partial orderings and hierarchies. *Indagationes Math.* **80**, 195–207. (doi:10.1016/1385-7258(77)90067-1)
31. Schütte K, Simpson S. 1985 Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen. Archiv für mathematische Logik und Grundlagenforschung 25, 75–89. (doi:10.1007/BF02007558)

32. Hasegawa R. 1994 Well-ordering of algebras and Kruskal’s theorem. In Logic, language and computation (eds N Jones, M Hagiya, M Sato), vol. 792. Lecture Notes in Computer Science, pp. 133–172. Berlin, Germany: Springer.

33. Freund A, Rathjen M, Weiermann A. 2022 Kruskal’s tree theorem and continuous transformations of partial orders. Talk at CCC 2020 held at Faro (online), abstract available at http://cid.uni-trier.de/additional_media/CCC-2020/CCC_2020_paper_4.pdf (accessed on 27 June 2022).